

MATHEMATICAL MODELING  
IN SUPPORT OF OPTIMIZATION  
OF TREATMENT OF LEUKEMIA

(Mathematisch Modelleren voor Optimale Behandeling van Leukemie)

PROEFSCHRIFT

Ter verkrijging van de graad van doctor  
aan de Erasmus Universiteit Rotterdam  
op gezag van de rector magnificus  
Prof.dr. P.W.C. Akkermans M.Lit.  
en volgens besluit van het College van Dekanen.

De openbare verdediging zal plaatsvinden op  
woensdag 5 oktober 1994 om 15.45 uur

door

Frank Willem Schultz

geboren te Djakarta

## PROMOTIE-COMMISSIE

**Promotoren:** Prof.dr. A. Hagenbeek  
Prof.dr.ir. J.A. Mulder

**Overige leden:** Prof.dr. D.W. van Bekkum  
Prof.dr. J.J. Broerse

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Schultz, Frank Willem

Mathematical modeling in support of optimization of  
treatment of leukemia / Frank Willem Schultz. - [S.l. :  
s.n.] (Meppel : Krips Repro). - Ill.

Thesis Rotterdam. - With ref. - With summary in Dutch.

ISBN 90-9007409-0

NUGI 743

Subject headings: leukemia / mathematical modeling.

This publication is a thesis for a doctoral degree at the Erasmus University Rotterdam.  
The work described was performed at the former Radiobiological Institute TNO / Institute  
for Applied Radiobiology and Immunology TNO, Rijswijk, The Netherlands

Front cover by Paul Schultz, Den Haag

Printed by Krips Repro, Meppel, 1994

## Table of Contents

|   |           |
|---|-----------|
| Preface   | 1         |
| <b>1 General Introduction and Scope of the Thesis</b>                               | <b>3</b>  |
| 1.1 Leukemia  | 3         |
| 1.1.1 Normal Hemopoiesis  | 3         |
| 1.1.2 Frequency of Occurrence of Leukemia   | 3         |
| 1.1.3 Types of Leukemia   | 4         |
| 1.2 Cytostatic Drugs  | 4         |
| 1.3 Pharmacodynamics and Pharmacokinetics   | 6         |
| 1.4 Modeling and Simulation   | 7         |
| 1.5 Mathematical Modeling and Cancer Chemotherapy                                   | 9         |
| 1.6 Problem Definition  | 13        |
| 1.7 Mathematical Modeling and Simulation to Optimize Treatment                      | 15        |
| 1.8 Aim of the Thesis   | 21        |
| 1.9 Experimental Data   | 26        |
| 1.9.1 Laboratory Animals  | 26        |
| 1.9.2 Brown Norway Rat Leukemia Model (BNML)  | 26        |
| 1.9.3 Generation of Data  | 27        |
| 1.10 Outline of the Thesis  | 27        |
| 1.11 References   | 29        |
| <b>2 Cell Population Dynamics of Leukemia Growth in the Brown Norway Rat</b>        | <b>31</b> |
| 2.1 Basic Experiments and <i>in vivo</i> Growth                                     | 31        |
| 2.1.1 Basic Experiments   | 31        |
| 2.1.1.1 Homing and Lodging of BNML Cells in the Recipient Rat                       | 31        |
| 2.1.1.2 BNML Cell Numbers Transferred and Leukemia Induction                        | 31        |
| 2.1.1.3 BNML Cell Numbers Transferred and Median Survival Time of the Recipient Rat | 32        |
| 2.1.2 Detection and Quantification Methods  | 32        |
| 2.1.2.1 Organ Weights   | 32        |
| 2.1.2.2 Counting based on Morphology (Quantitative Cytology)                        | 34        |
| 2.1.2.3 Flow Cytometry  | 34        |
| 2.1.2.4 Clonogenic Leukemic Stem Cell Assay   | 34        |
| 2.1.2.5 Survival Time Bio-assay   | 36        |
| 2.1.3 <i>In vivo</i> Leukemia Growth  | 38        |
| 2.2 BNML Growth Perturbation by Cyclophosphamide                                    | 41        |
| 2.2.1 Materials and Methods   | 43        |
| 2.2.1.1 The Rat Leukemia Model  | 43        |
| 2.2.1.2 The Cytostatic Agent Cyclophosphamide (CFA)                                 | 43        |
| 2.2.1.3 Experiments   | 43        |
| 2.2.1.4 Data Analysis; Basic Growth Curves  | 44        |
| 2.2.1.5 Data Analysis; Model Evaluation   | 48        |
| 2.2.2 Results and Discussion  | 51        |

|            |  |           |
|------------|--|-----------|
| 2.2.2.1    | Single LCFU-S Dataset  | 51        |
| 2.2.2.2    | All LCFU-S Datasets Combined   | 52        |
| 2.2.2.3    | The Dose-Survival Bio-assay and Flow Cytometry Data                              | 55        |
| 2.2.2.4    | Unperturbed BNML Growth in Liver and Spleen                                      | 59        |
| 2.2.2.5    | In Conclusion  | 59        |
| <b>2.3</b> | <b>Cell Cycle Specificity of AMSA? (A Simulation Study)</b>                      | <b>60</b> |
| 2.3.1      | Methods  | 60        |
| 2.3.1.1    | The Rat Leukemia Model (BNML)  | 60        |
| 2.3.1.2    | Experimental Animals   | 60        |
| 2.3.1.3    | LCFU-S Assay   | 60        |
| 2.3.1.4    | Unperturbed Leukemia Growth  | 61        |
| 2.3.1.5    | AMSA Administration  | 61        |
| 2.3.1.6    | Log Cell Kill (LCK)  | 61        |
| 2.3.1.7    | Surviving Fraction   | 61        |
| 2.3.1.8    | Leukemia Regrowth  | 62        |
| 2.3.1.9    | Mathematical Model for Simulating Leukemia (Re)Growth and Leukemic Cell Kill     | 62        |
| 2.3.1.10   | Software   | 64        |
| 2.3.2      | Results  | 64        |
| 2.3.3      | Discussion   | 69        |
| <b>2.4</b> | <b>Drug Resistance</b>   | <b>71</b> |
| 2.4.1      | Summary of Relevant Literature   | 71        |
| 2.4.1.1    | Evidence for Development of Drug Resistance                                      | 71        |
| 2.4.1.2    | Classification of Different Types of Drug Resistance                             | 72        |
| 2.4.1.3    | Mutation to Drug Resistance  | 73        |
| 2.4.1.4    | Mechanisms of Drug Resistance  | 74        |
| 2.4.2      | Development of Drug Resistant BNML Cell Lines                                    | 75        |
| 2.4.2.1    | A Cyclophosphamide Resistant Subline   | 76        |
| 2.4.2.2    | Mechanism of Cyclophosphamide Resistance   | 77        |
| 2.4.3      | Modeling Cyclophosphamide Resistance: Simulation Studies                         | 78        |
| 2.4.3.1    | Laboratory Experiments   | 79        |
| 2.4.3.2    | Modeling and Simulation  | 84        |
| 2.4.3.3    | Validation   | 88        |
| 2.5        | References   | 90        |
| <b>3</b>   | <b>Cell Population Dynamics of Childhood T-Cell Acute Lymphoblastic Leukemia</b> | <b>97</b> |
| 3.1        | Methods and Materials  | 98        |
| 3.1.1      | Detection and Quantification of T-ALL Cells by Immunological Marker Analysis     | 98        |
| 3.1.2      | Patients   | 100       |
| 3.1.3      | Case Report for Patient CDJ  | 100       |

|          |   |            |
|----------|---|------------|
| 3.1.4    | Mathematical Analysis   | 101        |
| 3.1.4.1  | Inspection of Raw Data  | 101        |
| 3.1.4.2  | Log-linear Regression   | 102        |
| 3.1.4.3  | Simulations   | 102        |
| 3.1.4.4  | Curve fitting   | 102        |
| 3.2      | Results and Discussion  | 106        |
| 3.2.1    | PB Data of Patient CDJ  | 106        |
| 3.2.2    | PB Data of Other Patients   | 110        |
| 3.3      | Concluding Remarks  | 116        |
| 3.4      | References  | 120        |
| <b>4</b> | <b>Pharmacokinetics</b>   | <b>123</b> |
| 4.1      | Materials and Methods   | 126        |
| 4.1.1    | Experimental Data Acquisition   | 126        |
| 4.1.2    | Multicompartment Models   | 127        |
| 4.1.2.1  | Model Structures  | 127        |
| 4.1.2.2  | Small-scale Model   | 127        |
| 4.1.2.3  | Medium-scale Model  | 131        |
| 4.1.2.4  | Large-scale Model   | 131        |
| 4.1.2.5  | Assumptions on Drug Transport   | 131        |
| 4.1.2.6  | Assumptions on Metabolism   | 136        |
| 4.1.2.7  | Model Equations   | 136        |
| 4.1.2.8  | Solving the Model Equations   | 141        |
| 4.1.3    | Parameter Estimation  | 143        |
| 4.1.3.1  | ML Technique  | 145        |
| 4.1.3.2  | Minimization Routines   | 147        |
| 4.1.4    | Error Estimates   | 157        |
| 4.1.4.1  | Estimation of the Variance-Covariance Matrix $Q$  | 157        |
| 4.1.4.2  | Estimation of the Variance-Covariance Matrix $V_p$  | 157        |
| 4.1.4.3  | Estimation of the Variance-Covariance Matrix $V_x$  | 158        |
| 4.1.5    | Performance Criteria  | 158        |
| 4.1.6    | Computation   | 159        |
| 4.1.7    | Set-up of Test and Evaluation Runs  | 161        |
| 4.1.7.1  | Data-sensitivity of the Identification Procedure  | 161        |
| 4.1.7.2  | Performance of the Different Minimization Routines  | 163        |
| 4.1.7.3  | Identification of the System of DAU-DOL Pharmacokinetics  | 167        |
| 4.2      | Results and Discussion  | 167        |
| 4.2.1    | Experimental Data   | 167        |
| 4.2.2    | Data-Sensitivity  | 170        |
| 4.2.2.1  | Test of the Optimization Routine  | 170        |
| 4.2.2.2  | Experiments A through C; reduction of the number of data-points while all compartments are observed | 172        |
| 4.2.2.3  | Experiments D through H; datapoints in a few compartments only                                      | 176        |
| 4.2.2.4  | Experiments I and A; the influence of model errors  | 178        |
| 4.2.2.5  | Conclusions   | 178        |
| 4.2.3    | Comparison of the Minimization Routines   | 184        |

|         |  |     |
|---------|--|-----|
| 4.2.3.1 | MGN method   | 184 |
| 4.2.3.2 | FD method  | 186 |
| 4.2.3.3 | DDH method   | 187 |
| 4.2.3.4 | Comparison of model responses, MGN and DDH methods       | 189 |
| 4.2.3.5 | Conclusion   | 191 |
| 4.2.4   | Identification of the Daunomycin Pharmacokinetics        | 194 |
| 4.2.4.1 | Comparison of the Small-scale Multicompartment Models    | 195 |
| 4.2.4.2 | Comparison of Intermediate-scale Multicompartment Models | 196 |
| 4.2.4.3 | Comparison of the Large-scale Multicompartment Models    | 196 |
| 4.2.4.4 | Conclusions  | 202 |
| 4.3     | References   | 202 |

## **5 Analysis of DNA Histograms 205**

|         |   |     |
|---------|---|-----|
| 5.1     | Introduction to DNA Histogram Analysis  | 205 |
| 5.1.1   | Purpose of Flow Cytometric Data (DNA Histogram) Analysis  | 205 |
| 5.1.2   | Flow Cytometric Data Acquisition  | 206 |
| 5.1.3   | Methods of FCM Data Analysis  | 207 |
| 5.2     | The Multiharmonic Technique   | 213 |
| 5.2.1   | Construction of the Mathematical Model  | 213 |
| 5.2.1.1 | Creating the Theoretical Histogram  | 213 |
| 5.2.1.2 | Fourier Expansion   | 214 |
| 5.2.1.3 | Application of Fourier Expansion to the S Phase Distribution  | 216 |
| 5.2.1.4 | Modeling the Dispersive Effects   | 216 |
| 5.2.2   | Parameter Estimation  | 220 |
| 5.2.2.1 | Introduction  | 220 |
| 5.2.2.2 | Maximum Likelihood Optimization   | 222 |
| 5.2.2.3 | Computation   | 229 |
| 5.3     | Application of the Multiharmonic Analysis Technique   | 233 |
| 5.3.1   | Test of the Multiharmonic Analysis Method   | 233 |
| 5.3.1.1 | Test Histograms   | 233 |
| 5.3.1.2 | Results for FRAMA, BGWLL2 and SMH Histograms  | 234 |
| 5.3.1.3 | Results for the Simulated Histograms I through V  | 243 |
| 5.3.1.4 | Results for the Simulated Histograms BA35, BA27 and BA21  | 254 |
| 5.3.2   | Sequential Histograms   | 254 |
| 5.3.2.1 | Influence of Growth Factor Presence on the Cell Kinetics of B-lymphocytes   | 254 |
| 5.3.2.2 | Influence of Imposing a Temporary Block in Early S-Phase on the Kinetics of a Simulated Exponentially Growing Cell Population | 260 |
| 5.3.3   | Histograms from BN Rat Tissues, With and Without infiltrated BNML Cells   | 266 |
| 5.4     | Concluding Remarks  | 271 |
| 5.5     | References  | 273 |

|          |  |            |
|----------|--|------------|
| <b>6</b> | <b>Bone Marrow Transplantation and Risk of Leukemia Relapse</b>  | <b>275</b> |
| 6.1      | Introduction to Bone Marrow Transplantation  | 275        |
| 6.1.1    | Transplantation  | 275        |
| 6.1.2    | Bone Marrow, Blood Cells and Hemopoiesis   | 275        |
| 6.1.3    | Bone Marrow Transplantation Procedure  | 276        |
| 6.1.4    | Types of Bone Marrow Transplantation   | 276        |
| 6.1.5    | Problems in Bone Marrow Transplantation for Leukemia; Allogeneic: Graft-versus-Host Disease, Autologous: Reinoculation of Leukemic Cells   | 276        |
| 6.1.6    | Mathematical Modeling  | 277        |
| 6.2      | On the Quantitative Role of Graft-versus-Host Disease in Decreasing the Leukemia Relapse Rate after Allogeneic Bone Marrow Transplantation | 278        |
| 6.2.1    | Methods  | 279        |
| 6.2.1.1  | Treatment Models   | 279        |
| 6.2.1.2  | Theoretical Probability of Leukemia Cure   | 279        |
| 6.2.1.3  | Choice of Parameter Values   | 281        |
| 6.2.2    | Results  | 282        |
| 6.2.3    | Discussion   | 285        |
| 6.3      | Contributions to Leukemia Relapse of Residual Leukemic Cells in the Patient and in the Graft after Autologous Bone Marrow Transplantation  | 288        |
| 6.3.1    | Methods  | 289        |
| 6.3.1.1  | Experimental Animals   | 289        |
| 6.3.1.2  | The Brown Norway Rat Acute Myelocytic Leukemia (BNML)  | 289        |
| 6.3.1.3  | BNML Growth <i>in vivo</i> after Cellular Transfer   | 289        |
| 6.3.1.4  | <i>In vivo</i> Growth of a Drug-Treated BNML Cell Line   | 290        |
| 6.3.1.5  | Determination of the ED <sub>50</sub> value for BNML   | 290        |
| 6.3.2    | Results  | 291        |
| 6.3.3    | Discussion   | 292        |
| 6.3.3.1  | The ED <sub>50</sub> Model Concept   | 292        |
| 6.3.3.2  | Application to ABMT  | 293        |
| 6.3.3.3  | The Probability of the Graft's Contribution to Leukemia Relapse  | 294        |
| 6.4      | References   | 296        |
| <b>7</b> | <b>General discussion, new developments and concluding remarks</b>   | <b>301</b> |
| 7.1      | Mathematical Modeling for Leukemia Treatment   | 301        |
| 7.2      | Detection and Regrowth of Minimal Residual Disease   | 306        |
| 7.2.1    | Development of Experimental Detection Methods  | 306        |
| 7.2.1.1  | In the BNML Rat Leukemia Model   | 306        |

|                                  |  |            |
|----------------------------------|--|------------|
| 7.2.1.2                          | In Human Leukemias   | 307        |
| 7.2.2                            | Regrowth of the Residual Leukemic Cell Population                                      | 309        |
| 7.3                              | Pharmacokinetics and Pharmacodynamic Drug Effects                                      | 310        |
| 7.3.1                            | Analysis of Existing Pharmacokinetics Data (BN or BNML and Adriamycin or Daunomycin)   | 311        |
| 7.3.2                            | Drug Sensitivity Testing <i>in vitro</i>   | 312        |
| 7.3.3                            | Drug Resistance  | 313        |
| 7.4                              | Cell (Cycle) Kinetics and Cell Population Dynamics                                     | 314        |
| 7.4.1                            | Flow Cytometry   | 315        |
| 7.4.2                            | Improving the Therapeutic Index of Cell Cycle Specific Cytostatic Agents; the Z-method | 316        |
| 7.4.3                            | Visualization of Tumor Growth  | 319        |
| 7.5                              | References   | 320        |
| <b>Summary</b>                   |  | <b>325</b> |
| <b>Samenvatting</b>              |  | <b>331</b> |
| APPENDIX A                       |  | 337        |
| APPENDIX B                       |  | 338        |
| APPENDIX C                       |  | 342        |
| APPENDIX D                       |  | 345        |
| APPENDIX E                       |  | 348        |
| APPENDIX F                       |  | 353        |
| <b>Abbreviations and Symbols</b> |  | <b>355</b> |
| <b>Naschrift</b>                 |  | <b>365</b> |
| <b>Curriculum Vitae</b>          |  | <b>367</b> |



*aan mijn ouders*



## Preface

Preclinical and clinical data on different aspects of leukemia and its treatment have been accumulated at the former Radiobiological Institute TNO during the past decades. The variety of data should be organized, sorted and combined properly, to be used for optimization of therapy. Mathematical modeling is a good method to reveal the important variables and their relations.

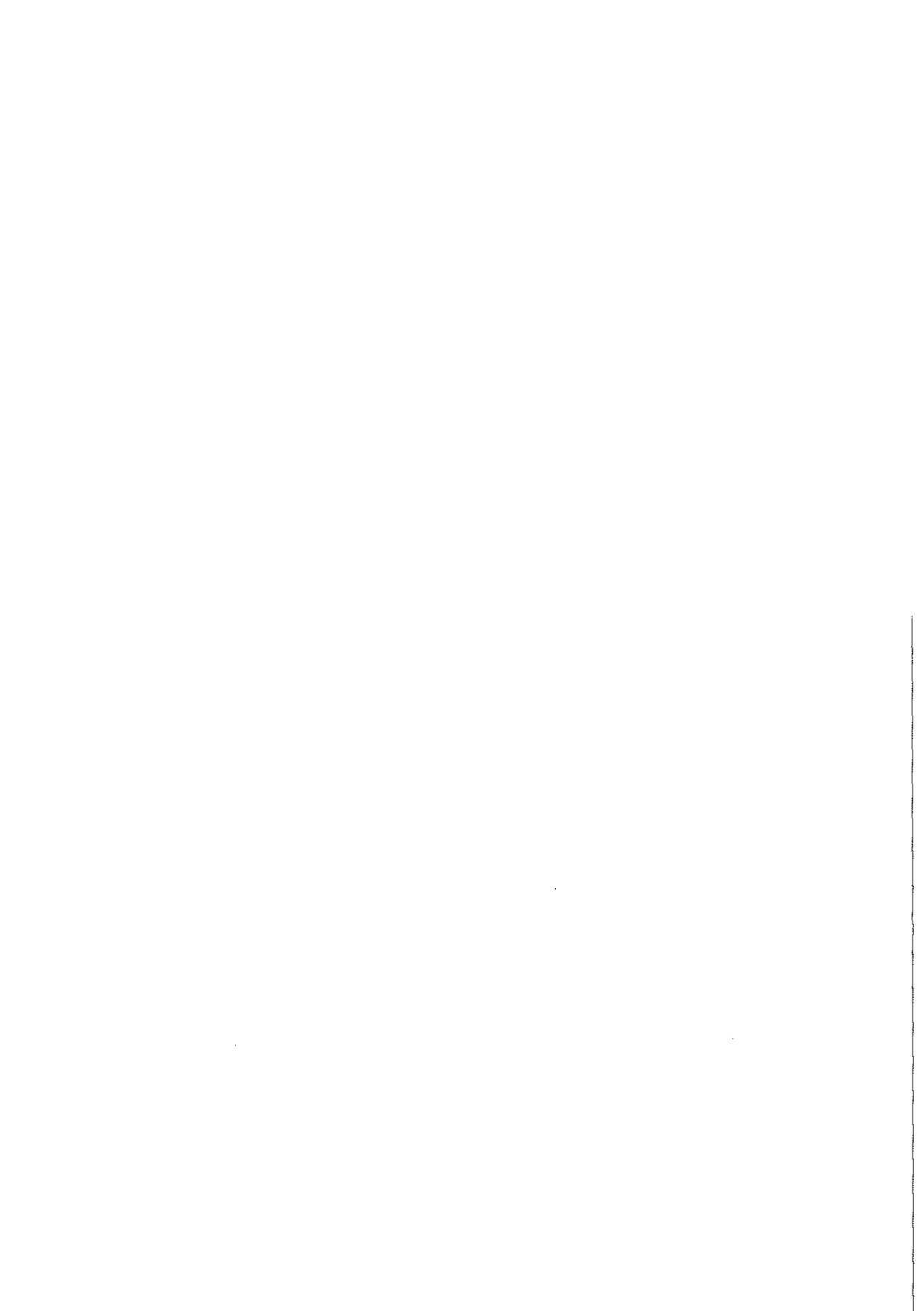
With this thesis archives are created that summarize (results of) scientific research efforts with respect to mathematical modeling and computer simulation of leukemia growth and chemotherapy, during the period 1981-1990, showing ways and giving recommendations for further investigations.

The investigations described in this thesis were supported by several grants of the 'Dutch Cancer Society' (Nederlandse Kankerbestrijding, 'Koningin Wilhelmina Fonds'), among which:

- Project RBI 80-1: Kinetics and treatment of minimal residual disease
- Project RBI 80-2: Mathematical modeling of drug effects in malignant cell proliferation
- Project RBI 82-5: Flow cytometer measurements of anthracyclines
- Project RBI 82-8: Late effects of supralethal chemoradiotherapy prior to bone marrow transplantation
- Project IKR 84-6: Computer simulation of leukemia growth with emphasis on minimal residual disease
- Project IKR 86-1: Quantitative analysis of chromosomal anomalies in malignant tumors employing dual beam flow cytometry
- Project IKR 87-12: Influence of tumor load on the pharmacokinetics and efficacy of anticancer drug treatment in rodent models for leukemias and solid tumors
- Project RRTI 88-8: Research and treatment of drug resistance in patients with solid tumors, leukemias and non-Hodgkin's lymphomas
- Project IKR 89-3: Molecular basis of cyclophosphamide resistance in neoplastic cells

Contributions by the Dutch Ministry of Welfare, Public Health and Cultural Affairs and by the Technical University Delft are acknowledged.

Financial support for printing this thesis was obtained from TNO-ME (Instituut voor Milieu- en Energietechnologie, Apeldoorn), the 'Dutch Cancer Society' (Nederlandse Kankerbestrijding, Amsterdam) and the Erasmus University Rotterdam.



# Chapter 1

## General Introduction and Scope of the Thesis

### 1.1 LEUKEMIA

Leukemia is a malignant disease of the hemopoietic tissues. This form of cancer causes a continuous large overproduction of malfunctioning immature blood cells, at the expense of the production of normal blood cells. Other blood cell cancers are the lymphomas and multiple myeloma, related to uncontrolled growth of, respectively, cells that make up the lymphatic system and of plasma cells in the bone marrow.

#### 1.1.1 Normal Hemopoiesis

Functional blood cells have a limited lifetime. They are destroyed when worn out. They also may get lost in disastrous events (e.g., a bleeding wound). Therefore, a continuous process of replacement is required, of which the regulation mechanisms now are getting better understood.

Normal blood cell production mainly takes place in the red bone marrow. According to the current concept, hemopoiesis starts with the pluripotent stem cell (PSC). This cell type is maintained by proliferation, i.e., cell division (mitosis) generates new PSCs. Some daughter cells are no longer PSCs but appear to have changed (differentiation). They have become committed stem cells (CSCs), i.e., the precursor cells of one line of specialized blood cells. Five such lines can be distinguished, each eventually yielding different functional end cells: erythrocytes (red blood cells), lymphocytes, granulocytes, monocytes (white blood cells), and thrombocytes. At each division CSCs and their daughter cells evolve to a more mature stage until, finally, the end cells result. They then have lost the ability of cell division. Migration out of the bone marrow takes place (e.g., to the lymph nodes for lymphocyte maturation).

Blood cells of different lines and in different stages of maturation can be recognized by morphology or other cytological and immunophenotypical characteristics. The same holds for the malignant counterparts of the normal hemopoietic cell types, although usually specific properties distinguishing the cell populations (normal—malignant) are missing.

#### 1.1.2 Frequency of Occurrence of Leukemia

In the year 1988 death from leukemia amounted to 2.8% of the mortality caused by any form of cancer, and to 0.8% of all deaths in The Netherlands (in numbers, per 100,000 inhabitants there were 7 deaths for leukemia, 239 for all

cancers and 839 for all causes, respectively [CBS, 1991]). Compared to other forms of cancer—e.g., cancer of the respiratory tract killed 59 out of every 100,000 inhabitants—leukemia must be regarded as a less frequent, but by far not insignificant cause of death. Death due to traffic accidents, for example, struck 9 out of 100,000 inhabitants. Therefore, all effort put into finding ways to cure this disease is easily justified.

### 1.1.3 Types of Leukemia

Van Dongen et al. [1988] have presented (hypothetical) schemes of the differentiation of hemopoietic cells. The schemes can be used to classify their malignant counterparts according to the stage of maturation where they turned malignant, i.e., started proliferation in an uncontrolled way without further maturation, thus overrunning and suppressing the functioning of the normal hemopoietic system. Two main types of leukemia can be distinguished: lymphocytic and myelocytic leukemia. *Lymphocytic* leukemia concerns the lymphoid paths of differentiation and comprises: acute undifferentiated leukemia, acute lymphoblastic leukemia (null, common, pre-B-cell, B-cell, immature T-cell, common thymocytic T-cell and mature T-cell), chronic lymphocytic leukemia (B-cell, T-cell), prolymphocytic leukemia (B- and T-cell) and hairy cell leukemia. *Myelocytic* leukemia concerns the myeloid paths, comprising: acute undifferentiated leukemia, acute myeloid leukemia (several subclasses according to the FAB classification [Bennett et al., 1985], among which progranulocytic, myelomonocytic, monocytic, erythroleukemia and megakaryocytic), chronic myeloid leukemia.

Many leukemias may occur in two forms, i.e., acute and chronic disease. In acute leukemia the immature hemopoietic cells are involved. Cells of the more mature differentiation stages are associated with the chronic form, which is characterized by an oscillating severity of illness. Acute leukemia often occurs early in life, chronic myeloid leukemia during mid-life, and chronic lymphocytic leukemia is mostly a geriatric disease.

For therapy, various cytostatic agents (most often as combinations) are applied. Especially for younger patients, high-dose chemo-/radiotherapy followed by bone marrow transplantation is a recent life-saving treatment.

In this thesis only acute myelocytic leukemia (AML) and childhood acute T-cell lymphoblastic leukemia (T-ALL) will be considered.

## 1.2 CYTOSTATIC DRUGS

Certain chemical compounds are able to kill cells directly or to interfere with cellular maturation processes and inhibit or prevent mitosis. These compounds are called cytostatic agents. Not only tumor cells are affected, but normal tissue cells as well. Loss of hair, damage to the intestinal tract, neurotoxicity and/or

suppression of hemopoiesis often result as side effect of the treatment. Therefore, in general, certain dose limits should not be exceeded to avoid unacceptable damage to normal tissues. Also, a therapy course may be interrupted to provide the opportunity for normal tissue cells to repair (sublethal) damage. Obviously, tumor cells then 'profit' as well from this rest period, but usually normal tissues recover faster than tumor tissue.

Cytostatic drugs that are commonly used for treatment of leukemia can be classified as follows [Perry, 1992]:

A) Agents that damage the DNA template

- 1) *Alkylating Agents* form electrophilic carbonium ions that alkylate (covalent bonds) nucleophilic groups (e.g., in guanine of DNA), causing cross linking and abnormal base pairing, thus interfering with the DNA replication function. Also, reactions with sulfhydryl, phosphate or amine groups result in multiple lesions in both dividing and non-dividing cells. Subclasses are
  - a) *Nitrogen mustards*, e.g., cyclophosphamide (cytoxan), chlorambucil (leukeran), melphalan and mechlorethamine
  - b) *Nitrosoureas*, e.g., carmustine (BCNU) and semustine (methyl/CCNU)
  - c) *Others*, e.g., triethylenethiophosphoramidate (Thio-TEPA), busulfan (myeleran), hexamethylmelamine, dacarbazine and mitomycin C

2) Agents that cause double-strand cleavage via topoisomerase II

- a) *Antibiotics*, e.g., doxorubicin (adriamycin), daunorubicin (daunomycin), mitoxantrone, idarubicin, epirubicin and amsacrine
  - b) *Podophylotoxins*, e.g., etoposide and teniposide
- B) *Antimetabolites* are compounds which inhibit protein or DNA synthesis, or the formation of various enzymes necessary for a cell's normal metabolic processes. Subclasses are
  - a) *Dihydrofolate reductase*, e.g., methotrexate (amethopterin)
  - b) *DNA polymerase*, e.g., cytosine arabinoside (cytarabin)
  - c) *Phosphoribosylpyrophosphate aminotransferase*, e.g., 6-mercaptopurine (purinethol) and 6-thioguanine
  - d) *Ribonucleotide reductase*, e.g., hydroxyurea (hydra)
  - e) *Adenosine deaminase*, e.g., deoxycoformycin (pentostatin)

C) *Spindel Poisons*. These are the *Vinca alkaloids*, e.g., vinblastine (velban), vincristine (oncovin) and vindesin

D) *Hormonal and Antihormonal Agents*

Various classes of hormones and antihormones have shown the ability to cause regression of malignant cell populations. A well-known example of the class of *Adrenocorticosteroids*, useful against leukemia, is prednisone

E) *Biological Response Modifiers*, e.g., the interferons. Retinoic acid derivatives, now clinically used against acute promyelocytic leukemia, may be classified in this category. They induce the leukemic cells to relinquish their malignant phenotype and enter a program of normal cellular differentiation

and death [Warrell et al., 1993; Cline 1994]

F) *Miscellaneous Compounds*, e.g., l-asparaginase.

Effects of only a few drugs will be discussed in this thesis.

### 1.3 PHARMACODYNAMICS AND PHARMACOKINETICS

Pharmacodynamics is the study of the time course and the intensity of biological responses arising from exposure to or treatment with particular chemicals [e.g., Conolly and Andersen, 1991]. Biological responses manifest both as beneficial, therapeutic effects (drugs) and as noxious, deleterious effects (toxic chemicals). To be able to describe quantitatively the relationship between the exposure of an organism and the time course of the response, a pharmacodynamic (PD) model must be used. Such a model may—in order of increasing sophistication— A) be of *correlational* nature (e.g., a correlation of percentages surviving bone marrow cells observed at different time points as biological effect and the blood concentration of a certain cytotoxic drug at the observation times); B) consist of *empirical* equations, whose parameter values are obtained by fitting to observed data (in an empirical model a mathematical structure is selected to be consistent with the observed data, but it need not necessarily be a precise description of the underlying physical processes); and C) be *biology-based* (BB), i.e., if knowledge exists about the biology of the test species it should be possible to describe, in a physically correct way, 1) the concentration—time course of an administered chemical at a site of interest; 2) how the chemical interacts with exposed tissues; and 3) how these tissues respond to this interaction.

Item 1) can be seen as a separate subject of study, which is known as the field of pharmacokinetics (PK). In other words, a PK model is part of a BB-PD model. While the BB-PD model describes (the time-dependent) relation between blood concentration of a chemical and a biological effect, the PK model gives the relation between the chemical's concentrations in blood and in the tissue of interest. Schematically,

PK = exposure  $\Rightarrow$  blood concentration  $\Rightarrow$  tissue concentration

PD = exposure  $\Rightarrow$  blood concentration  $\Rightarrow$  tissue concentration  $\Rightarrow$

$\Rightarrow$  interaction, chemical—tissue  $\Rightarrow$  tissue response  $\Rightarrow$  biological effect

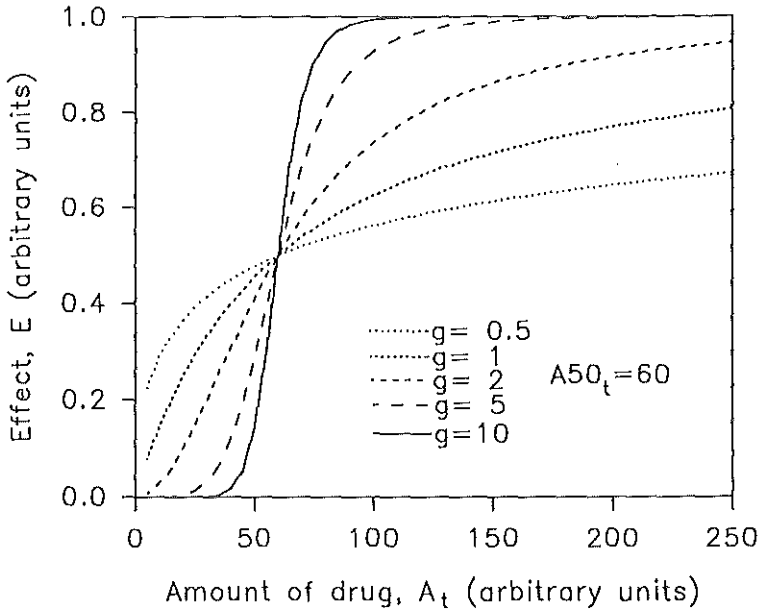
A pharmacokinetic model often is multicompartamental, each compartment corresponding with a tissue region or body fluid, and consists of polyexponential equations that describe passive diffusion processes. For example,

$$\frac{dA_t(T)}{dT} = k_{t,b} \cdot A_b(T) - k_{b,t} \cdot A_t(T), \quad (1.1)$$

where T denotes time;  $A_t$  and  $A_b$ , amount of drug in a tissue (t) and in the



Fig. 1.1 General curves that describe a relation between a drug effect in a tissue and the amount of drug in that tissue (Eq. 1.2)



blood (b), respectively;  $k_{t,b}$  and  $k_{b,t}$ , transfer rate constant, to t from b and in the opposite direction, respectively.

A biological effect is most often described with an empirical function, which shows (log)linear behavior at low concentrations and saturation behavior at high concentrations, e.g.,

$$E(T) = \frac{A_t(T)^g}{A_t(T)^g + A50_t^g}, \quad (1.2)$$

where  $A50_t$  denotes the amount of drug required for 50% effect in t; and E, effect, which may be the death of the malignant cells. Exponent g is a numerical constant. Its value determines the shape of the curve (Fig. 1.1). In this thesis 'pharmaco' modeling is limited to pharmacokinetics.

#### 1.4 MODELING AND SIMULATION

A general definition of a system is an efficiently arranged set of related objects and their components. To deal with its complexity, the world around us is sim-

plified by partitioning it into systems, subsystems, subsystems, etc. Thus broken down, the (independent) parts can be understood separately, often with the help of models that are further simplifications and idealizations of reality. Mathematical models are representations of reality, describing it in terms of mathematical formulae. Static and dynamic models can be distinguished [e.g., Rice, 1983]. Static models are a collection of equations, formulae, definitions, tables, relationships and data that describe a situation or phenomenon, presumably with sufficient completeness. Dynamic models are a collection of the same objects that describe how a situation changes from one state to the next one. An advantage is that in this form the model becomes explicit and can (relatively easily) be implemented on a (digital) computer to produce answers from the formulae that are a function of data, parameters and variables.

Next, simulations can be performed, i.e., subsequent states of the model as function of time can be calculated. Predictions of the real-life behavior of the system thus are acquired. Instead of a system, a process may be considered as well. A process is defined here as a sequence of actions within or between systems.

Simulation (including modeling) serves two main purposes, *process description* (including identification) and *process control*. Both correspond with a more or less systematic way of generating knowledge in contrast to the empirical methods of trial and error.

*Process description* yields insight into the process in an efficient way by extracting relations from a large quantity of diverse information. In particular, simulation is useful when no sensors are available to observe essential variables directly, as often happens with biomedical systems or processes. Through modeling, assorted experimental observations can be correlated and interpreted better than with the empirical approach, by explaining phenomena from the viewpoint of the underlying mechanisms that are incorporated in the model. Simulation techniques comprise the following components [e.g., Avula, 1987]: *modeling*, i.e., arranging the knowledge about the process in a model; performing *calculations* with the model, e.g., a systematic series of input/output calculations enabling the reconstruction of (the time history of) the state of the studied system; *comparing* the model predictions—simultaneously adjusting model parameters, i.e., characteristic constants in the model equations, to their optimal values—with actual observations to assess the probability of the assumptions used when building the model. If various models can be conceived of, this approach enables discrimination between them and selection of the most appropriate one. There is a systematic validation of theories based on interactions between results of simulation and measured results, leading to the (re)-drafting of work hypotheses. The model is revised and refined until the simulation results no longer conflict with any observed natural phenomenon, unless discrepancies can be explained and their reasons understood. A qualitative and

quantitative description of the process results. Through quantification of the model parameters the (relative) sensitivity of the process for certain input stimuli is obtained.

Based on the possibility to predict the behavior of the now identified process, even beyond the observed range, *process control* deals with finding those input stimuli that will result in some desired process output as function of time. Thus, simulation is used to learn how to employ and control a process in the most favorable way.

Computer simulation techniques can very well be applied in the field of biomedicine as a tool to direct experimental research and synthesize results of various experiments into general concepts. Knowledge and understanding of the behavior of biomedical systems often is difficult to obtain through measurements only. The implications of work hypotheses—expressed in the model equations that also account for established facts—can be tested through simulations with the computer model. Comparing the results to experimental observations yields the probability of the various assumptions.

The other way around, simulation results may lead to the formulation of new hypotheses about the mechanisms behind a process. Furthermore, based on simulation results it is possible to determine what additional experiments must be conducted to verify the chosen model. Model validation, i.e., checking whether the model predictions do not conflict with any actual observations, is always required to establish the model's appropriateness and usefulness for the intended applications.

## 1.5 MATHEMATICAL MODELING AND CANCER CHEMOTHERAPY

Cancer is a disease characterized by disturbances in the normal development of somatic cells causing an uncontrolled sequence of cell divisions. The malignant cells invade and destroy healthy tissues. Three basic methods are used to cope with the various forms of cancer, i.e., removal (*surgery*), application of ionizing radiation (*radiotherapy*) and treatment with cytostatic drugs that suppress tumor growth (*chemotherapy*). The action of the drugs may be based on different principles (see the Section on Cytostatic Drugs). During the past thirty years chemotherapy has become increasingly successful. On the one hand because of the introduction of new, more effective cytostatic agents, on the other hand because of the improved administration schemes based on accumulated (empirical) knowledge from clinical experience and (pre)clinical research (experiments with laboratory animals). For example, for leukemia the percentage cure, in terms of five year disease free survival, has increased from 15 to 40-50, depending on the type of leukemia: acute, chronic, lymphocytic, myelocytic; and other factors, like the age of the patient and the treatment strategy employed. Still, many

problems remain to be solved, among which the treatment of minimal residual disease (MRD) and handling the development of drug resistance.

Relations and interactions between elements of scientific theories can be understood as system models [Brock and Schneider, 1984]. In a model the actual conditions are deliberately reduced to a limited number of important aspects. If the general elements (variables) are of a quantitative nature, i.e., if numerical values can be assigned to them, then mutual links and relationships can also be expressed quantitatively in the form of mathematical functions or equations. Thus, the system becomes a mathematical model.

Biological systems are inherently complex. In general, they are more complex than most technical systems [Garfinkel, 1984] and less observable. Furthermore, often only little theoretical knowledge is available concerning the even elementary processes in living organisms, though much experimental information is available and still rapidly being accumulated. Scattered detail information should be combined in comprehensive models. Thus, a considerable data reduction can be achieved, scientific knowledge is consolidated, and a general picture is presented in a formalized, distinct and clear way. The state of the art of computer modeling shows a shift from saving computations (trying to reduce the need for expensive computation time) to considerations of how to represent biology better, and pulling together knowledge embodying both heuristics and incomplete or even contradictory information. Another important function of modeling is the interaction with experiments, not only for interpretation of observed results, but to help design experiments, as well as to see to it that the experiments actually do determine what is needed to be measured.

For cancer chemotherapy [Brock and Schneider, 1984] the experiments with living tumor-bearing animals are the basis for both a) the evaluation of scientific theories (modeling: models are used for checking theories and they provide insight and new information, leading to new model concepts); and b) the inference to clinical conditions (e.g., induction therapy to remove the bulk of the malignant cells, maintenance therapy to consolidate the disease-free state, toxic side effects).

Continuing development and integration of rational biomathematical models based on principles already identified, and testing them for compatibility with much already available experimental and clinical data will lead to models that will help in planning more effective treatment regimens for cancers that are now classified as moderately or very refractory to chemotherapy [Skipper, 1986]. For fruitful progression it is important that both the biological concepts and the arithmetic to be used for modeling are equally good.

Variables to be incorporated in the models comprise tumor and treatment variables. To the former category belong a) the *initial composition* of the tumor with respect to sensitive, single and multiple drug resistant subpopulations; b)

the *mutation rates* from sensitive to single drug resistant cells, and subsequently from single to multiple drug resistant cells; and c) the *population doubling time* (which is assumed to be constant for all sublines, as well as time invariant, i.e., there is exponential growth). Treatment variables are d) the *log cell kill factor* per dose of each drug; and e) the *treatment combination*, i.e., the number of doses of each drug and the time interval between them. Skipper [1986] performed computer simulations with these variables and a model for two-drug combinations.

But also the tumor environment should be considered. Saiga et al. [1985] have derived a coefficient of cell variation to describe the heterogeneity of a malignant population, considering mutational transformations at different rates (cell sublines adapt better or worse to the tumor environment). Michelson and Slate [1987] modeled the influences of a drug therapy on the tumor environment as a reduced capability of supporting a tumor burden of certain size. Rosen [1986] introduced an element of competition between sensitive and resistant cells, beside time-dependent drug influence functions.

Acheraya and Sundareshan [1984] point out that systems theory principles can and should be used in the development of optimal drug administration strategies. Finding the minimum drug dosage that can be administered for killing the number of tumor cells present, while maintaining the proliferation of normal cells at a safe level, constitutes an optimization problem. This approach differs from the trial and error simulation runs as performed by Skipper [1986].

While Acheraya and Sundareshan [1987] do not consider resistance phenomena in their model, Birkhead et al. [1984, 1986] do. They allow repeated doses of a single cytostatic agent, as well as the presence and accumulation of drug resistance. The latter comprises two elements: that already present at diagnosis and that acquired in response to and during treatment (e.g., spontaneous mutation or host defense reactions). The log cell kill hypothesis is assumed, as well as exponential growth and constant treatment intervals, though possibilities for model extensions are indicated. This model results in expressions for quantities such as the fractional tumor reduction due to each drug dose, the minimum tumor size achieved by a treatment, the changing composition of a tumor, etc. It is useful with respect to evaluating the influences of several *a priori* assumed variables and exploring the consequences of the hypotheses on which the clinician builds his treatment strategies. The likely results of his choices can be studied more fundamentally, replacing decisions on purely empirical basis.

Elements from the theories advocated by Skipper [1986] are incorporated in the models by Coldman and Goldie [1985, 1986a,b] and Goldie et al. [1979, 1982, 1983, 1985, 1986], who illustrated in their papers the use of mathematical modeling in relation to the development of chemotherapy strategies. They assumed a spontaneous somatic mutation rate from drug sensitivity to drug resistance and related it to the drug response of tumors as well as to the expectation

of cure. Allowing the emergence of cells that are resistant to multiple drugs, which in fact result in incurability, they could give a rationale for the application of alternating non-cross resistant chemotherapy as the most effective, risk minimizing strategy.

The somatic mutation model predicts a specific quantitative relationship between tumor size and probability of cure. The latter is related to both tumor size and mutation rate to resistance. The most effective way to utilize two equivalent (i.e., same killing capacity and same rate of mutation) drugs—not concurrently, e.g., because of toxicity—is in an alternating fashion.

In a later model Goldie and Goldman [1985] found an explanation for the fact that advanced stage tumors (slow growers) are less curable than smaller tumors (fast growers). Allowing more cell loss to account for slow growth, it is obvious that more replication cycles are necessary to reach a certain tumor size. By consequence, a slowly growing tumor will contain more resistant cells, because of the higher probability that mutations have occurred, than a fast growing tumor of equal size.

To evaluate the link between the clinical situation and mathematical modeling Dembo [1984], for example, looked into the implications of the somatic mutation model by Goldie et al. with respect to the management of ovarian cancer. As this model explains clinical observations (e.g., better response to first than to later course chemotherapy) better than the constant log cell kill model, the former's recommendations of early timing of chemotherapy and applying alternating non-cross resistant combination chemotherapy should be considered seriously.

Hokanson et al. [1986] presented a computer based model that simulates the characteristic features of the clinical time course of human myeloma (Kahler's disease, a neoplastic disorder of plasma cells). In this model therapy resistance is caused by kinetic differences between myeloma cells. Faster cycling cells are more sensitive. If the duration of the cycle time is a property with a high degree of heredity, the model is compatible with clinical results for various therapy schemes (pulsed intermittent or low dose continuous administration).

The purpose of mathematical modeling [Goldie and Coldman, 1986] is to provide deeper insight into natural phenomena and through this insight to make accurate predictions about the behavior of such phenomena. Verification of these predictions will produce further understanding of the processes being studied, which in turn may lead to modification and enhanced sophistication of the basic model, which was derived from a set of hypotheses. The experiment for model validation has to be a fair test of these hypotheses. This requires careful attention to the assumptions of the model when the experiment is planned. It is better to utilize models in which the assumptions and relationships are explicit and well-defined rather than ones that are ill-defined and purely phenomenological.

*Mathematical models incorporating descriptions of tumor growth kinetics, including drug resistance development, and the effects of cytostatic chemotherapy on established tumors and their microenvironment can be used to investigate the potential of hypothetical chemotherapy strategies and to identify general principles for successful treatment.*

## 1.6 PROBLEM DEFINITION

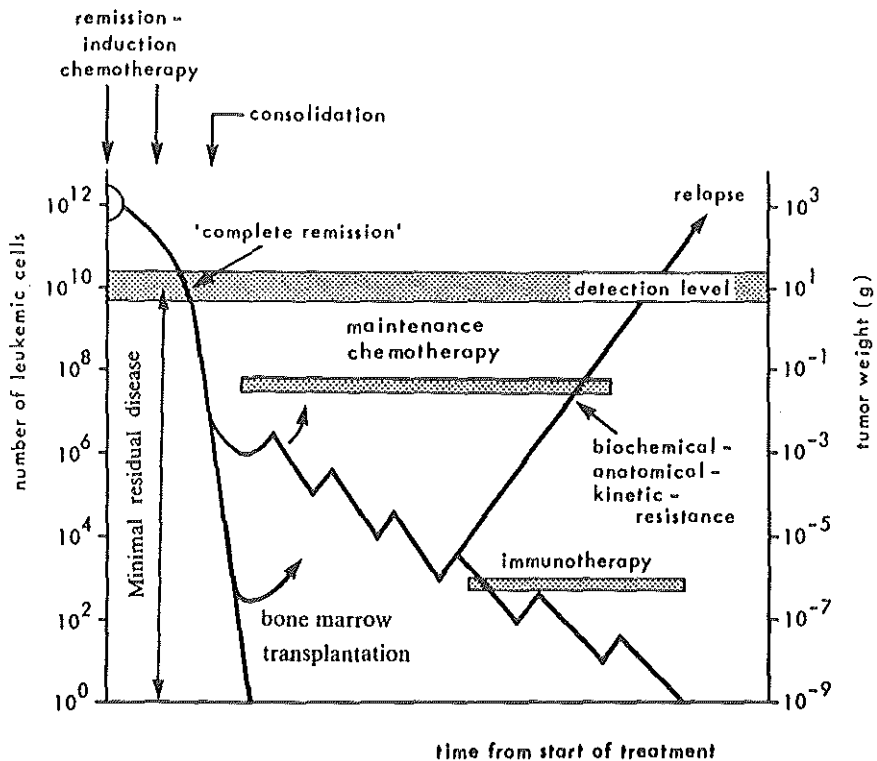
In acute leukemia (Fig. 1.2) at diagnosis the average patient (75 kg) already carries a leukemic cell load of approximately  $10^{12}$  cells (1 kg). With chemotherapy it is possible to induce a complete remission (CR) in the majority of cases, i.e., the leukemia disappears below the clinical detection level. The problem is, that with conventional cytological tools one can detect 1 abnormal cell in 20. One in 100 is the lowest (clinical) detection level. This means that the patient in remission, feeling well and apparently cured, may still invisibly carry a leukemic cell burden of up to  $10^{10}$  cells. This situation is called the state of minimal residual disease (MRD). Effort is put in shrinking the region of invisibility. But even under ideal circumstances, with advanced techniques like the polymerase chain reaction (PCR; see Chapter 7) the level of detection is at best 1 in  $10^5$  to  $10^6$  cells [Hagenbeek, 1992]. In spite of maintenance therapy courses, of which the efficacy inevitably is much a matter of guesswork due to the unmeasurable response during MRD, the residual leukemic cells often grow out, sooner or later causing a relapse of the disease when their numbers have once again reached detectable levels.

A second problem is the possible presence of drug resistant subpopulations of leukemic cells. As stated before they may be naturally present or may develop through natural mutation. They may develop as well under influence of the exposure to the drug (acquired resistance). From clinical experience it is known that identical chemotherapy courses, given sequentially, tend to become less effective. Therapy outcome can be improved by switching to other cytostatic drugs.

The clinician's problem therefore is, whether, which and how long maintenance chemotherapy must be given during the phase of complete remission. It would be a great help if the growth of the malignant cell population during MRD and the development of drug resistance could be monitored or predicted.

The success of cancer chemotherapy depends on a large number of variables whose values must be within certain ranges. They must operate in concert, but few variables can be controlled directly. When administration schemes are designed, decisions must be made about monotherapy or combination therapy and, for each drug, about the choice of 1: the individual *dosage level*; 2: the *interval* between doses; 3: the *number of administrations*; 4: the *duration* of the

Fig. 1.2 Definition of Minimal Residual Disease in Acute Leukemia



treatment ( $=2 \times 3$ ); 5: the *total dose* ( $=3 \times 1$ ); and 6: the *dose intensity* ( $=5/4$ ). The route of administration also plays a role, e.g., oral administration or intravenous injections (intermittent) or infusion (continuous). Because, these variables determine the processes of drug distribution in and elimination from the body, as well as the metabolic processes, that, in turn, determine how long and with what concentration the cytostatic agent can interact with the tumor. The problem is that knowledge about the influence of each (separate) variable is very poor. Yet, the ability to steer the concentration—time courses is very important. On the one hand the tumor must be exposed as intensely as possible to achieve the maximum therapeutic effectivity. It is desired that all tumor cells disappear quickly. On the other hand, just as do most drugs, cytostatic agents show adverse effects. Nausea and vomiting, loss of hair commonly occur, but also damage to the bone marrow, intestines and cardiac muscles. Therefore, the organs concerned should be exposed as little as possible.

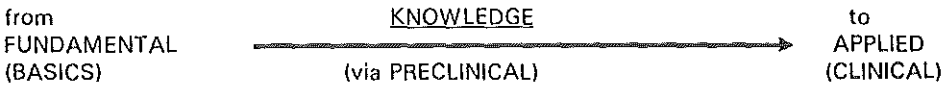
Through mathematical modeling and computer simulation the three discussed problems can be tackled.



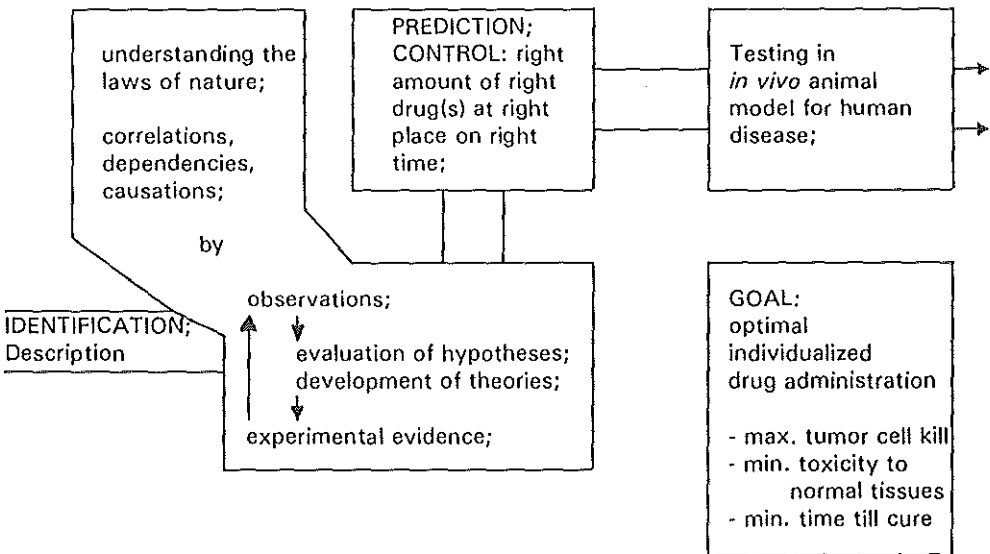
Fig. 1.3 Survey of the Research Area in Cancer Chemotherapy

MULTIDISCIPLINARY INTERESTS

Medicine: therapy strategy for cures  
 Biology: knowledge of the biological mechanisms and principles  
 Pharmacology/Chemistry: drug effects/drug development  
 Technical Sciences: system dynamics and control



MODELING & SIMULATIONS FOR

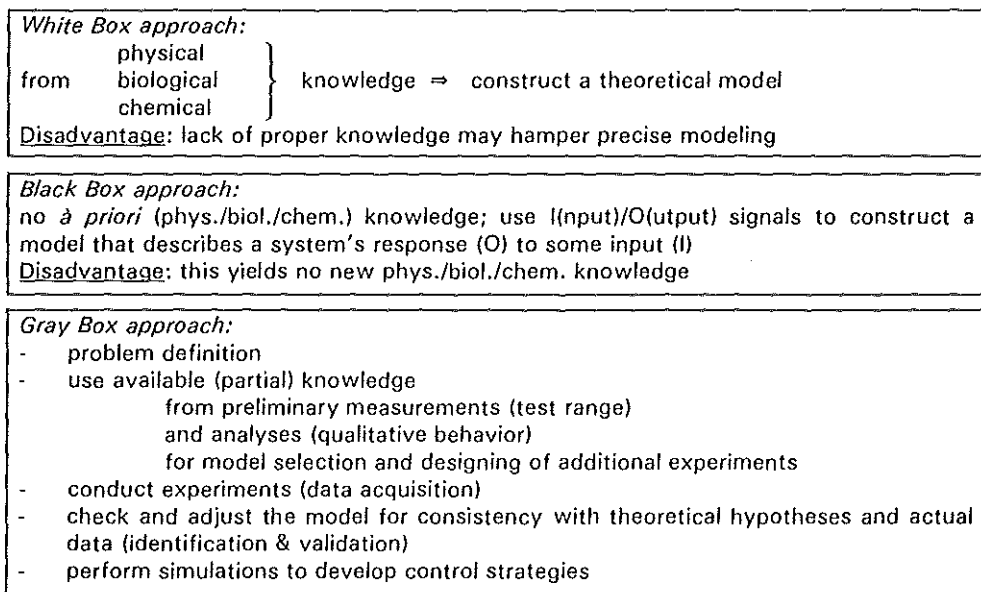


**1.7 MATHEMATICAL MODELING AND SIMULATION TO OPTIMIZE TREATMENT**

As discussed above, cancer chemotherapy necessitates seeking compromises, weighing profits and penalties to derive the optimal therapeutic ratio. Therefore, chemotherapy is a very suitable subject for the application of optimization techniques in order to improve—in advance, by evaluating several possible treatment strategies—the probability of good treatment results. The modeling approach should be applicable throughout the whole trajectory from basics to clinical application (Fig. 1.3).

Given a certain patient, suffering from a certain type of cancer, and given

Fig. 1.4 Approaches to MODEL CONSTRUCTION for Process or System Simulation [Schweppe, 1973]



the present spectrum of available cytostatic drugs, ideally it should be possible to determine 1) the drug or combination of drugs, 2) the total dose, 3) the dose rate, i.e., the time schedule for continuous or intermittent administrations, and 4) the route of administration that would yield the best results.

The objective is to cure the patient by killing all malignant cells as quickly as possible, with a minimum of complications and discomfort, as caused for instance by inevitable adverse effects of the presently used cytostatic agents.

Of course it will be difficult, if possible, to generalize treatment strategies because of interpatient variability, not only concerning tumor load and tumor location (even for a same type of cancer) but the response of both tumor and healthy tissues to the cytostatic drugs as well. However, best treatment strategies for certain classes of patients and tumors may be identified.

To that purpose, a patient and the cancer growing inside him can be regarded as a system, which is to be perturbed by the chemotherapy. The state of the system may be given by the number of tumor cells and the number of normal tissue cells at risk.

A requirement is the existence of (mathematical) models, or the feasibility to develop them. A mathematical model allows concise but compact description of the qualitative and quantitative system behavior in terms of the dynamic respon-

Fig 1.5 Scheme of FACTORS to be Considered for MODELING CHEMOTHERAPY of LEUKEMIA

Cell (Sub)Population Size as Function of Time

is determined by

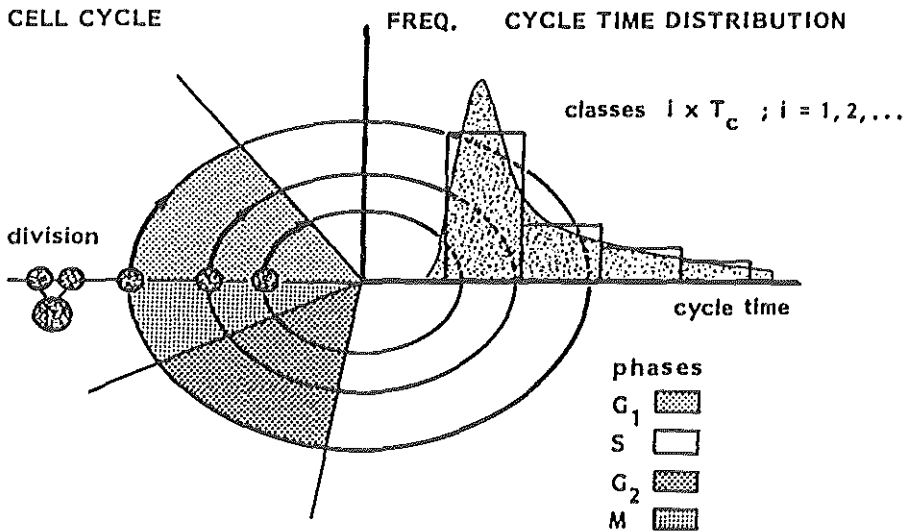
- ENVIRONMENTAL FACTORS  
(e.g., site: growth differences in bone marrow, spleen or liver)
- INITIAL CONDITIONS  
(population size at diagnosis, type of leukemia)
- GROWTH MODEL  
(cell proliferation, migration, mutation to resistance, relation with chronological time and/or cell population size)
- CHEMOTHERAPY MODEL  
(administration regimen, pharmacokinetics (processes of distribution, metabolism, elimination on tissue level; natural processes vs artificial enhancement ("magic bullets")), drug effect (log cell kill), pharmacodynamics (processes on cellular/molecular level), recruitment (accelerated cell proliferation triggered by treatment), relation with chronological time and/or cell population size)
- BOUNDARY CONDITIONS  
(e.g., max growth rate, toxicity related restrictions, patient characteristics (e.g., age, weight, organ function))

se to a certain input in the form of a drug administration schedule. Moreover, next to providing convenient means of analytical evaluation mathematical modeling also allows time-saving numerical manipulations on the computer. Such models must be based upon the insight into the population dynamics of tumor cells. This means that it must be known how a tumor grows, first without interference, next when a certain drug administration regimen is applied. Such knowledge often is not, or only fragmentarily present, or must be deduced first from several unrelated sources. Therefore, the most appropriate way of setting up a model is the gray box approach (Fig. 1.4 [Schweppe, 1973; Bohlin, 1994]). Both theoretical and experimental considerations are used to select the proper model and thus identify the studied system. Numerical values of model parameters can be obtained by, for instance, maximum likelihood estimation.

Maximum Likelihood Estimation (MLE) requires the formulation of a function that expresses the likelihood of the measured data in terms of a set of unobserved parameters that represent the source distribution. This likelihood function is defined by the joint probability density function (pdf) of the measured data in terms of the unobserved parameters to be estimated in the task. Maximizing this likelihood function with respect to the unobserved parameters yields estimates with which the measured data are most consistent [Brailean et

Fig 1.6 Schematic View of the Cell Cycle

Just after birth the cell is functionally active in the  $G_1$  phase. After a certain time preparations for cell division are started; in the S phase the cell synthesizes DNA (duplication of genetic material) to double the original amount. The cell then pauses some time in  $G_2$  phase. Next, actual cell division takes place in M phase (mitosis). Two daughter cells are produced, each starting in  $G_1$ . The cycle time,  $T_c$  required to complete the cell cycle is a variable distributed between a certain minimum value and infinity (the exact shape of the distribution depends on the type of cell); after a certain time interval, cells that started at the same time point thus may be at different positions in the cycle. Some (resting) cells stay in  $G_1$  for a very long time. An equivalent model can be built by allowing each cell a constant time interval  $T_c$  to complete the cycle and then assuming that it yields on average  $A$  ( $1 \leq A \leq 2$ ) daughter cells. The quantity  $A-1$  characterizes the cell production in relation to the potential cell production, and is called the growth fraction, GF. Together with  $T_c$  and the cell loss factor,  $\Phi$  (a measure of the difference between cell production and actual population doubling time), GF determines the growth of the cell population.



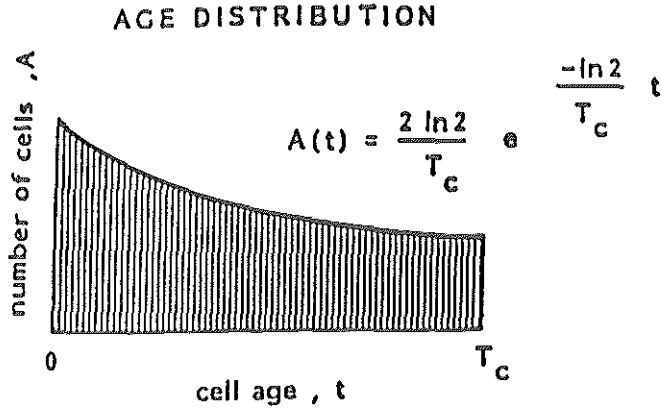
al., 1992].

Figure 1.5 gives a subdivision of factors that should be considered for modeling. On the one side is the tumor growth model with initial and boundary conditions, on the other side there is the chemotherapy model with pharmacokinetic processes (distribution/elimination) and dose—effect relationships (log cell kill per unit dose).

Of course, as with many biomedical systems, there remains the problem of observability, or better, lack of observability. Therefore, look separately at subsystems first: unperturbed tumor growth, drug distribution and clearance, response of tumor and normal tissue cells to exposure to drug concentrations.

Fig. 1.7 Stochastic Model for Population Growth

A) Given a certain initial population size, the cells are distributed over a number of age compartments, corresponding with the cell cycle which is assumed to last a constant value of  $T_c$ . The allocation takes place with the aid of a random number generator (RNG), assuming that an unperturbed cell population has an exponential age distribution (twice as many new born cells as mitotic cells [Matthews, 1988]).



B) The development of the population is simulated by moving cells from one compartment to the next one during small time increments. When moving a cell from the last compartment back into the first one, first it is determined (again with the RNG) whether the cell will disappear (chance  $p_2$ , e.g., natural cell death). If not, it is checked whether the cell will divide (chance  $p_1$ ) into two cells.

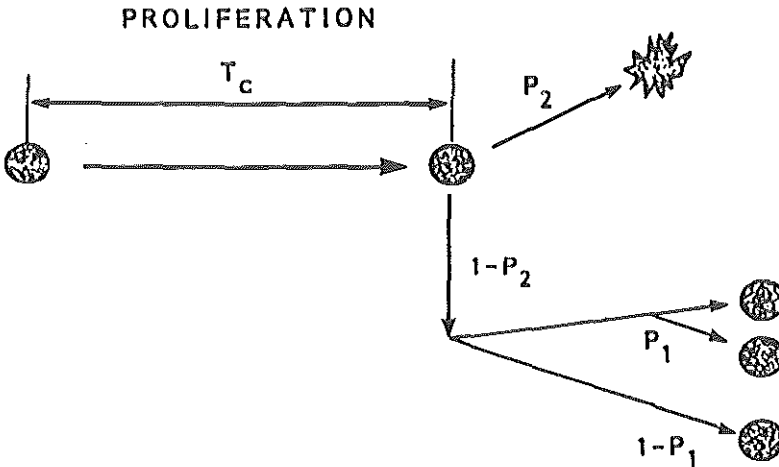
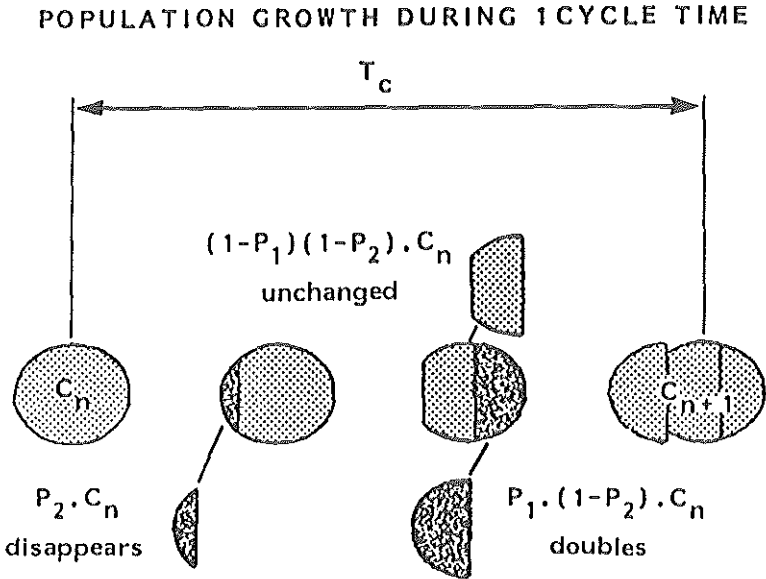


Fig. 1.7 (Continued)

C) For large cell populations the relation  $C_n = C_{n-1} \cdot (1 + p_1 - p_2 - p_1 \cdot p_2)$  can be derived for growth during consecutive cycle times. In this way growth curve are generated (population size as function of time). The chances  $p_1$  and  $p_2$  may vary as function of the population size. For the leukemias considered in Fig. 1.8,  $T_c \approx 14$  h (BNML) and  $T_c \approx 3.2$  d (median value for AML, range 0.7-12.2, determined by *in vivo* BrdUrd and  $^3\text{H}$ -Thymidine double labeling (Raza et al., 1987)).



This is best done in laboratory animal studies.

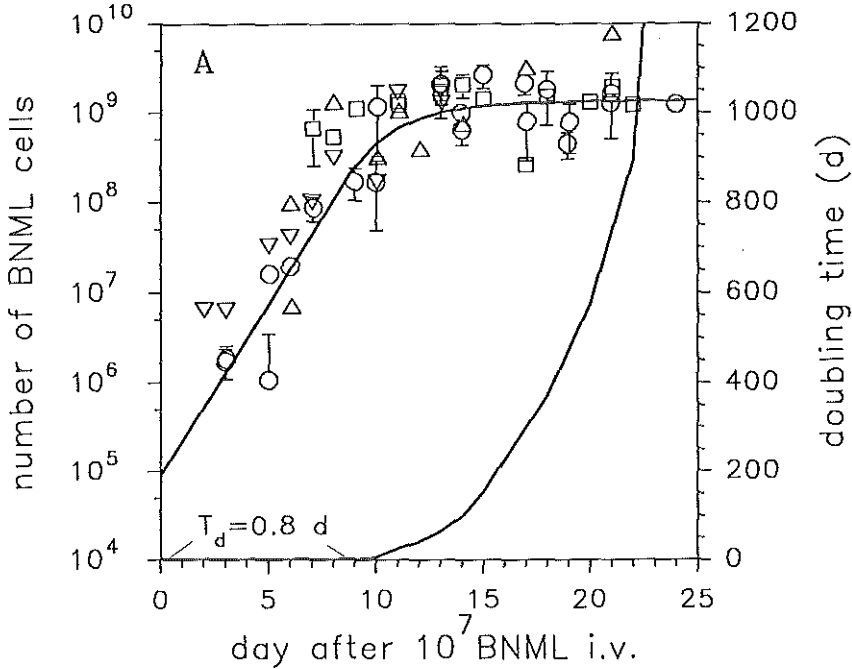
With a few illustrations an impression is given of how simulation for cancer chemotherapy may be applied. Note that not all model assumptions used in the simulations have been verified (yet).

Figure 1.6 shows how cells move through a cell cycle before reproduction. A model for cell proliferation is shown in Fig. 1.7. Figure 1.8 demonstrates that such a model can yield realistic growth curves, assuming that during tumor development cells keep a constant chance of dying while their chance of producing offspring decreases linearly with the population size (the latter property is also seen in normal tissues: an organ grows only to a certain size).

Figure 1.9 divides a cell population into fractions that are sensitive or resistant to two cytostatic drugs. Simulation results for two different treatment strategies based on this model are given in Fig. 1.10. The validity of the log cell kill hypothesis [Skipper, 1986] is assumed when modeling the effect of a drug dose, i.e., a same dose always kills a constant fraction of the (sensitive) cell

Fig. 1.8 Growth of Leukemic Cell Populations

Panel A shows the growth of leukemic (BNML) cells in the bone marrow of the BN rat, after *i.v.* inoculation of  $10^7$  BNML cells on day zero. The measured points are derived from various laboratory experiments. The fitted growth curve (least squares method) consists of an exponential component (constant population doubling time, 0.8 d) and a contiguous Gompertz curve (from day 8.6 on, at a population size of  $1.8 \times 10^8$  cells, the doubling time  $T_d$  increases exponentially, with retardation constant  $0.4 \text{ d}^{-1}$ ).



population. The magnitude of such a fraction is experimentally determined for each cytostatic drug, e.g., as shown in Fig. 1.11 for the drugs cyclophosphamide (CFA) and acridinyl anisidide (AMSA). By doing simulations with the cell cycle model it is possible to check whether the drug might be cell cycle phase specific, i.e., whether the drug affects cells only when they are in a certain phase of the cell cycle (Chapter 2.3). Likewise, the instantaneous nature of the cell killing effect of a drug can be tested by using instantaneous and gradual drug influence models and checking simulation results against experimental observations (Chapter 2.2).

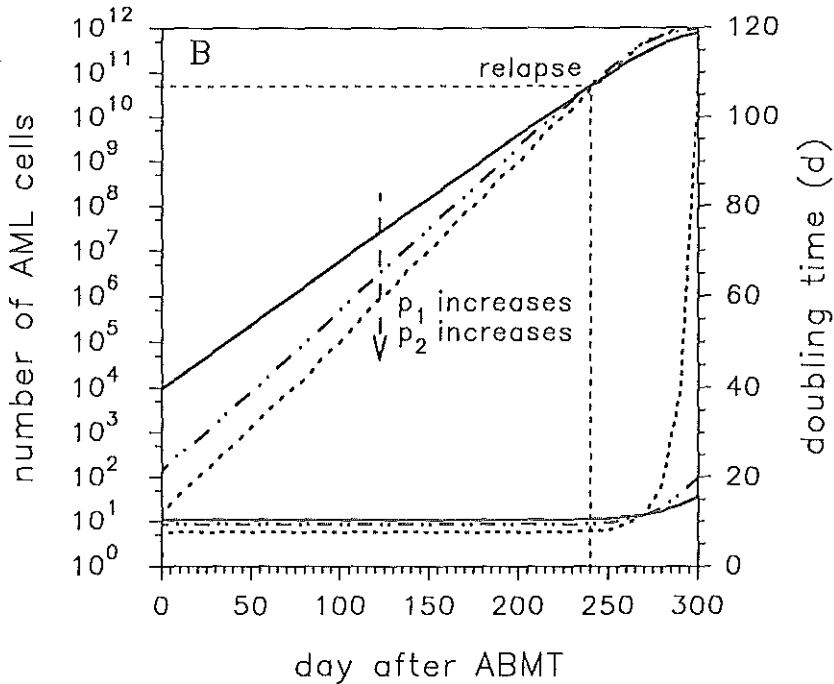
## 1.8 AIM OF THE THESIS

The purpose of this thesis is to show advantages of the application of techniques

Fig. 1.8 (Continued)

Panel B shows similar curves, obtained for human acute myelocytic leukemia (AML, observed median value of  $T_c \approx 3$  d) with the simulation model of Fig. 1.7 by keeping the death probability ( $p_2$ ) constant while decreasing the probability of cell division ( $p_1$ ) linearly with population size.

Relapse ( $5 \times 10^{10}$  cells) on day 240 after treatment (clinically observed median value after autologous bone marrow transplantation (ABMT)) may originate from different quantities of residual cells as long as the true course of  $T_d$  remains unknown.



developed in general systems theory to solve biomedical problems and generate knowledge. Biological systems are complex and often understood to a very limited extent only. In general, the predictability of system responses is correspondingly poor. Through modeling, it will be possible to explain at least part of the observed system responses. Thus, fluctuations in a response may not fully disappear—e.g., due to variation in individual patients—but they will be reduced.

Two different worlds have to be united. That of the clinicians, who have to live with large unexplained variations in observed (system) responses, face apparently contradicting results of a multitude of different (experimental) treatments that each throw light on only a limited aspect of the studied phenomena, therefore are obliged to base pragmatic solutions ("rules of thumb") for their individ-



Fig. 1.9 Schematic View on Relations between Subpopulations

There are four subpopulations: cells in compartment S are sensitive for both cytostatic agents A and B; cells in  $R_A$  are resistant for drug A; cells in  $R_B$  are resistant for drug B; and cells in  $R_{AB}$  are resistant for both drugs. Double resistant cells develop from single drug resistant cells by mutation; single drug resistant cells in turn develop from sensitive cells. Per unit time and per cell in every subpopulation  $b$  cells are born and  $d$  cells disappear. A fraction  $m$  of the difference mutates into another subpopulation, the remainder stays in the same subpopulation. In principle the value of  $m$  may differ for the different subpopulations, but in this case the somatic mutation rate  $m = 10^{-6}$  leukemias kept constant. This also applies to the  $b$  and  $d$  values in the subpopulation; no difference between subpopulations, though  $(b-d)$  may be dependent on total population size.

The equations for population size as function of time ( $C$  denotes the total number of cells) can be written as:

$$C(t+\Delta t) = C(t) + \Delta t \cdot dC(t)/dt, \text{ where}$$

$$dC(t)/dt = dS/dt + dR_A/dt + dR_B/dt + dR_{AB}/dt \text{ and}$$

$$\begin{bmatrix} dS/dt \\ dR_A/dt \\ dR_B/dt \\ dR_{AB}/dt \end{bmatrix} = (b-d) \cdot \begin{bmatrix} (1-2m) & 0 & 0 & 0 \\ m & (1-m) & 0 & 0 \\ m & 0 & (1-m) & 0 \\ 0 & m & m & 1 \end{bmatrix} \cdot \begin{bmatrix} S \\ R_A \\ R_B \\ R_{AB} \end{bmatrix}$$

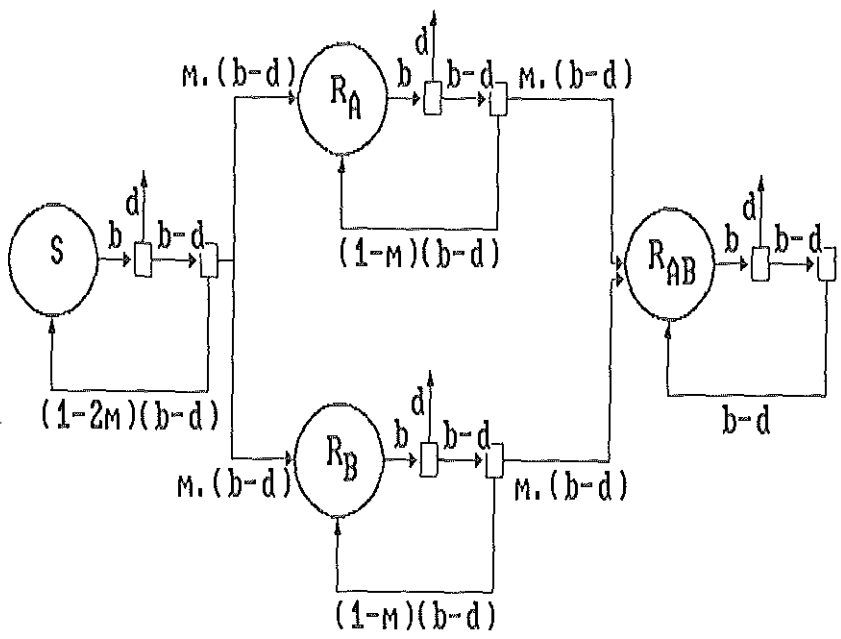


Fig. 1.10 Example of Chemotherapy Simulations in the BNML

At the start of the simulation the leukemic cell population consists of  $1.65 \times 10^9$  sensitive cells that would double in 19 h if unperturbed exponential growth is allowed. Drug resistance develops at a spontaneous mutation rate of 1 cell in  $10^6$  divisions. Drug administration reduces the sensitive population with a dose dependent number of decades (q log cell kill means a reduction with a factor of  $10^q$ ) according to an experimentally determined dose—effect relationship (see Fig. 1.11). Different treatment strategies are compared.

**Monotherapy.** Two high doses of AMSA (20 mg/kg causing 4.7 LCK each) at 12 and 84 h are given, and nine low doses (5 mg/kg causing 1.2 LCK each) divided over two series, e.g. to prevent toxicity problems, at 24, 36, 48, 60, 72 and 96, 108, 120 and 132 h. The total dose amounts to 85 mg/kg, administered in 132 h, so dose intensity is  $85/132 = 0.64$  mg/(kg.h). At first the cell population decreases but after 330 h it appears to reach the initial size again. The simulation model predicts that AMSA treatment could have been stopped after 72 h. Further treatment has been in vain as all sensitive cells had already been eradicated at that moment and the remaining resistant cells are not affected by AMSA. Worse, normal tissues have been unnecessarily put at risk.

**Combination therapy.** Next to low doses of AMSA (5 mg/kg causing 1.2 LCK) at 12, 24, 36, 48 and 108, 120, 132, 144 h a few doses of CFA are given (60 mg/kg causing 2.7 LCK) at 60, 72, 84 and 96 h. The total dose, expressed in AMSA equivalents—the same LCK is attained with a CFA dose which is 5.2x as large as the required AMSA dose—as  $8 \times 5 + 4 \times 60 / 5.2 = 86$  mg/kg is given in 144 h, so dose intensity is  $86/144 = 0.60$  mg/(kg.h). Despite the almost equal exposure of the patient with respect to the monotherapy the model now predicts cure ( $< 1$  cell left). The total population is destroyed in 120 h. The last two AMSA administrations at 132 h and 144 h have apparently been superfluous.

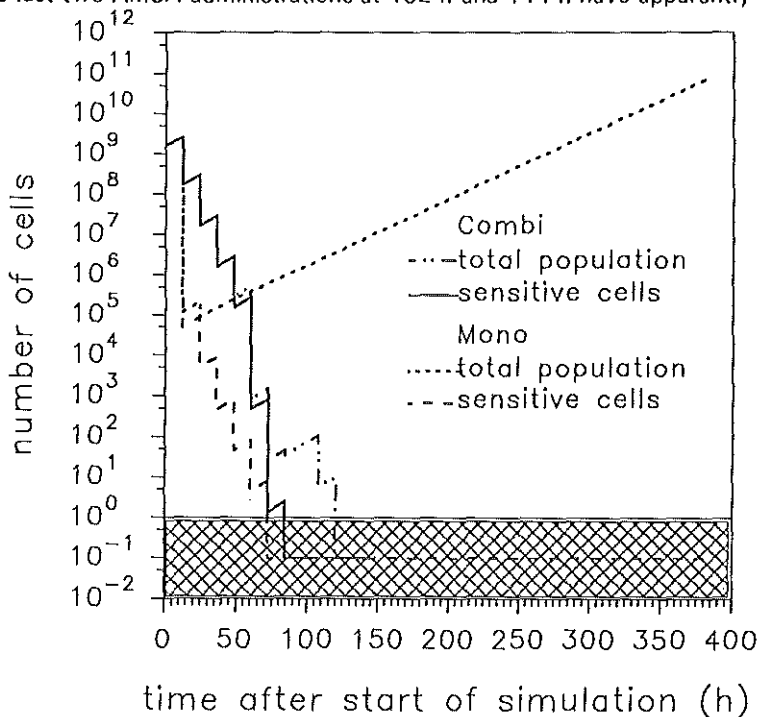
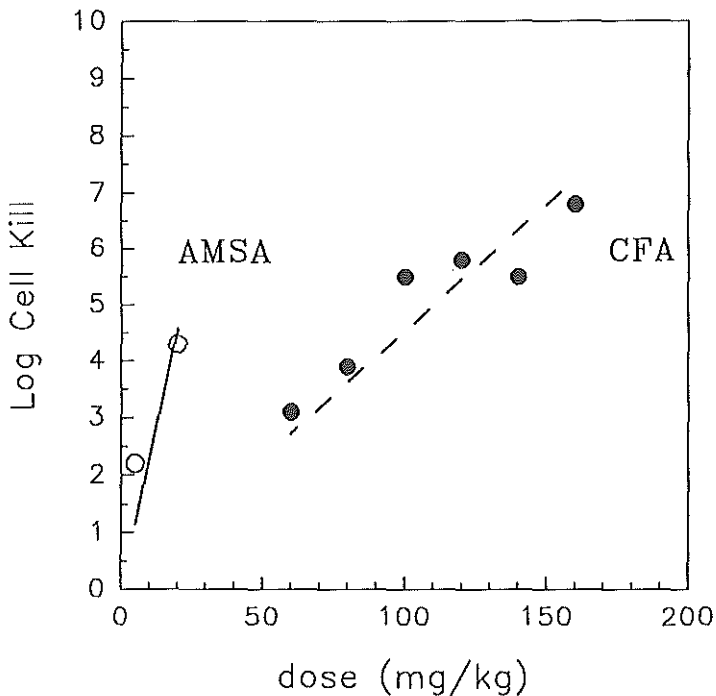


Fig. 1.11 Dose Effect Relationships for BNML and Two Cytostatic Drugs  
 Observations for AMSA (acridinyl anisidide) and CFA (cyclophosphamide) and BNML cells.  
 Regression lines show linear relationships between dose and log cell kill (correlation  
 coefficient,  $r=0.93$  and  $r=0.87$ , respectively).



ual patients on some hazy notions about the true mechanisms behind the processes. And that of the theoretic physicists, striving for exact formalisms to describe and understand the processes, but who sometimes are more fixed on the elegance of their mathematics than on gathering experimental "evidence", perhaps by lack of laboratory facilities and skilled technicians' assistance, and who, conversing in "cryptic" formulae, can happily live with 1.75 children and allow 0.24 viable cells. Obviously, there must be fruitfulness in sharing views. Other interested parties may be those educated in biology, biochemistry or pharmacology.

As an illustrative example the problem of optimization of chemotherapy of leukemia is chosen. In this rather vast field a limited number of subproblems are addressed. They mainly concern a myelocytic leukemia growing in a laboratory rat (BNML). The subproblems considered include the identification of a system of pharmacokinetics of a cytostatic agent, and leukemia growth under various boundary conditions.

## 1.9 EXPERIMENTAL DATA

### 1.9.1 Laboratory Animals

The *in vivo* data discussed in this thesis were obtained, in general, from experiments conducted with inbred Brown Norway rats, BN/Bi/Rij, raised in the Rijswijk colony. Animals used were 10-12 weeks old at the start of the experiments. The total body weights of male animals were about 165 g, of female animals about 150 g. Occasionally, older rats were used, with ages between 13 and 16 weeks and average weight of males: 220 g. In view of the expected life span of 36 months for rats, these animals can be considered young adults. The rats always had free access to water and pelleted rat food.

### 1.9.2 Brown Norway Rat Leukemia Model (BNML)

In the early seventies a leukemia was discovered at the Radiobiological Institute TNO in a female BN rat that had been repeatedly injected with carcinogenic dimethyl-benzanthracene (DMBA). The disease appeared to be an acute promyelocytic leukemia and was called Brown Norway rat Myelocytic Leukemia (BNML). The BNML cell line can be maintained by transplantation, i.e., leukemic cells that in a late stage of the disease have been isolated from the spleen or the bone marrow of a leukemic animal are able to replicate in a normal BN rat. This ability persists if the isolated BNML cells have been kept stored in frozen condition for quite some time. Transfer of small numbers of donor BNML cells by means of intravenous (*i.v.*), subcutaneous (*s.c.*) or intraperitoneal (*i.p.*) injection of single cell suspensions causes leukemia in recipient rats. As few as 25 BNML cells *i.v.* yield a 50% risk of leukemia development (see Chapter 6.2).

Because of various similar properties, the BNML is considered a very good model for human acute (pro)myelocytic leukemia (AML) [Hagenbeek and Van Bekkum, 1977; Van Bekkum and Hagenbeek, 1977]. The most important among such properties are:

- 1) Striking similarities with respect to cytology, cytochemistry and histology
- 2) Absence of leukemia-specific antigens
- 3) No viral activity has been demonstrated to serve as an etiological agent
- 4) Relatively slow growth, due to a low growth fraction (40-50%) and a high cell loss rate (80-90%) at later stages of the disease; the median survival time of BN rats after *i.v.* inoculation of  $10^7$  BNML cells amounts to 22 days, the leukemic cell load then being about  $4 \times 10^{10}$  BNML cells (see Chapter 2.1)
- 5) Signs of diffuse intravascular coagulation when the leukemia progresses

- 6) A dramatic decrease in numbers of normal hemopoietic stem cells results in a severe suppression of normal hemopoiesis
- 7) Similar response of BNML and human AML to chemotherapeutic agents

Another property that makes the BNML such a convenient model is the presence of leukemic clonogenic cells. As BNML can be quantified in an *in vivo* colony-formation assay (Chapter 2.1), a variety of experiments can be conducted to generate data on the behavior of the leukemic cell population. In contrast, clinical data can never be that extensive.

The BNML rat leukemia model has been, and still is, used in over a dozen European and American research centers for various preclinical investigations on the diagnosis and treatment of human AML. A comprehensive survey of BNML studies performed at the former Radiobiological Institute TNO can be found in Hagenbeek and Martens [1991] and Martens et al. [1990a,b].

### 1.9.3 Generation of Data

For cell kinetics experiments the BN rats were inoculated intravenously (*i.v.*) with leukemic cells, usually  $10^7$ , so as to induce disease. The development of the leukemia as function of inoculum size and growth perturbation (none, cytostatic drug administration or total body irradiation) was monitored in several ways (see Chapter 2.1).

For pharmacokinetics experiments leukemic BN rats were injected (*i.v.*) with a single dose of an anticancer drug. Its concentrations were measured in various organs at various time points after the treatment (see Chapter 4).

The BNML has been studied extensively with respect to problems met in the area of bone marrow transplantation (BMT) [Martens et al., 1990a]. Among these studies are the conditioning regimens (marrow ablative chemo-radiotherapy) before BMT, the removal of malignant cells from an autograft, and treatment-related complications such as interstitial pneumonia and graft-versus-host disease [e.g., Hagenbeek and Martens, 1983, 1987; Hagenbeek et al., 1989; Kloosterman et al., 1993; Van Bekkum, 1993]. Mathematical analyses and deductions about BMT made in this thesis (Chapter 6) are based on the basic experiments conducted with the BNML and on clinical observations (survival data of patients that are treated in different ways) derived from the literature.

## 1.10 OUTLINE OF THE THESIS

In the following chapters the modeling approach is applied to address the problems raised in Chapter 1. Mathematical models are built and—employing parameter estimation and system identification techniques—tested on the basis of available biological data, and results of computer simulations are analyzed to

better understand various processes that are relevant for chemotherapy of leukemia.

First, more must be learnt about how the leukemic cell population develops *in vivo*, especially in the situation of MRD. In *Chapters 2 and 3* the cell population dynamics of leukemia growth will be dealt with. In Chapter 2 the BN rat leukemia (BNML) will be discussed. To show how the necessary input data for the developing the models were obtained, in *Chapter 2.1* an overview is given of the basic experiments. Their results are important for the remainder of the thesis. In *Chapter 2.2* growth of the BNML cell population is examined, as well as the influence of chemotherapy, applied as a single dose of the cytostatic agent cyclophosphamide. In *Chapter 2.3* a simulation study is described to further investigate the influence of chemotherapy. Here, the cytostatic drug AMSA is given as a series of daily doses. A first approach to modeling the development of drug resistance is the subject of *Chapter 2.4*. The model for BNML growth should be validated to see whether the knowledge gained can be applied to human leukemia. Clinical data on human AML—probably the most appropriate for testing the model—was not available. However, data on the growth, and regrowth during and after therapy, of childhood T-cell acute lymphocytic leukemia was willingly supplied by the Department of Immunology of the Erasmus University Rotterdam. In *Chapter 3* these data are analyzed for malignant cell growth in peripheral blood.

Control of *in vivo* drug distribution is another important factor for optimal chemotherapy. Before control is possible, more must be learnt about the mechanisms behind the distribution processes. This requires studies in laboratory animals. In *Chapter 4* system identification techniques are applied to the pharmacokinetics of the drug daunomycin and its metabolite daunomycinol in the Brown Norway rat.

The system identification technique developed in the previous chapter is generally applicable. In *Chapter 5* it is used in a new method to analyze flow cytometric DNA histograms. With this tool the cell kinetics of (malignant) cell populations can be studied. This may yield new information on the progression of the leukemic cell populations, in addition to the results of Chapter 2.

Bone marrow transplantation is a recent development in the treatment of leukemia. In *Chapter 6* the application of mathematical analysis to a few of its problems is discussed. After an introduction to the subject (*Chapter 6.1*) possibilities for improving the probability of disease free survival after allogeneic bone marrow transplantation are examined (*Chapter 6.2*). The contributions of minimal residual disease and leukemic cells reinfused with an autologous bone marrow graft to the risk of leukemia relapse are compared (*Chapter 6.3*).

Finally, *Chapter 7* comprises a general discussion and comments on a few developments. It is concluded that in the vast area of optimization of treatment of leukemia there are many problems that can, and should, be analyzed using

mathematical models. The problems focussed upon in this thesis form a small, but important, subset to start with.

## 1.11 REFERENCES

- Acheraya RS, Sundareshan MK (1984) Development of optimal drug administration strategies for cancer chemotherapy in the framework of systems theory. *Int J Biomed Comput* 15:139-150
- Avula XJR (1987) Mathematical modeling. In: Meyers RA (ed) *Encyclopedia of physical science and technology*. Vol 7, Io-Mec, 719-728. Academic Press, New York
- Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DAG, Gralnick HR, Sultan C (1985) Proposed revised criteria for the classification of acute myeloid leukemia; a report of the French-American-British cooperative group. *Ann Intern Med* 103:626-629
- Birkhead BG, Gregory WM (1984) A mathematical model of the effects of drug resistance in cancer chemotherapy. *Math Biosci* 72:59-69
- Birkhead BG, Gregory WM, Slewin ML, Harvey VJ (1986) Evaluating and designing cancer chemotherapy treatment using mathematical models. *Eur J Cancer Clin Oncol* 22:1-8
- Bohlin T (1994) A case study of grey box identification. *Automatica* 30:307-318
- Brailean JC et al. (1992) Application of the EM algorithm to radiographic images. *Med Phys* 19:1175-1182
- Brock N, Schneider B (1984) Models in cytostatic chemotherapy. *Cancer* 15 (supplement):1229-1238
- CBS, Centraal Bureau voor de Statistiek (1991) *Statistisch Jaarboek 1991*. SDU Uitgeverij, Cbs Publicaties, 's-Gravenhage, p22, p424, p440
- Cline MJ (1994) The molecular basis of leukemia. *New Engl J Med* :328-336
- Coldman AJ, Goldie JH (1985) Role of mathematical modeling in protocol formulation in cancer chemotherapy. *Cancer Treatm Rep* 69:1041-1045
- Coldman AJ, Goldie JH (1986a) Factors affecting the development of permanent drug resistance and its impact upon neoadjuvant chemotherapy. *Recent Results Cancer Res* 103:69-78
- Coldman AJ, Goldie JH (1986b) Variation in growth parameters and their effect on the acquisition of drug resistance. *Prog Clin Biol Res* 223:103-111
- Conelly RB, Andersen ME (1991) Biologically based pharmacodynamic models: tools for toxicological research and risk assessment. *Annu Rev Pharmacol Toxicol* 31:503-523
- Dembo AJ (1984) Spontaneous mutation to chemotherapy resistance: implications of the Goldie-Coldman model for the management of ovarian cancer. *J Clin Oncol* 2:1311-1316
- Garfinkel D (1984) Modeling of inherently complex biological systems: problems, strategies, methods. *Math Biosci* 72:131-139
- Goldie JH, Coldman AJ (1979) A mathematical model for relating the drug sensitivity of tumors to their spontaneous mutation rate. *Cancer Treatm Rep* 63:1727-1733
- Goldie JH, Coldman AJ, Gudauskas GA (1982) Rationale for the use of alternating non-cross resistant chemotherapy. *Cancer Treatm Rep* 66:439-449
- Goldie JH, Coldman AJ (1983) Quantitative model for multiple levels of drug resistance in clinical tumors. *Cancer Treatm Rep* 67:923-931
- Goldie JH, Coldman AJ (1985) A model for tumor response to chemotherapy: an integration of the stem cell and somatic mutation hypotheses. *Cancer Invest* 3:553-564
- Goldie JH, Coldman AJ (1986) Application of theoretical models to chemotherapy protocol design. *Cancer Treatm Rep* 70:127-131
- Hagenbeek A, Van Bekkum DW (eds) (1977) *Proceedings of an international workshop on "Comparative evaluation of the L5222 and the BNML rat leukemia models and their relevance for human acute leukemia"*. *Leukemia Res* 1:75-256

- Hagenbeek A, Martens ACM (1983) Cell separation studies in autologous bone marrow transplantation for acute leukemia. In: Gale RP (ed) Recent advances in bone marrow transplantation. UCLA Symposia on Molecular and Cellular Biology, New Series, 7:717-735, Alan R Liss Inc, New York
- Hagenbeek A, Martens ACM (1987) Conditioning regimens before bone marrow transplantation in acute myelocytic leukemia. In: Dicke KA, Spitzer G, Jagannath (eds) Autologous bone marrow transplantation. Proc Third Int'l Symposium, The University of Texas, MD Anderson Hospital and Tumor Institute, Houston, 99-104
- Hagenbeek A, Schultz FW, Arkesteijn GJA, Martens ACM (1989) Animal models of bone marrow transplantation for acute myelocytic leukemia. In: Gale RP, Champlin (eds) Bone marrow transplantation: Current controversies. UCLA Symposia on Molecular and Cellular Biology, New Series, 9:179-189, Alan R Liss Inc, New York
- Hagenbeek A, Martens ACM (1991) Minimal residual disease in leukemia: Preclinical studies in a relevant rat model (BNML). In: Jasmin C, Proctor SJ (eds) Ballières "Clinical Haematology" 4 (3) Chapter 3:609-635. Ballière Tindall, WB Saunders Company, London
- Hagenbeek A (1992) Minimal residual disease in leukemia: State of the art 1991. *Leukemia* 6, Suppl 2:12-16
- Hokanson JA, Brown BW, Thompson JR, Jansson B, Drewinko B (1986) Mathematical model for human myeloma relating growth kinetics and drug resistance. *Cell Tissue Kinet* 19:1-10
- Kloosterman TC, Tieleman MJC, Martens ACM, Van Bekkum DW, Hagenbeek A (1993) Graft-versus-leukemia after allogeneic bone marrow transplantation. *Blood* (submitted)
- Martens ACM, Van Bekkum DW, Hagenbeek A (1990a) The BN acute myelocytic leukemia (BNML). A rat model for studying human acute myelocytic leukemia. *Leukemia* 4:241-257
- Martens ACM, Van Bekkum DW, Hagenbeek A (1990b) Minimal residual disease in leukemia. Studies in an animal model for acute myelocytic leukemia (BNML). *Int J Cell Cloning* 8:27-38
- Mathews JH (1988) Cell cycle characteristics in acute nonlymphoblastic leukemia. *Blood* 71:532-534
- Michelson S (1987) Facilitation of emergence of multidrug resistant state by alteration of tumor environment: implications from competitive ecology models. *Cancer Treatm Rep* 71:1093-1094
- Perry MC (ed) (1992) The chemotherapy source book. Williams and Wilkins, Baltimore
- Raza A, Maheshwari Y, Preisler HD (1987) Differences in cell cycle characteristics among patients with acute nonlymphoblastic leukemia. *Blood* 69:1647-1653
- Rice JR (1983) Matrix computation & mathematical software. McGraw-Hill, New York
- Rosen R (1986) Role of mathematical modeling in protocol formulation in cancer chemotherapy. *Cancer Treatm Rep* 70:1461-1462
- Saiga T, Horie K, Tabuchi K, Midorikawa O (1985) Mathematical models of cell variation seen in a heterogeneous malignant cell population. *Exp Pathol* 28:21-30
- Schweppe FC (1973) Uncertain dynamic systems. Prentice Hall Inc., Englewood Cliffs, NJ
- Skipper HE (1986) On mathematical modeling of critical variables in cancer treatment (goals: better understanding of the past and better planning in the future). *Bull Math Biol* 48:253-278
- Van Bekkum DW, Hagenbeek A (1977) The relevance of the BN leukemia as a model for human acute myeloid leukemia. *Blood Cells* 3:565-574
- Van Bekkum DW (1993) Biological aspects and clinical implications of GvHD. *Bone Marrow Transplantation* 11(suppl 1):99-102
- Van Dongen JJM, Adriaansen HJ, Hooijkaas H (1988) Immunophenotyping of leukemias and non-Hodgkin lymphomas. *Neth J Med* 33:298-314
- Warrell RP, De Thé H, Wang ZY, Degos L (1993) Acute promyelocytic leukemia. *New Engl J Med* :177-189



## Chapter 2

### Cell Population Dynamics of Leukemia Growth in the Brown Norway Rat

In this chapter the dynamics of *in vivo* leukemia growth will be examined. The unperturbed development of leukemic cell populations, as well as the influence of radio-/chemotherapy in terms of log cell kill and development of drug resistance, were studied first in the BNML laboratory model. Later, the application of preclinical findings to actual clinical data is discussed, i.e., the treatment of several cases of childhood T-cell acute lymphocytic leukemia (Chapter 3).

#### 2.1 Basic Experiments and *in vivo* Growth

##### 2.1.1 BASIC EXPERIMENTS

Various experiments have been conducted to generate basic data on *in vivo* growth of BNML in the BN rat [Hagenbeek and Martens, 1981, 1982, 1985, 1987a].

###### 2.1.1.1 Homing and Lodging of BNML Cells in the Recipient Rat

Labeling with radioactive chromium (yielding measurable quantities of  $^{51}\text{Cr}$  in more than 90% of the BNML cells) revealed that the bulk of the BNML cells, injected *i.v.*, goes to and settles in the liver, the spleen and the bone marrow. Within a few hours 55%, 10% and 2.5%, respectively, of the inoculum is found in these organs. Very low percentages arrive in other organs. Apparently, only in the liver, the spleen and the bone marrow BNML cells may find favorable spots where they can start the process of reproduction (proliferation).

###### 2.1.1.2 BNML Cell Numbers Transferred and Leukemia Induction

When a number of malignant cells is inoculated *i.v.* only a fraction will home to and lodge in favorable places in the recipient such that offspring can be produced that leads to overt disease. A certain size of the inoculum is required to let the disease develop in 50% of the recipients. This inoculum size is called the  $\text{ED}_{50}$  value [Hewitt, 1958].

The  $\text{ED}_{50}$  value for BNML was determined by inoculating rats *i.v.* with low numbers of BNML cells (1 to 1000). In each group of 6 rats, who were given a same dose of BNML cells, the percentage of leukemia-free survivors was recorded. A probit analysis of the datapoints yielded the  $\text{ED}_{50} = 24.7$  BNML cells *i.v.* being required to induce leukemia in 50% of the cases (see Fig. 2.1).

### 2.1.1.3 BNML Cell Numbers Transferred and Median Survival Time of the Recipient Rat

Various amounts of BNML cells—dose (C) in the range from  $10^3$  to  $10^7$ —were inoculated *i.v.* into recipient rats (6 rats per datapoint). Death by leukemia was awaited and the survival time since inoculation was noted. The median survival time (MdST) was calculated for each dose of BNML cells. By linear regression of the MdST—log(C) datapoints the following relationship was established (Fig. 2.2):

$$C = 10^{-0.25 \times MdST + 12.5} \quad (2.1)$$

The MdST after  $10^7$  BNML cells *i.v.* amounts to 22 days. Equation (2.1) implies that each tenfold reduction in cell dose (1 log) corresponds with 4 days extra survival. The other way around, from an observed increase in life span the corresponding decrease in cell numbers (dose) can be calculated from Eq.(2.1).

### 2.1.2 DETECTION AND QUANTIFICATION METHODS

Several methods are available to detect and quantify leukemic cells in a (tissue or body fluid) sample. The methods mentioned below—in order of increasing sensitivity, i.e., ability to detect cells that are present in decreasing frequencies—are applicable only to preclinical studies. There, experimental animals can be sacrificed and/or many (large) samples can be obtained. Exceptions may be the methods of direct counting and of flow cytometry, which may be used with human (blood) samples and/or (tissue) biopsies. In those cases, however, malignant cells will only be detected if their frequency in the sample is relatively high. Because clinical samples generally are small, due to processing time limitations and/or ethics in patient handling. The clinical detection level—below which the presence of malignant cells goes unnoticed—amounts to more than  $10^{10}$  malignant cells, i.e., 1% of the typical leukemic cell load of  $10^{12}$  leukemic cells (1 kg) at diagnosis. With new detection methods (see Chapters 3 and 7) malignant cells may be detected when they are present in frequencies of one in  $10^4$  or  $10^5$ .

#### 2.1.2.1 Organ Weights

When, after *i.v.* inoculation, BNML cells home to and lodge in the liver and spleen of the rat the weights of these organs slowly increase with time as several BNML cells will form growing colonies. Therefore, the excess weight (= total weight - normal organ weight in a healthy animal) may be attributed to the leukemic cell load. The weight of  $10^9$  BNML cells being 1 g, the excess organ

Fig. 2.1 Determination of the  $ED_{50}$  value for BNML. Shown are observed datapoints and the dose-response relationship with 95% confidence limits after probit analysis

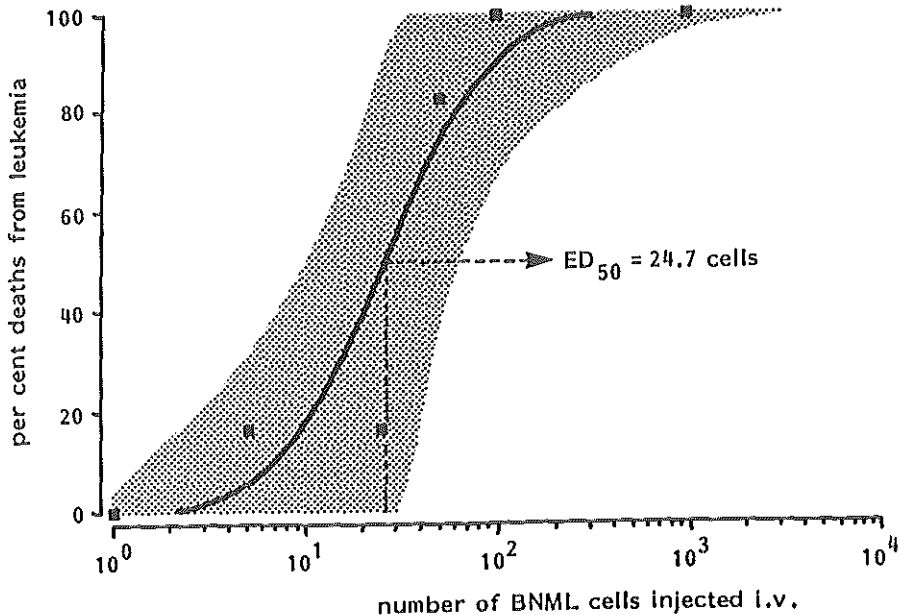
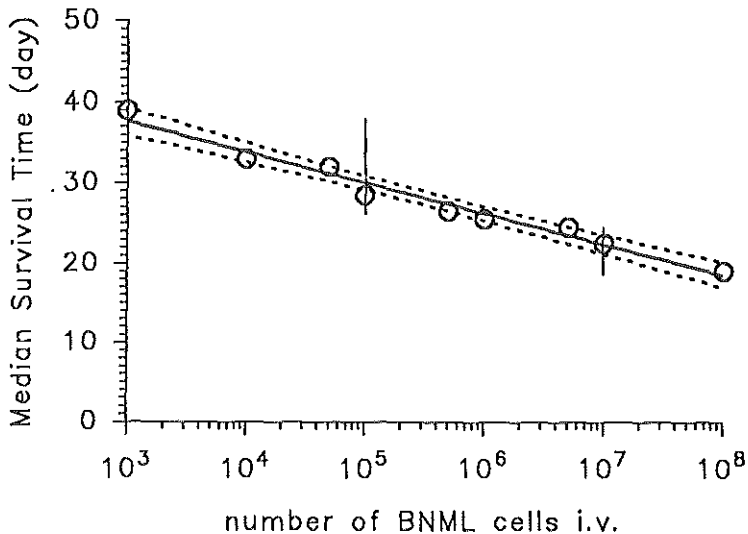


Fig. 2.2 Relationship between *i.v.* inoculated dose,  $C$  of BNML cells and median survival time, MdST of recipient rats. Bars show ranges of observed survival times. Log-linear regression yields:  $\log(C) = -0.25 \text{ MdST} + 12.50$ , with correlation coefficient  $r = -0.9962$  and 95% confidence limits as shown by the broken lines



weight can be converted to a number of cells.

The organ weight parameter obviously yields very rough estimates of the leukemic cell burden. The increase in weight must be considerable before a significant deviation from normal values (spleen: 0.5 g, liver: 10 g) can be detected. Thus, basically, the method is useful in late stages of the disease, when the fraction of leukemic cells is larger than 0.1.

In Fig. 2.3 the development of spleen, liver and total body weight after inoculation of  $10^7$  BNML cells is illustrated.

#### 2.1.2.2 Counting based on Morphology (Quantitative Cytology)

Based on cytological staining procedures and morphological characteristics it is possible to detect and count leukemic cells under the microscope if they are present in frequencies of 0.01 or larger. The leukemic cell load is estimated from the sample counts by correcting for the total numbers of cells per organ as determined by cell counting or organ weights.

#### 2.1.2.3 Flow Cytometry

The monoclonal antibody (MCA), Rm-124, specifically binds to an antigen that is present in high density on the surface of BNML cells. This MCA itself can be conjugated with fluorescein isothiocyanate (FITC). Thus, when running a cell suspension through a flow cytometer, BNML cells—if they are present in frequencies larger than 0.0001—can be distinguished from other cells because of the high intensity fluorescent signal evoked from the FITC+Rm-124 label by laser light excitation. (See [Martens et al., 1984; Martens and Hagenbeek, 1985] for details on this method; see Chapter 5 for the principle of flow cytometry).

#### 2.1.2.4 Clonogenic Leukemic Stem Cell Assay

This method is also known as the "leukemic colony forming unit-spleen" assay (LCFU-S; see Table 2-1). It is based on the fact that a certain fraction of the BNML cells that are inoculated *i.v.* into a normal recipient rat will cause the development of cell colonies on the surface of the spleen, i.e., white spots that can be easily seen and counted on day 19-20.

It is assumed that each colony is founded by one inoculated BNML cell. It was experimentally determined [Van Bekkum, 1977] by an assay based on the one for normal hemopoietic stem cells [Lahiri et al., 1970]—as briefly described in Appendix A—that, on average (with an accuracy of approximately 10%), one out of every 500 BNML cells present in the inoculum yields a spleen colony.

Thus, if an inoculum from a donor rat contains leukemic cells, the actual number can be deduced from the counted number of colonies on the surface of

Fig. 2.3 Development of spleen, liver and total body weights after *i.v.* transfer of  $10^7$  BNML cells on day 0 (solid lines). On day 13 a single dose of cyclophosphamide (100 mg/kg) is given *i.v.* (dashed lines). Bars represent 1 SE (sometimes hidden within symbol);  $n=3$  rats per datapoint

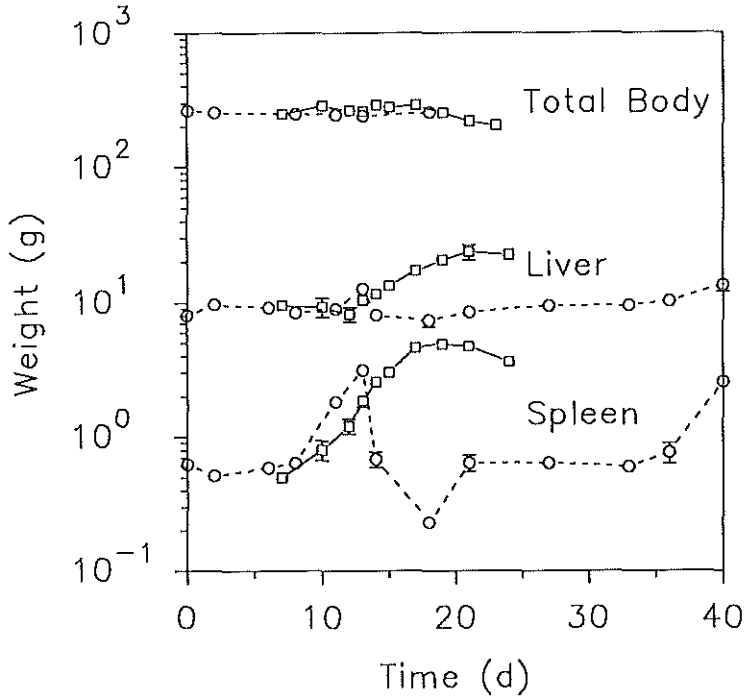


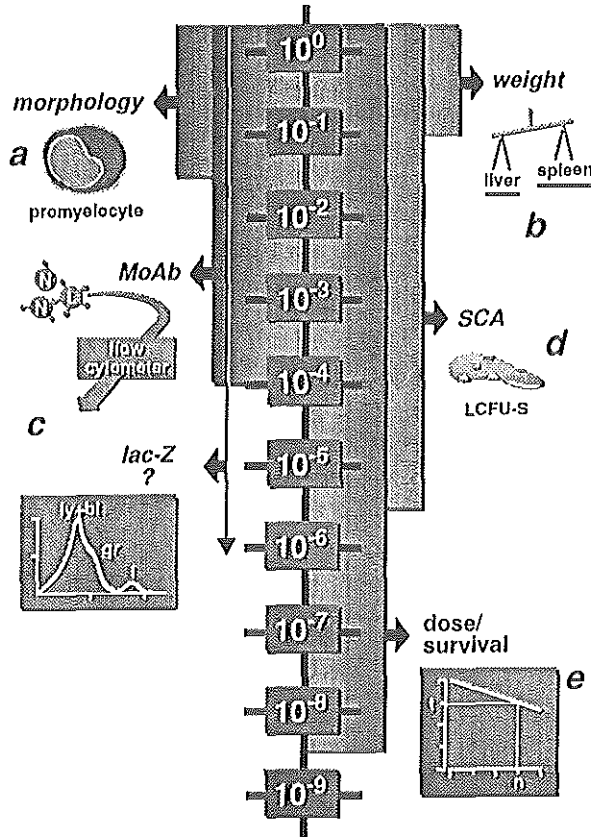
Table 2-1 Injection of known numbers of leukemic (L) cells ( $a, b$ ) results in a number of colonies to be counted on the spleen ( $f \times a, f \times b$ ). An experimental method to determine the dilution factor  $f$  is described in Appendix A. The other way round, when an unknown quantity of L-cells is injected and  $c$  colonies are counted, it can be deduced that this unknown quantity must have been  $c/f$  L-cells

### LCFU-S ASSAY

linear relation between the number of leukemic cells injected *i.v.* and the number of colonies on the spleen surface

| nr. of L-cells<br>on day 0 | nr. of spleen colonies<br>on day 20 |
|----------------------------|-------------------------------------|
| $a$                        | $f \times a$                        |
| $b$                        | $f \times b$                        |
| $c/f$                      | $c$                                 |

Fig. 2.4 Limits of detection of leukemic cells in the BNML [from Hagenbeek, 1992]

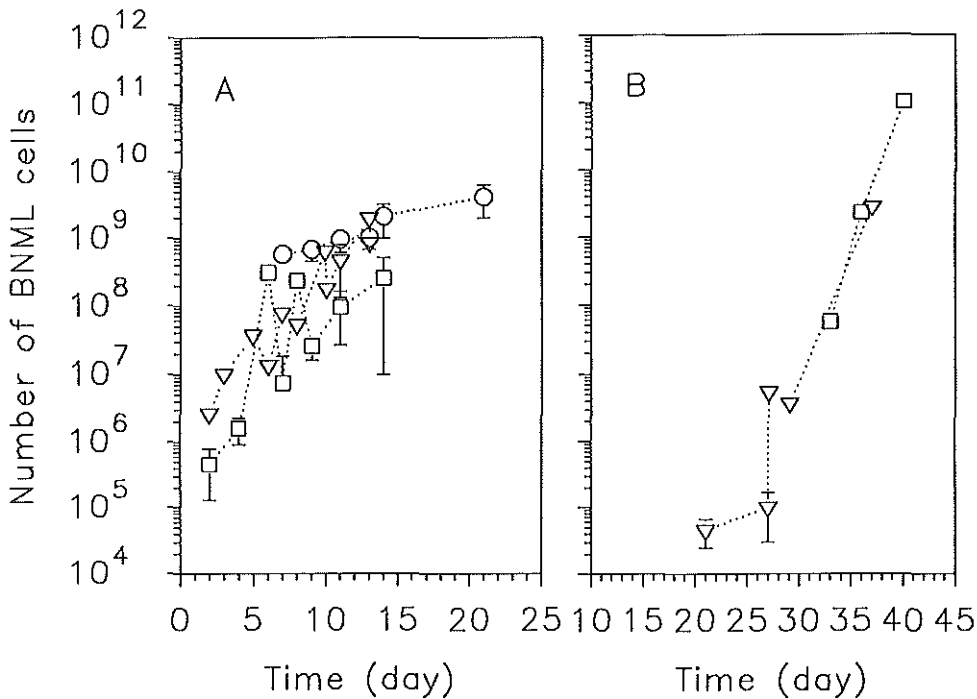


the spleen of the recipient rat, multiplied by 500. Subsequently, by assuming a homogeneous distribution of leukemic and normal cells in the organ from which the inoculated sample was taken, the leukemic cell load in that organ can be estimated. The detection limit of this method is at the level of one leukemic cell in  $10^5 - 10^6$  normal cells.

### 2.1.2.5 Survival Time Bio-Assay

This assay is based on the established relationship (Section 2.1.1.3) between the number of *i.v.* inoculated BNML cells and the median survival time. Several rats are inoculated with cells from a sample, e.g., a bone marrow sample from a leukemic rat that has been treated a certain time before with a cytostatic drug. After registering the rats' survival times the MdST is determined. The unknown

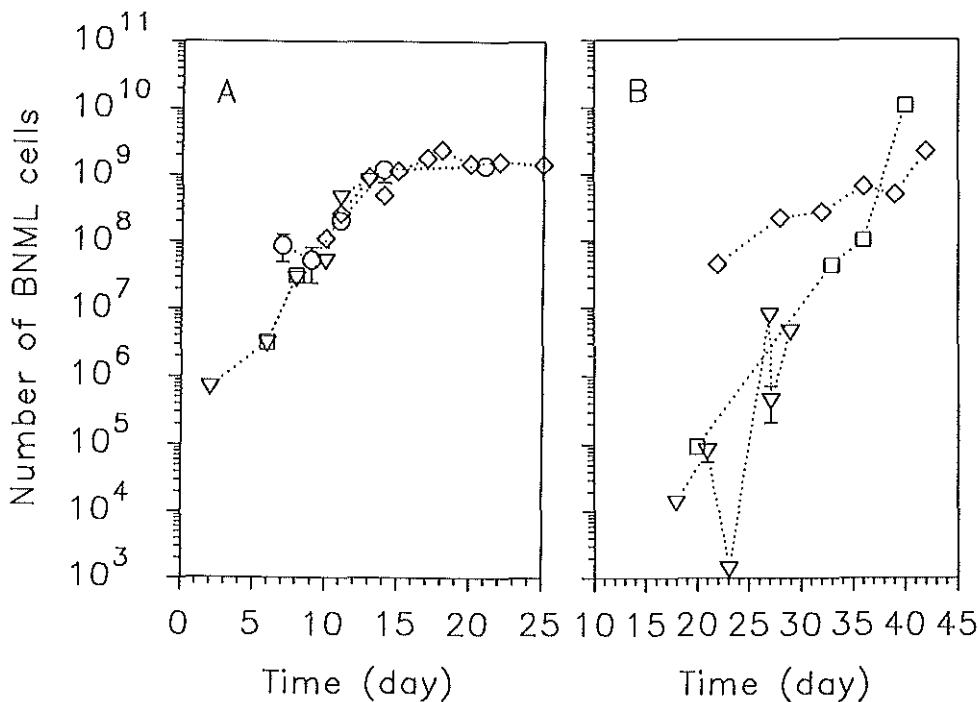
Fig. 2.5 Growth of leukemia in the liver after *i.v.* inoculation of  $10^7$  BNML cells on day 0, as observed with various detection methods during unperturbed growth (A) and during regrowth after cyclophosphamide treatment (single *i.p.* dose, 100 mg/kg) on day 13 (B). (o-o: microscopy counts;  $\nabla$ - $\nabla$ : flow cytometry;  $\square$ - $\square$ : survival time bio-assay. Bars represent 1 SD; n=3)



quantity of BNML cells in the donor rat's sample,  $C$ , then can be estimated using Eq.(2.1). Strictly speaking, the method is valid for the quantification of leukemic cell numbers between  $10^3$  and  $10^7$ , corresponding with MdSTs between 38 and 22 days. Outside this range the direct  $C$ -MdST relationship may no longer hold with good accuracy. If used for detection purposes only, however, the method is suitable for leukemic cell frequencies from  $10^{-8}$  to 1.

Figure 2.4 summarizes the detection limits of the methods mentioned above.

Fig. 2.6 Growth of leukemia in the spleen after *i.v.* inoculation of  $10^7$  BNML cells on day 0, as observed with various detection methods during unperturbed growth (A) and during regrowth after cyclophosphamide treatment (single *i.p.* dose, 100 mg/kg) on day 13 (B). ( $\diamond$ - $\diamond$ : organ weight;  $o$ - $o$ : microscopy counts;  $\nabla$ - $\nabla$ : flow cytometry;  $\square$ - $\square$ : survival time bio-assay). Bars represent 1 SD;  $n=3$ )



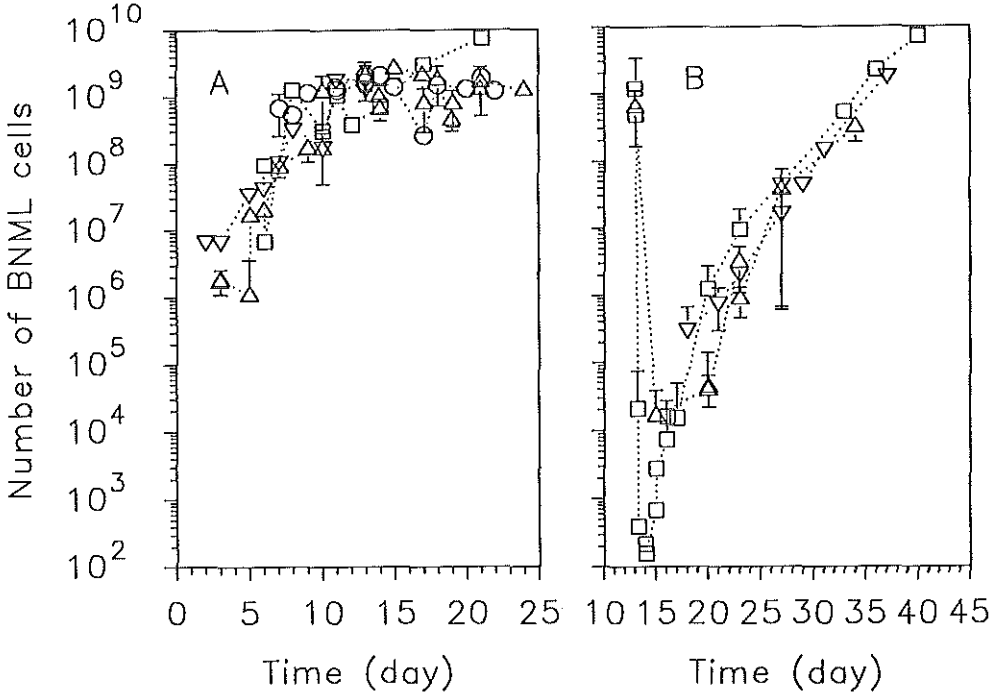
### 2.1.3 *IN VIVO* LEUKEMIA GROWTH

At various time points during 25 days after *i.v.* inoculation of  $10^7$  BNML cells groups of 5 BN rats were sacrificed and the leukemic cell loads in four main target organs—liver, spleen, bone marrow and blood—were determined. All of the detection and quantification methods discussed were used.

Actually, the leukemic cell numbers in liver and spleen samples were measured; the total leukemic cell loads in these organs were then calculated on the assumption of a homogeneous distribution of BNML and normal cells. The leukemic cell load in bone marrow is based on measurements in femoral bone marrow; it is assumed that femoral bone marrow accounts for 2.5% of the total amount of bone marrow [Colly et al., 1984a]. In peripheral blood the number of leukemic cells was determined per ml.



Fig. 2.7 Growth of leukemia in the bone marrow after *i.v.* inoculation of  $10^7$  BNML cells on day 0, as observed with various detection methods during unperturbed growth (A) and during regrowth after cyclophosphamide treatment (single *i.p.* dose, 100 mg/kg) on day 13 (B). (o-o: microscopy counts;  $\nabla$ - $\nabla$ : flow cytometry;  $\Delta$ - $\Delta$ : LCFU-S assay;  $\square$ - $\square$ : survival time bio-assay. Bars represent 1 SD; n=3)



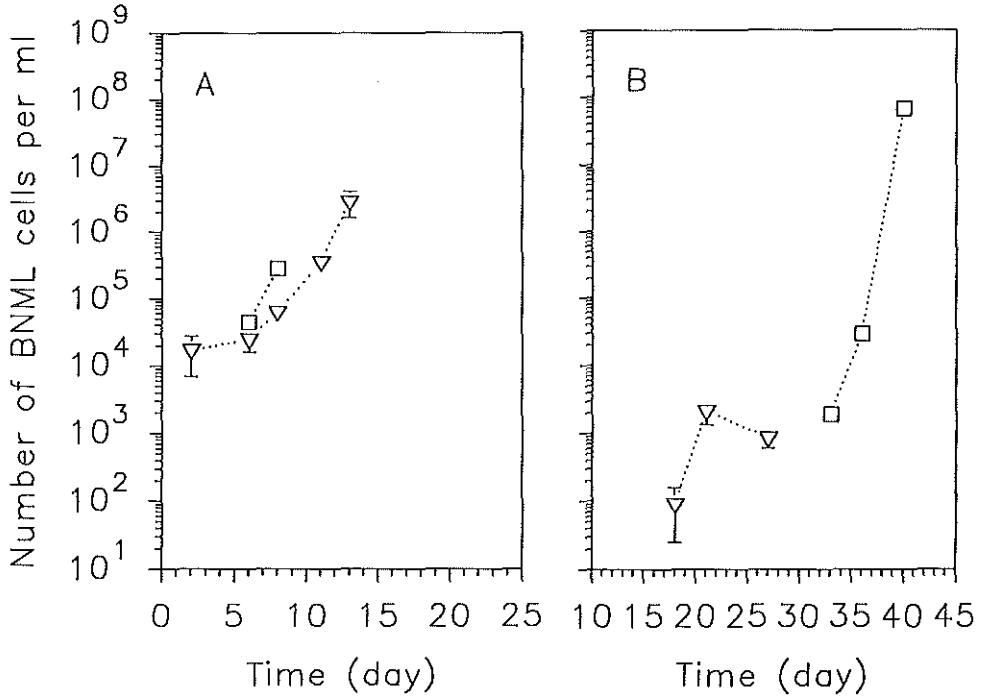
By repeating the experiments several times, with quite large time intervals in between, it was checked that the resulting growth patterns are reproducible.

Results are shown in panels A of Fig. 2.5 (liver), Fig. 2.6 (spleen), Fig. 2.7 (bone marrow) and Fig. 2.8 (peripheral blood). In panels B of the same figures results are presented from similar experiments, which examine leukemia regrowth after treatment rather than unperturbed leukemia growth. The animals were now followed for 42 days. Treatment consisted of a single *i.p.* dose of 100 mg/kg cyclophosphamide (CFA) on day 13.

For any random measurement time point, the spread in the observations is considerable, both for individual data from a same detection method and for data from different detection methods. General patterns in the time courses, however, can be recognized. They are similar for all organs. Concave time-courses of the data on the semi-log plots indicate decelerating growth rates.

The experimental data were used for modeling unperturbed and chemotherapy perturbed leukemia growth. Best fit parameter values of growth curves, cor-

Fig. 2.8 Growth of leukemia in the peripheral blood after *i.v.* inoculation of  $10^7$  BNML cells on day 0, as observed with various detection methods) during unperturbed growth (A) and during regrowth after cyclophosphamide treatment (single *i.p.* dose, 100 mg/kg) on day 13 (B). ( $\nabla$ - $\nabla$ : flow cytometry;  $\Delta$ - $\Delta$ : LCFU-S assay;  $\square$ - $\square$ : survival time bio-assay. Bars represent 1 SD; n=3)



responding with several different models, were determined and compared to identify the best description of the development of the leukemic cell population (see Chapter 2.2).

## 2.2 BNML Growth Perturbation by Cyclophosphamide<sup>1</sup>

At present, treatment of human acute leukemia often starts successfully with the induction of a complete remission by chemotherapy. Initial treatment reduces a typical load of  $10^{12}$  leukemic cells (1 kg) by at least a factor 100. As it is impossible to detect a tumor of 10 g or less by standard clinical cytologic methods, all that can be said is that the actual number of surviving cells may be anything between zero and  $10^{10}$ . Because no information is available on the residual tumor load, also called minimal residual disease (MRD), maintenance of the state of complete remission is still a major problem. To try and eradicate MRD either maintenance chemotherapy or high-dose chemotherapy (e.g., with cyclophosphamide, CFA) combined with total body irradiation (TBI) and bone marrow transplantation (BMT) is applied. Failure of these methods, i.e., not *all* clonogenic leukemic cells are eliminated, means a relapse of the disease eventually.

Therefore, it is very important to gain knowledge about the kinetics of growth of the leukemic cell population, especially in the "invisible" area of MRD. If changes in the tumor load, due to both natural (re)growth and under influence of chemotherapy, can be accurately quantified, then relapse can be predicted earlier and more effective treatment strategies can be designed.

A way to track down the time course of the size of the leukemic population is by performing computer simulation studies, using mathematical models for growth under both unperturbed and under therapy conditions. Various hypothetical growth curves can be tested for adequate description of data obtained from *in vivo* experiments. The best fitting curves may be used for extrapolation into the experimentally invisible area. As long as they do not contradict any other known physical facts, they may be considered as the most likely description of the growth properties of MRD.

In the past many investigators have employed mathematical models for tumor growth. It is generally assumed that a natural human tumor develops from a single transformed cell. Starting at time  $t = 0$  with a single cell the growth of a population may be modeled with a linear birth process. The population size,  $C$ , will increase exponentially with time if a constant probability of cell birth,  $p$ , is assumed [Iversen and Arley, 1950], i.e.,

---

<sup>1</sup>A manuscript has been published as:

F.W.Schultz, A.C.M.Martens and A.Hagenbeek (1987) Computer simulation of the progression of an acute myelocytic leukemia in the Brown Norway rat. *Comput Math Applic* 14:751-761, and in: M.Witten (ed) *Mathematical Models in Medicine*, *Advances in Mathematics in Computers and Medicine* 2 (1988) Pergamon Press, New York.

$$C(t) = \exp(p \cdot t). \quad (2.2)$$

This will still be the case if, more realistically, a death process is modeled as well, again with a constant probability [Neyman and Scott, 1967]. The growth rate will be

$$dC(t)/dt = GF \cdot C(t), \quad (2.3)$$

where the growth fraction GF is a constant depending on the characteristics of the specific tumor investigated (i.e., cell cycle time, fraction of resting cells, death rate).

In practice, experiments starting with a single cell generally will yield some distribution of population sizes at time t. At some time after initiation the tumor may even become extinct. Thus, to some extent tumor growth is a stochastic process, requiring stochastic models. Tumor initiation itself also may be modeled as a stochastic transformation process [Whittemore and Keller, 1978].

Also, at some time during their development most tumors appear to deviate from growing exponentially. The growth fraction obviously depends on population size. To account for this phenomenon many researchers have investigated various growth laws to explain their data, e.g., the logistic equation [Steel, 1977] or the Gompertz equation [Simpson-Herren and Lloyd, 1970]. The appropriateness of these functions is based on empirical curve fitting rather than on any underlying physical or biological arguments. Especially when a large range of tumor sizes is involved the Gompertzian law has been demonstrated to be the most applicable [Steel, 1977; Simpson-Herren and Lloyd, 1970; Laird, 1964; Hanson and Tier, 1982; Sullivan and Salmon, 1972]. The growth fraction then decreases with increasing tumor size according to

$$GF = A \cdot \ln(C_{\max}/C(t)), \quad (2.3)$$

where  $C_{\max}$  is the maximum population size for  $t \rightarrow \infty$ , and A is either a constant or, if an aspect of randomness in growth must be introduced, A may be replaced by a Gaussian white noise process [Smith and Tuckwell, 1974]. Effects of heredity, i.e., absolute randomness being restricted by the fact that daughter cells behave more like their mother than do their nieces, may be considered [Prajneshu, 1979]. For similar results with respect to the qualitative aspects tumor growth with a size dependent growth fraction also may be modeled by a non-Gompertz equation [Wette et al., 1974]. Again different models consider tumor size dependent birth and death rates [Dubin, 1976; Swan, 1977].

In the present study the growth characteristics of an acute myelocytic leukemia in various organs of the Brown Norway rat, and in the bone marrow

also the influence of chemotherapy, will be evaluated along the above indicated lines of curve fitting and simulations. The cell numbers dealt with are relatively large. The experimental data are assumed to reveal the development of the mean population size. Therefore, no stochastic effects will be considered, but only a population size dependent growth fraction.

## 2.2.1 MATERIALS AND METHODS

### 2.2.1.1 The Rat Leukemia Model (See also Chapters 1.9 and 2.1.1)

The Brown Norway acute myelocytic leukemia (BNML) was induced with 9,10-dimethyl-1,2-benzanthracene in a female Brown Norway (BN) rat. Within the BN/Bi/Rij rat strain BNML is transplantable by cellular transfer. As few as 25 *i.v.* inoculated BNML cells cause the development of leukemia in 50 % of the test animals [Hagenbeek and Martens, 1985]. The principal target organs are the liver, the spleen and the bone marrow. The leukemic cell load at the time of death amounts to  $2 \times 10^{10}$ . The fact that BNML behaves very much like human acute myelocytic leukemia (AML) explains the relevance of the disease for use in experimental tumor treatment studies. Some major properties are: a) a slow growth rate; b) the presence of clonogenic leukemic cells; c) a severe suppression of hemopoiesis owing to an absolute decrease in normal hemopoietic stem cells (CFU-S); d) a response to chemotherapy similar to that of AML.

### 2.2.1.2 The Cytostatic Agent Cyclophosphamide (CFA)

CFA was discovered in 1958 and since has been the most commonly used alkylating agent in studies of clinical and experimental cancer chemotherapy and immunosuppression [Friedman et al., 1979]. Applied in doses of 50-250 mg/kg the compound is metabolized by hepatic microsomal enzymes and thus activated into its cytotoxic form.

### 2.2.1.3 Experiments

Data concerning the growth pattern of BNML in the BN rat's bone marrow could be derived from the following experiments that had been conducted with 14 weeks old male BN/Bi/Rij rats.

a) *a dose—response experiment* (See Chapter 2.1.1.3 and Fig. 2.2).

b) *clonogenic leukemic stem cell assays (LCFU-S)* (See Chapter 2.1.2.4). At various times after *i.v.* inoculation with  $10^7$  BNML cells groups of five rats were sacrificed and the leukemic cell load in the bone marrow was determined in the following way. Femoral bone marrow was inoculated *i.v.* into recipient rats that were sacrificed 19 days later. It was experimentally determined—by an

assay based on the one for normal hemopoietic stem cells [Lahiri et al., 1970]—that, on average (with an accuracy of approximately 10%), one out of every 500 BNML cells present in an inoculum produces a spleen colony (seen as a white spot on the spleen surface). Thus, by counting the spleen colonies, the total number of BNML cells originally present in the donor rat's marrow can be calculated. It is assumed that a femur contains 2.5 % of the total bone marrow, and that the BNML cells in the bone marrow are homogeneously distributed.

To check the reproducibility of the results this experiment was repeated several times. Four times with and twice without the inclusion of chemotherapy consisting of a single *i.p.* bolus dose of 100 mg/kg CFA at day 13. For both cases median survival times were determined as well.

c) *dose-survival time bio-assay* (See Chapter 2.1.2.5). This experiment resembles the one described under b. The recipient rat is not sacrificed for spleen examination, but its time of death owing to leukemia is awaited. Using the dose—response relationship (Eq.(2.1) from experiment a), the number of BNML cells originally present in the donor rat's bone marrow can be calculated from the median survival time (MdST) of the recipient rats.

d) *flow cytometry* (See Chapter 2.1.2.3). Flow cytometry measurements have the advantage that no recipient rats are required for the quantification of the number of (residual) leukemic cells in a bone marrow sample. The donor rat's bone marrow is incubated in suspension with a monoclonal antibody (MCA), RM124, that binds to BNML cells [Martens and Hagenbeek, 1985]. This MCA is conjugated with fluorescein isothiocyanate (FITC), a dye that emits fluorescence when excited by laser light. When the cell suspension is run through a flow cytometer, in this way the fluorescing cells are recognized as the malignant ones. This allows their quantification with respect to the total number of bone marrow cells.

(Un)perturbed growth of leukemic cells in the liver and the spleen after  $10^7$  BNML cells inoculated *i.v.* was monitored using the same detection methods as for bone marrow. Perturbation consisted of a single rapid CFA dose (100 mg/kg *i.p.*) on day 13.

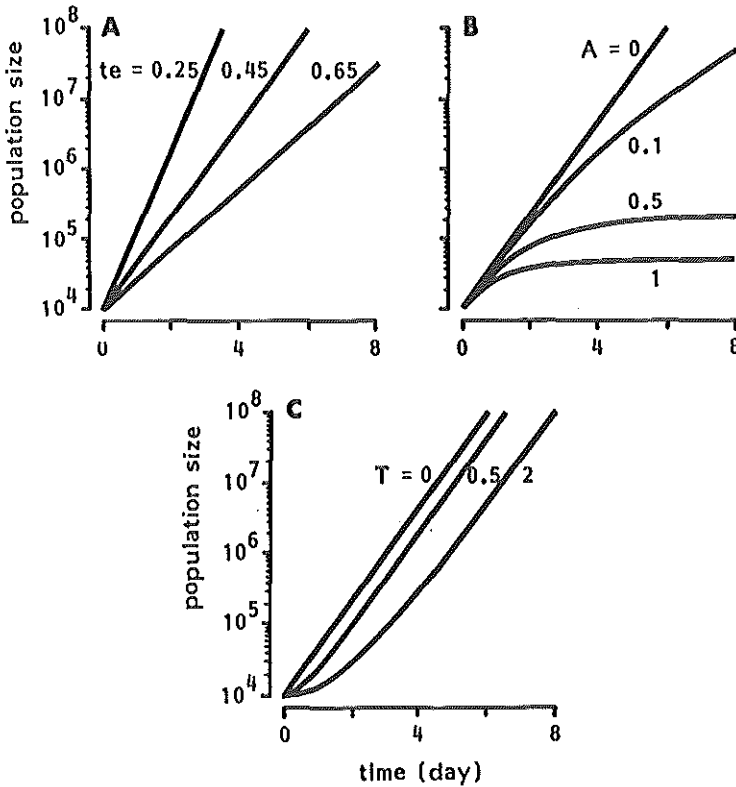
#### 2.2.1.4 Data Analysis; Basic Growth Curves

**Unperturbed Growth.** A general equation that expresses the growth of a cell population can be written as:

$$dC(t)/dt = GF(C) \cdot C(t), \quad (2.5)$$

in which  $dC/dt$  denotes the growth rate,  $C(t)$  the population size and  $GF(C)$  the growth fraction.  $GF(C)$  can be regarded as the fraction of the population that doubles its size during time interval  $(t, t+dt)$ , and includes many contributory

Fig. 2.9 Basic Growth Curves; A) exponential growth, at various doubling times,  $t_e$  (d); B) Gompertz growth, for several values of the retardation constant,  $A$  ( $d^{-1}$ ); and C) exponential growth with time delay,  $T$  (d)



factors like natural cell loss and tumor environment. If the growth fraction as a function of the population size is known, then Eq.(2.5) can be solved, yielding an analytical expression for the development of the population size with time  $t$ .

Plotting the LCFU-S datapoints on semi-log paper reveals that unperturbed growth is characterized by an exponential phase (constant population doubling time) followed by a plateau phase (constant population size). This growth might be empirically best described by either a contiguous exponential and Gompertz curve, or by a Gompertz curve alone [Hanson and Tier, 1982].

In *exponential* phase  $GF(C)$  is a constant,  $k_e$ , and so is the population doubling time, given by  $t_e = \ln 2/k_e$ . The growth rate is directly proportional to the population size and increases exponentially (Fig. 2.9a). Equation(2.5) becomes a *Gompertz* curve if  $GF(C)$  decreases exponentially, i.e.,

$$dGF(C)/dt = -A \cdot GF(C). \quad (2.6)$$

The larger the retardation constant A, the faster the curve reaches a plateau level (Fig. 2.9b). The growth rate first increases to a maximum value, then decreases to zero. A state of accelerating growth also can be simulated by making use of a *time delay*  $\tau$ :

$$GF(C) = k_e \cdot \{1 - \exp(-t/\tau)\}. \quad (2.7)$$

If  $\tau = 0$ , then GF reduces to  $k_e$ ; if  $\tau > 0$  and t becomes large, then the growth curve will tend to the original exponential curve, but shifted a distance  $\tau$  forwards in time (Fig. 2.9c).

**Perturbed Growth.** The influence of chemotherapy can be modeled by subtracting a growth inhibition term,  $D(C,t)$ , from Eq.(2.5). This term can be regarded as the cell quantity removed from the population during time increment  $dt$ . It depends on the therapy regimen and the tumor sensitivity, and includes environmental changes and augmented cell loss caused by the drug. If D is large enough the growth rate becomes negative and the tumor regresses. It was found for laboratory cell lines growing exponentially that a certain drug level kills a fixed percentage of the population. In clinical practice, where Gompertzian tumor growth is rather common, medium size tumors often are the most sensitive for chemotherapy. A hypothetical expression for D that agrees with these findings was derived by Norton and Simon [1977]. In a multiplicative way D is proportional to the level of therapy,  $L(t)$ , which comprises drug dose, route of administration, bioavailability and other pharmacokinetic factors; to the population size; and to the (unperturbed) growth fraction. Thus:

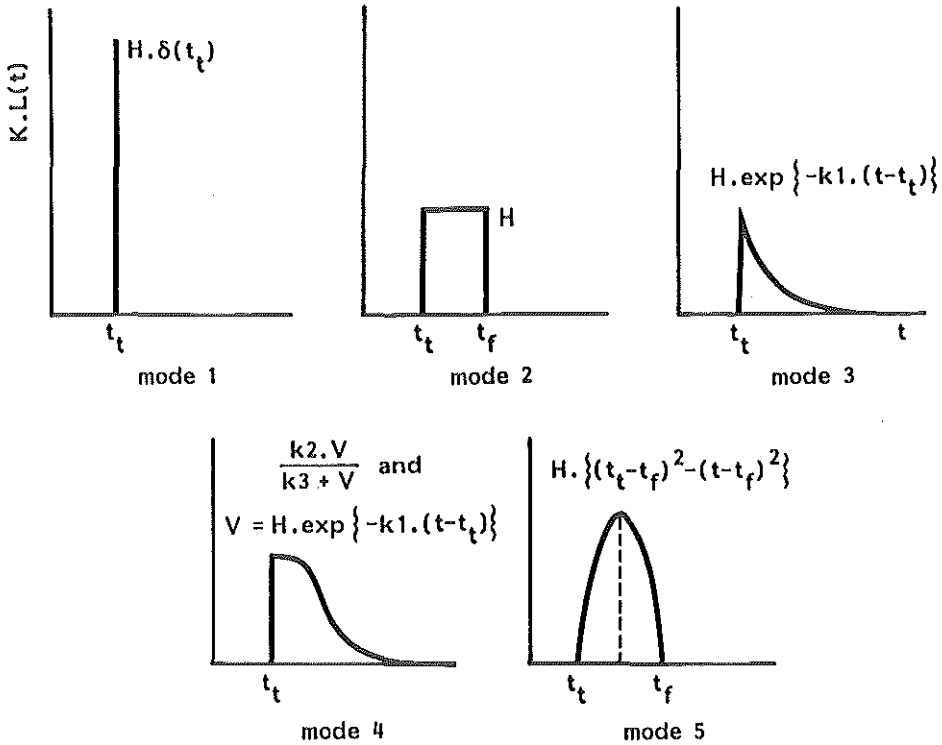
$$dC(t)/dt = GF(C) \cdot C(t) - K \cdot L(t) \cdot GF(C) \cdot C(t). \quad (2.8)$$

In exponential phase GF is a constant, so, for given level  $L(t)$ , growth inhibition in Eq.(2.8) becomes proportional to  $C(t)$ . In Gompertzian growth the relative effect of some given L is maximal for intermediate C, for  $GF \cdot C$  then is largest. Therefore, Eq.(2.8) is potentially in accordance with clinical experience.

Several assumptions can be made concerning the shape of the function  $L(t)$ . Rapid administration combined with a short plasma half-time (20' [Donelli et al., 1984]) of CFA allows an instantaneous drug effect model, in which  $L(t)$  is pulse-shaped. On the other hand CFA must first be metabolized by liver microsomes to become an active (alkylating) compound. Secondly, prolonged presence of CFA in the bone marrow has been demonstrated (e.g., 14 % of the given dose still after 12 h [Houghton et al., 1976]). Furthermore, cells show an



Fig. 2.10 Hypothetical functions to describe therapy level (Eq.(2.9)); 1) pulse (instantaneous); 2) block function; 3) exponential decay; 4) 2 and 3 combined; 5) parabolic increase and decrease (modes 2-5: gradual drug influence models)



enhanced rate of death for some three days after a 3 h exposure to CFA [Pohl J, personal communication, 1986]. So, longer lasting drug influence models also should be considered. For instance, the therapy level might remain constant for some time. In view of the nature of drug distribution processes in general (diffusion and possibly saturation effects) an exponentially decreasing  $L(t)$  is another obvious choice. Also, a combination might be adequate. Finally, the course of the drug level also might be well described with a parabolic increase and decrease. The model equations are, for subsequent modes:

$$\begin{aligned}
1 \quad K \cdot L(t) &= H \cdot \delta(t_f) \quad \text{for } t \geq t_f, \\
2 \quad K \cdot L(t) &= H \quad \text{for } t_f \leq t \leq t_p, \\
3 \quad K \cdot L(t) &= H \cdot \exp\{-k_1 \cdot (t-t_f)\} \quad \text{for } t \geq t_f, \\
4 \quad K \cdot L(t) &= k_2 \cdot H \cdot \exp\{-k_1 \cdot (t-t_f)\} / [k_3 + H \cdot \exp\{-k_1 \cdot (t-t_f)\}] \quad \text{for } t \geq t_f, \\
5 \quad K \cdot L(t) &= H \cdot \{(t_f - t_f)^2 - (t-t_f)^2\} \quad \text{for } t_f \leq t \leq 2 \cdot t_f - t_f,
\end{aligned} \tag{2.9}$$

where  $t_f$  is time of drug administration,  $\delta(t_f)$  is the unit pulse and  $H$ ,  $t_f$ ,  $k_1$ ,  $k_2$  and  $k_3$  are positive constants (model parameters). Outside the given time intervals  $K \cdot L(t)$  equals zero. The functions are shown in Fig. 2.10. Integration of  $K \cdot L(t)$  over time results in an "area under the curve" (AUC) that reflects the intensity and/or duration of the drug influence.

### 2.2.1.5 Data Analysis; Model Evaluation

**Curve Fitting.** For unperturbed growth the model equations (Eq.(2.5)) were solved analytically, using the basic exponential and contiguous Gompertz curves. The time course of the population size then is given by:

$$\begin{aligned}
C(t) &= C_g \cdot \exp\{\ln 2 \cdot (t-t_g)/t_e\} \quad \text{for } 0 \leq t \leq t_g, \\
C(t) &= C_g \cdot \exp\{\ln 2 \cdot [1 - \exp\{-A \cdot (t-t_g)\}] / (A \cdot t_e)\} \quad \text{for } t > t_g.
\end{aligned} \tag{2.10}$$

The model parameters are the doubling time in exponential phase,  $t_e$ ; the transition point: time  $t_g$  and population size  $C_g$ ; and the retardation constant  $A$ . If only Gompertzian growth is concerned, then  $t_g = 0$ . To find the optimum parameter values (considered as elements of the vector  $\mathbf{p}$ ) the curves (Eq.(2.10)) were fitted to the  $M$  observed datapoints  $C_{\text{obs}}(t_i)$  using a nonlinear least squares method. The computer algorithms for either a gradient or a grid search, starting from some initial estimate, were derived from [Bevington, 1969]. In minimizing the error function:

$$E = \sum W_i \cdot \{C_{\text{obs}} - C_{\text{calc}}(\mathbf{p})\}_i^2, \quad i = 1, \dots, M \tag{2.11}$$

three weighting options were available:  $W = 1$  (normal situation, equal weights);  $W = 1/C_{\text{obs}}(t_i)$  (statistical weight, favoring smaller observed values over larger ones);  $W = 1/\text{sd}_{\text{obs}}(t_i)$  (instrumental weight, putting emphasis on the more accurate data by looking at the observed standard deviations). For the separate optimization runs each option was evaluated and the one yielding the best fit was maintained.

To simulate chemotherapy the optimum parameter values for unperturbed growth were substituted in Eq.(2.8). Additional parameter values related to each particular mode of drug action, Eq.(2.9), were estimated by fitting curve:

$$C(t+\Delta t) = C(t) + \Delta t \cdot (dC(t)/dt) \quad (2.12)$$

to observed datapoints (where  $\Delta t = 0.1$  and  $dC/dt$  is given by Eq.(2.8)), using the same least squares routines.

**Goodness of Fit.** The growth curve that results in the best fit of the observed data is the best model—if it does not contradict any known physical facts. As a criterion for the fit the residual sum of squares, SSR can be used:

$$SSR = \sum \{dev_i^2 / var_i\}, \quad i = 1, \dots, M \quad (2.13)$$

where

$$dev_i = \{C_{obs} - C_{calc}(D)\}_i^2, \quad (2.14)$$

$$var_i = \{dev_i - [(\sum dev_i) / M]\}^2 / (M-1). \quad (2.15)$$

Low SSR values mean better fits. Also, especially to compare the fit of several curves to a same set of observations, the total correlation coefficient, TCC is a useful measure. As it is defined as:

$$TCC = \{1 - \sum dev_i^2 / \sum C_{obs,i}^2\}^{1/2}, \quad (2.16)$$

the closer TCC tends to 1, the better the fit.

In general, the simplest model that explains observed data adequately is to be preferred to any other adequate but more complex model. The Akaike Information Criterion, AIC, balances the goodness of fit and the model complexity by considering the number of observations, M; the SSR; and the number of model parameters, p, as follows [Akaike, 1974]:

$$AIC = M \cdot \ln(SSR) + 2 \cdot p. \quad (2.17)$$

Thus, the model with the lowest AIC value is to be preferred.

**Computation.** All programs were written in FORTRAN 5 and used on a DATA GENERAL Eclipse MV/10000 minicomputer.

Fig. 2.11 Bone marrow: Fitted growth curves and datapoints (with standard deviations) from a single series of LCFU-S experiments; instantaneous CFA influence model unperturbed growth

A: Exponential + Gompertz, TCC = 0.98895, SSR = 4.00,  $t_0 = 8.20 \pm 1.88$  d,  
 $C_0 = 1.24 \times 10^8 \pm 1.25 \times 10^8$ ,  $t_0 = 0.83 \pm 0.31$  d,  $A = 0.19 \pm 0.23$  (1/d);

B: Gompertz, TCC = .98480, SSR = 4.70,  $t_0 = 0.0$  d,  $C_0 = 3.07 \times 10^3 \pm 3.03 \times 10^3$ ,  
 $t_0 = 0.30 \pm 0.03$  d,  $A = 0.15 \pm 0.02$  (1/d);

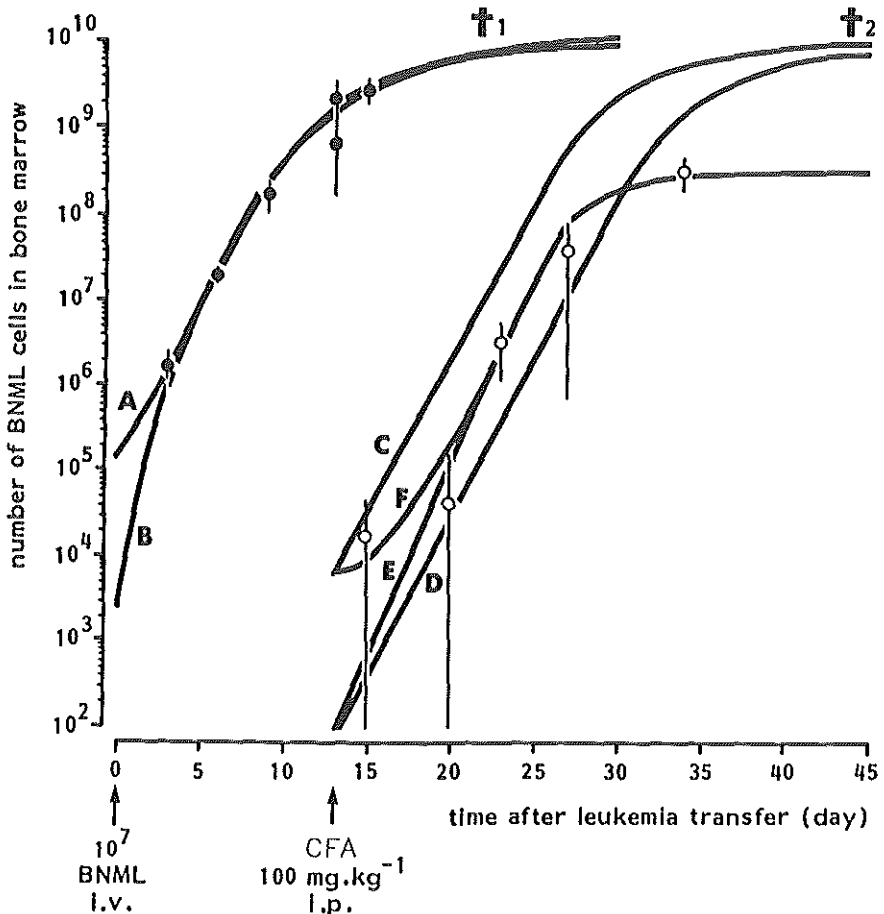
CFA treated

C: TCC < 0.00001, SSR = 5.30, 5.5 LCK, regrowth as A;

D: TCC < 0.00001, SSR = 5.30, 7.4 LCK (correction for MdST), regrowth as A;

E: Exponential + Gompertz, TCC = 0.98367, SSR = 4.06, 7.36 LCK,  $t_0 = 24.85 \pm 1.27$  d,  
 $C_0 = 1.45 \times 10^7 \pm 1.37 \times 10^7$ ,  $t_0 = 0.67 \pm 0.32$  d,  $A = 0.34 \pm 0.21$  (1/d);

F: TCC = 0.98367, SSR = 4.05, as E, 4.5 d time delay to correct for 5.5 LCK



Test Data. To test the least squares optimization routines' performance with respect to the reliability of the resulting parameter values two sets of test data were created. Twelve points, equidistant in the range 3-25 d, and seven non-

equidistant points, were calculated from Eq.(2.10), using  $t_e = 0.75$  d,  $t_g = 8.5$  d,  $C_g = 0.2 \times 10^9$  and  $A = 0.45$  d<sup>-1</sup>. Reconstruction of the original curve from these sets was attempted, starting with initial parameter estimates both 10 % above and below true values. Good fits were obtained (TCC > 0.9999). Some starting point dependency was observed. Parameter uncertainty (the estimated standard deviations can be rather large) decreases slightly with increasing number of observations. The optimum estimates for  $t_e$  and  $t_g$  as returned by the routines tend to their true values, larger relative deviations occur in  $C_g$  (10 %) and, to lesser extent, in  $A$  (5 %).

## 2.2.2 RESULTS AND DISCUSSION

The observations in the bone marrow pertaining to a single set of LCFU-S experiments were taken first to find the best descriptive model for unperturbed growth. Next, using this model, the various functions describing drug influence were evaluated. The resulting best fitting models for growth and regrowth were subsequently examined with respect to the datapoints from all LCFU-S experiments taken together. The results are presented below.

### 2.2.2.1 Single LCFU-S Dataset

**Unperturbed Growth of BNML in Bone Marrow.** The datapoints were fitted with an exponential and contiguous Gompertz curve (EG), and a Gompertz curve alone (G). The results are shown in Fig. 2.11, curves A and B, respectively. The EG curve yields the best fit in terms of TCC and SSR. The population starts with  $1.3 \times 10^5$  cells (1.3 % of inoculum) and at time of death (day 22) has grown to  $7.3 \times 10^9$  (18.3 % of overall BNML burden at death). The G curve starts with very few cells (0.03 % of dose), and, though this model also fits the data, a main (biological) argument that can be raised against its validity is that the initial population doubling time amounts to 7.2 h, whereas the cell cycle time is almost twice as large (14 h according to previous autoradiography experiments). Being too fast in early phase, this model is not further considered.

**Perturbed Growth; Instantaneous Drug Influence Model.** First two simulations were performed, based on the assumption that after instantaneous size reduction the population will regrow in a way identical with the unperturbed pattern.

CFA treatment at day 13 prolongs the rats' life by 22 days (MdST = 44 d instead of 22 d). According to the experimentally established dose response (Eq.(2.1))—i.e., every 4 d increase in lifespan means a factor of 10 reduction in cell number—this should correspond to a leukemic cell load reduction of 5.5 decades (= 22/4), or 5.5 log cell kill (LCK). In other words, at day 13 the cell

population drops from  $1.7 \times 10^9$  to  $8.5 \times 10^3$ . Subsequent regrowth identical with growth (curve C, Fig. 2.11) results in an expected MdST = 38.8 d.

To match expected and observed MdST a second simulation shifts the growth curve (D) to the right. This, however, necessitates 7.4 LCK, which is rather large. Furthermore, neither simulation fits the observed datapoints.

Abandoning the "regrowth equals growth" hypothesis an EG curve fitted to the data (E, Fig. 2.11) proves much better although 7.4 LCK is found here too. By allowing a time delay,  $\tau = 4.5$  d in early exponential phase this value can be reset to 5.5 (curve F), without improving the goodness of fit. In this model regrowth after chemotherapy is faster than unperturbed growth, the plateau phase is reached sooner and death occurs for fewer BNML cells.

**Perturbed Growth; Gradual Drug Influence Models.** The hypothetical functions that describe gradual drug influence (Fig. 2.10, modes 2 through 5) were tested by fitting them to the observed datapoints, using the unperturbed EG curve as basis. For each mode the best fits are shown in Fig. 2.12, in sequence curves G, H, I, J. Mode 2, constant therapy level for some time, is the least satisfactory. The other modes do not differ very much in goodness of fit. The TCC, SSR and AIC criteria appear to be not quite unanimous.

For intermediate cell numbers all curves lie rather to the right of the observations. Also, the predicted MdST lies beyond the observed value, unless death occurs at a somewhat lower malignant cell load after treatment than after unperturbed growth. By modifying the growth equation (Eq.(2.8)) improvement (left shift) can be achieved: gradual drug influence is modeled with therapy level,  $K \cdot L(t)$ , according to mode 4, but the GF factor is omitted in the second right-hand term and it is made a function of C times a positive constant F,  $F > 1$ , in the first term on the right:

$$dC(t)/dt = GF(F \cdot C) \cdot C(t) - K \cdot L(t) \cdot C(t). \quad (2.18)$$

Thus, by the former modification, a given therapy level  $L(t)$  causes cell loss proportional to the population size  $C(t)$  (constant percentage kill). The biological meaning of the latter modification is not yet clear: after therapy the GF at a certain population size equals the GF of an F times larger unperturbed population. In other words, therapy makes the transition to Gompertzian growth occur sooner, and the plateau phase level lower. Curve K in Fig. 2.12 shows the best fit of the last model to the datapoints.

### 2.2.2.2 All LCFU-S Datasets Combined

The procedure as described above was repeated with all datapoints acquired from several similar experiments, whose purpose was the confirmation of the

Fig. 2.12 Bone marrow: Fitted growth curves and datapoints (with standard deviations) from a single series of LCFU-S experiments; gradual CFA influence models unperturbed growth

A: see Fig. 2.11;

CFA treated

G: TCC=0.95844, SSR=6.53,  $K \cdot L(t) = 5.74$ ,  $13 < t < 16.7$ , AUC = 21.2;

H: TCC=0.99335, SSR=6.34,  $K \cdot L(T) = 8.58 \cdot \exp\{-0.4 \cdot (t-13)\}$ , AUC = 21.4;

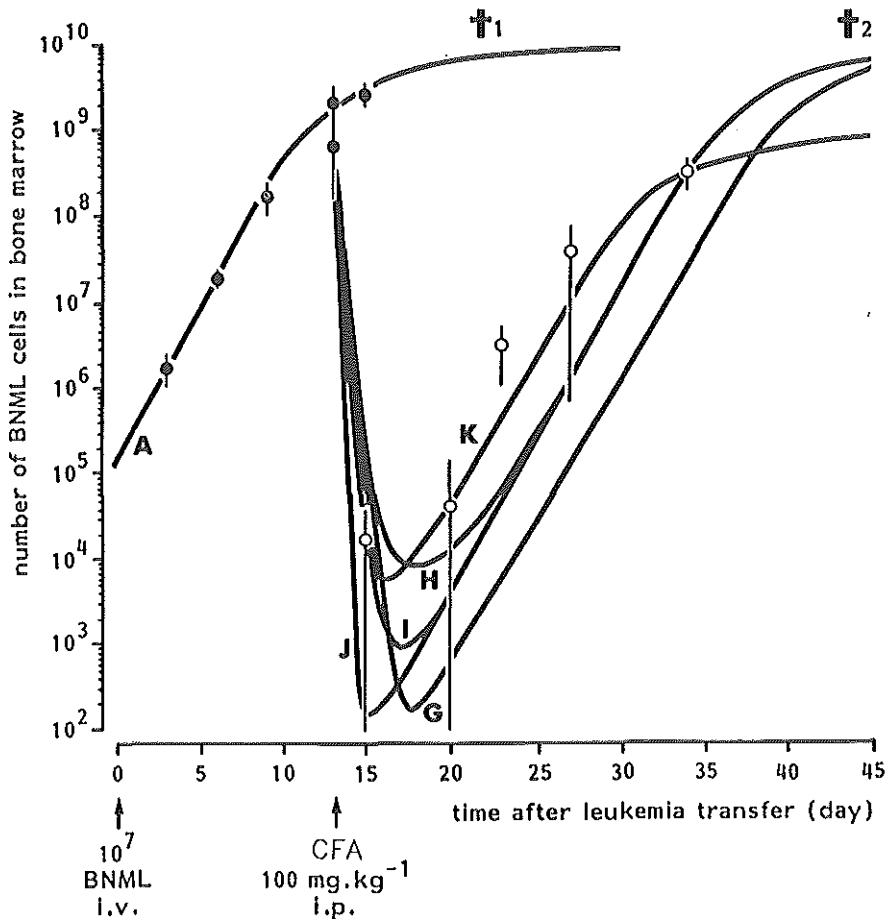
I: TCC=0.99333, SSR=6.57,

$K \cdot L(t) = 8.61 \cdot 15 \cdot \exp\{-1.2 \cdot (t-13)\} / \{1 + 15 \cdot \exp\{-1.2 \cdot (t-13)\}\}$ , AUC = 19.9;

J: TCC=0.99327, SSR=6.12,  $K \cdot L(t) = 8 \cdot \{(13-14.2)^2 - (t-14.2)^2\}$ , AUC = 16.1;

K: TCC=0.99733, SSR=4.90, Eq.(2.13),  $F = 10.02$ , max 5.8 LCK at day 16.5,

$K \cdot L(t) = 5.12 \cdot 86.01 \cdot \exp\{-1.6 \cdot (t-13)\} / \{1.5 + 86.01 \cdot \exp\{-1.6 \cdot (t-13)\}\}$ , AUC = 13.0



slope of the growth curve in exponential phase and of the level of the plateau phase. The different datasets (Fig. 2.13) show some shift along the time axis, probably owing to variance in the inoculation circumstances. As development of leukemic cells in the spleen and bone marrow is likely to correlate independent

Fig. 2.13 Bone marrow: Fitted growth curves and datapoints (with their standard deviations) from all series of LCFU-S experiments  
unperturbed growth

L: Exponential + Gompertz,  $TCC=0.88075$ ,  $SSR=22.1$ ,  $t_g=8.60 \pm 3.35$  d,  
 $C_g=1.82 \times 10^8 \pm 3.76 \times 10^8$ ,  $t_0=0.78 \pm 0.40$  d,  $A=0.44 \pm 0.30$  (1/d);

CFA treated, instantaneous model

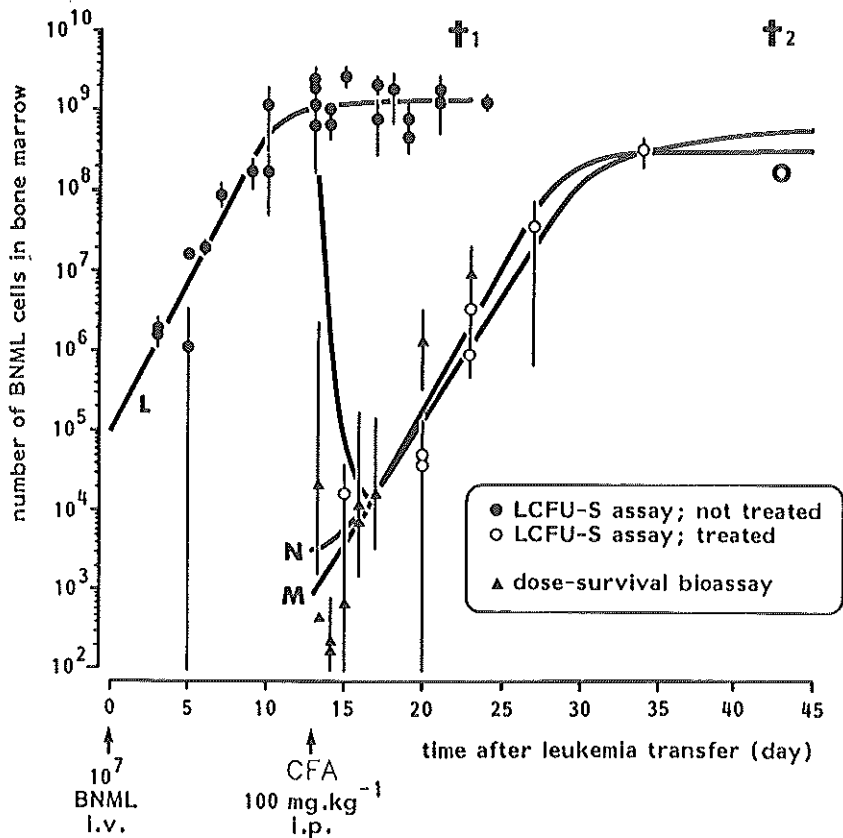
M: Exponential + Gompertz,  $TCC=0.85023$  (overall: 0.88030),  $SSR=7.82$  (overall: 29.8), 6.03 LCK,  $t_g=27.40 \pm 5.90$  d,  $C_g=2.70 \times 10^7 \pm 5.90 \times 10^7$ ,  $t_0=0.97 \pm 0.53$  d,  $A=0.23 \pm 0.54$  (1/d);

N:  $TCC=0.85023$  (0.88030),  $SSR=7.82$  (29.8), as M, time delay  $T=1.7$  d to correct for 5.5 LCK;

CFA treated, gradual model

O:  $TCC=0.84915$  (0.88028),  $SSR=7.99$  (21.4), Eq.(2.14),  $F=4.39$ ,

$K \cdot L(t) = 4.78 \cdot 86.02 \cdot \exp\{-1.6 \cdot (t-13)\} / [1.5 + 86.02 \cdot \exp\{-1.6 \cdot (t-13)\}]$ ,  $AUC=12.2$



on inoculum size, using spleen weight as reference in future experiments, rather than absolute time, may result in better reproducibility.

In exponential phase the EG curve that fits the joined unperturbed data best



(Fig. 2.13, curve L) does not differ much in slope, nor in predicted day zero population size, from the best single dataset fit. A difference exists in the plateau phases, which now attains a lower level with  $1.3 \times 10^9$  BNML cells in the marrow at death (3 % of total burden).

Chemotherapy again is modeled as an instantaneous or a gradual event. In the former case, by fitting a EG curve to the perturbed datapoints, 6.0 LCK results at day 13. Introducing a time delay,  $\tau = 1.72$  days in the initial exponential phase allows the size reduction to stop at a 5.5 LCK value, again without effect on the goodness of fit (Fig. 2.13, curves M and N).

The necessity for the time delay is not immediately apparent from the shown data. The 5.5 value is based on the observed delay in MdST after therapy and thus concerns what happens on average in the total animal. It might be possible that the drug effect is larger in the bone marrow, and accordingly smaller in, say, the liver (an organ that by nature better copes with toxic compounds). On the other hand preliminary data from running experiments with both chemo- and radiotherapy combined suggest that the LCK cannot be this large, otherwise many more 'cures' should have been seen.

In contrast to the single dataset case, regrowth after therapy here appears to be slower ( $t_e = 0.97$  d) than unperturbed growth ( $t_e = 0.78$  d). Though the standard deviations of the observations are rather large, the difference in doubling time therefore not very significant, this phenomenon would agree with the reported greater CFA sensitivity of rapidly proliferating cells [Dewys, 1972]. It would imply that the BNML cells are heterogeneous with respect to cell cycle time, faster cycling cells preferably being killed. This hypothesis may be tested by reinoculating rats with BNML cells that have survived CFA treatment, and see whether the growth kinetics differ from the growth pattern of non-treated cells. If not, another hypothesis is that cellular growth properties themselves remain constant, but greater cell loss occurs in the treated rat because some natural defense mechanism (immunology?) has had time to develop. CFA influence on the tumor environment also may play some role.

To model gradual drug influence, mode 4 therapy level was chosen in combination with the modified growth equation (Eq.(2.18)). The resulting curve (Fig. 2.13, O) fits the data only slightly poorer than the instantaneous model (curve N). The difference in goodness of fit is even smaller for the whole curves (fit to both perturbed and unperturbed data).

Table 2-2 summarizes again all curves that, in their category, fitted the LCFU-S in bone marrow data best. It shows the values of the goodness of fit criteria.

### 2.2.2.3 The Dose-Survival Bio-assay and Flow Cytometry Data

The dose-survival bio-assay experiments were conducted to yield more informa-

TABLE 2-2 GOODNESS OF FIT CRITERIA OF CURVES FITTED TO LCFU-S DATAPOINTS (BONE MARROW)

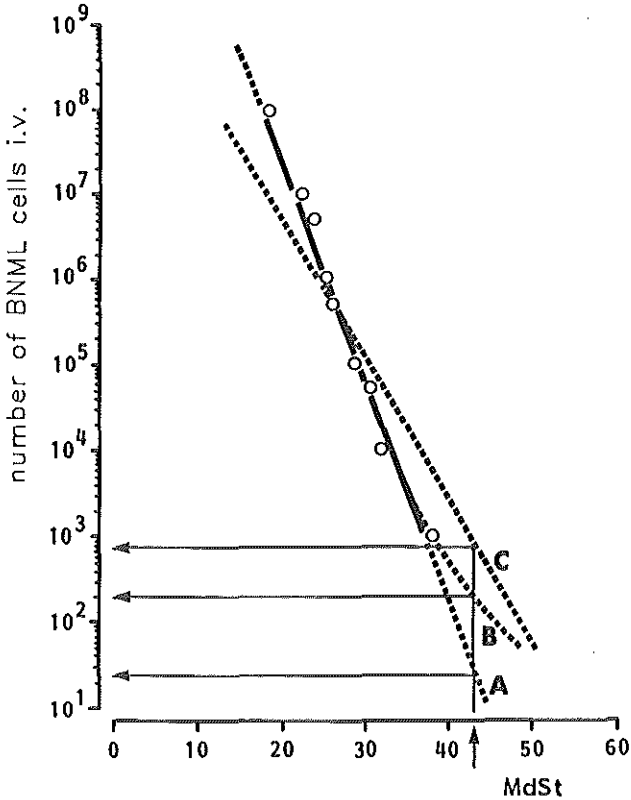
See Figs. 2.11 through 2.13;

EG: exponential + Gompertz; EGT: exponential + Gompertz + time delay;  
 In: instantaneous drug effect; Gr: gradual drug effect; mode numbers: see Eq.(2.9) (mode 4a: Eq.(2.18));  
 TCC: total correlation coefficient (Eq.(2.16)); SSR: sum of squared residuals (Eq.(2.13)); AIC: Akaike's Information Criterion (Eq.(2.17))

| curve | type | drug mode | nr of parameters | nr of datapoints | TCC      | SSR   | AIC    |
|-------|------|-----------|------------------|------------------|----------|-------|--------|
| A     | EG   | -         | 4                | 6                | 0.98895  | 4.00  | 16.32  |
| B     | G    | -         | 4                | 6                | 0.98480  | 4.70  | 17.28  |
| C     | EG   | In, 1     | 5                | 5                | <0.00001 | 5.30  | 18.34  |
| D     | EG   | In, 1     | 5                | 5                | <0.00001 | 5.30  | 18.34  |
| E     | EG   | In, 1     | 4                | 5                | 0.98367  | 4.06  | 15.00  |
| F     | EGT  | In, 1     | 5                | 5                | 0.98367  | 4.05  | 16.99  |
| G     | EG   | Gr, 2     | 6                | 5                | 0.95844  | 6.53  | 21.38  |
| H     | EG   | Gr, 3     | 6                | 5                | 0.99335  | 6.34  | 21.23  |
| I     | EG   | Gr, 4     | 8                | 5                | 0.99333  | 6.57  | 25.41  |
| J     | EG   | Gr, 5     | 6                | 5                | 0.99327  | 6.12  | 21.05  |
| K     | EG   | Gr, 4a    | 9                | 5                | 0.99733  | 4.90  | 25.95  |
| L     | EG   | -         | 4                | 25               | 0.88075  | 22.10 | 85.39  |
| M     | EG   | In, 1     | 4                | 7                | 0.85023  | 7.82  | 22.40  |
| N     | EGT  | In, 1     | 5                | 7                | 0.85023  | 7.82  | 24.40  |
| O     | EG   | Gr, 4a    | 9                | 7                | 0.84915  | 7.99  | 32.55  |
| L+M   | EG   | In, 1     | 8                | 32               | 0.88030  | 29.80 | 124.62 |
| L+N   | EGT  | In, 1     | 9                | 32               | 0.88030  | 29.80 | 126.62 |
| L+O   | EG   | Gr, 4a    | 9                | 32               | 0.88028  | 21.40 | 116.03 |

tion on the residual malignant cell load just after therapy. As the method is able to detect BNML cells at lower frequencies compared with the LCFU-S assay a convincing justification for introducing a time delay might be found. However, the results, also plotted in Fig. 2.13, show a very steep descent of the growth curve, suggesting an instantaneous rather than gradual chemotherapy effect. The reduction in population size is high (7 LCK) and, in contrast with the LCFU-S data, subsequent regrowth is very fast. Too fast in fact to be explained by proliferation of the surviving BNML cells alone; the population doubling time of 12 h is less than the cell cycle time (14 h). Besides, it disagrees with the expected CFA effects of enhanced death and greater kill of cells with short cycle times. An explanation might be a BNML cell repopulation by migration from

Fig. 2.14 The inverse of Fig.2.1: Inoculated BNML cell population size related to median survival time (MdSt); A) linear regression of the datapoints (o) yields Eq.(2.1). For some high MdSt value a larger BNML load is found if: B) the true relationship deviates from the extrapolated regression line A); or C) another (linear) relationship is valid in the post-treatment situation



other organs like liver and spleen. Therefore, the BNML growth kinetics in those organs must be examined first.

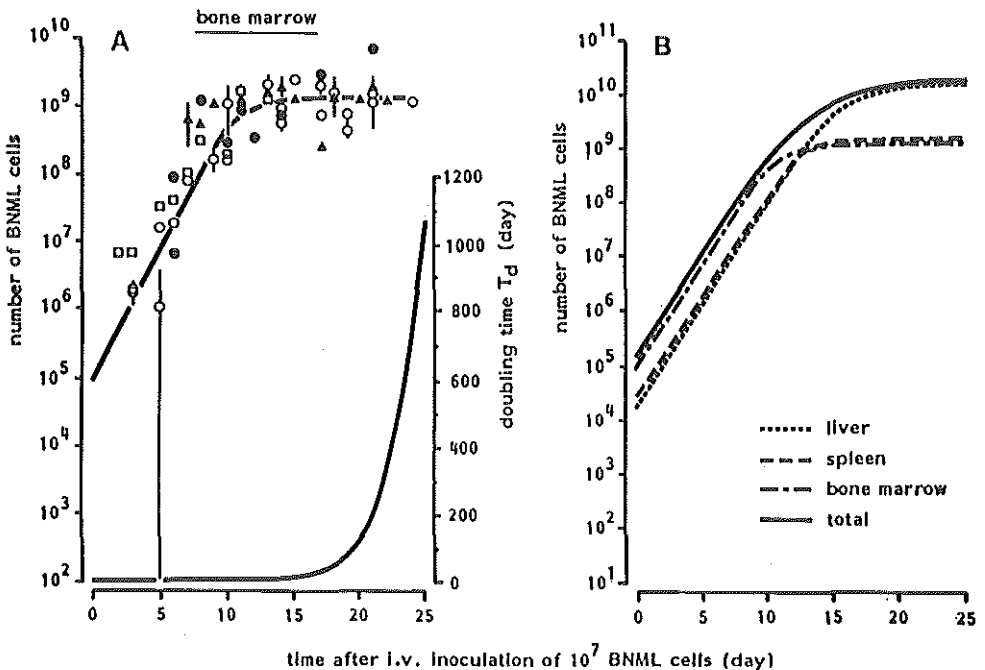
The discrepancy, however, between the LCFU-S and bio-assay data remains to be explained. The results of the flow cytometry measurements seem to confirm the LCFU-S data (Fig. 2.13). Of course, they are available only in the region of relatively large population sizes. Still, two of the three independent experimental methods yielding comparable results, this throws doubt on the survival time bio-assay data.

Perhaps the straight line relationship between MdSt and the log of the inoculated cell dose has been extrapolated too far. That might make that the cell doses calculated for large MdSTs are underestimated (Fig. 2.14). Another explanation for the bio-assay data being so unexpectedly far below the LCFU-S

TABLE 2-3 UNPERTURBED GROWTH CURVE PARAMETER VALUES (EQ. 2.10)

|             | $t_e$ (d) | $t_g$ (d) | $C_g \times 10^8$ | $A$ (d <sup>-1</sup> ) |
|-------------|-----------|-----------|-------------------|------------------------|
| liver       | 0.8467    | 13.308    | 14.84             | 0.3369                 |
| spleen      | 0.8120    | 11.900    | 4.91              | 0.7294                 |
| bone marrow | 0.7854    | 8.605     | 1.82              | 0.4394                 |

Fig. 2.15 A) BNML growth curve for bone marrow, as in Fig. 2.13. More datapoints ( $\Delta$ - $\Delta$  microscopy counts,  $\square$ - $\square$  flow cytometry,  $\circ$ - $\circ$  LCFU-S assay,  $\bullet$ - $\bullet$  survival time bio-assay) and exponential increase of population doubling time are shown; B) Fitted curves for unperturbed growth of BNML cells in liver (short dash), spleen (long dash) and bone marrow (dash-dot) after inoculation of  $10^7$  BNML cells *i.v.* on day zero. Total population growth (solid line) is estimated by adding the organ values at each time point. On day zero the population starts with  $1.5 \times 10^5$  proliferating cells. This is about half the number of injected ED<sub>50</sub> units (of 25 cells each; see Chapter 6.3)



results is, that the mentioned dose—response relationship may not hold true for the CFA-treated animals. Using this relationship when the actual (CFA-treated) BNML cell population grows at a slower rate, also yields an underestimation of the cell population size (Fig. 2.14). This last explanation is supported by the fact

that it implies that for low MdSTs the calculated cell population sizes should be overestimations, as compared to the measured LCFU-S data. This indeed appears to be the case.

#### 2.2.2.4 Unperturbed BNML Growth in Liver and Spleen

Parameters of the unperturbed growth curves (Eq. 2.10) for liver and spleen were determined by fitting the respective observed datapoints. Results are listed in Table 2-3. The fitted curves are shown in Fig. 2.15. The variation in doubling time during exponential phase is not large; in bone marrow the population grows faster than in the liver. Transition to Gompertz phase occurs first in bone marrow, perhaps due to the more rigid spatial confinement in this organ, then in the spleen and last in the liver. In the spleen the plateau phase is reached relatively soon after the transition, the retardation constant A being much larger in this organ as compared to the other organs. The increase in population size continues longest in the liver.

Neglecting growth in other organs, the total population growth is determined by adding the numbers of BNML cells in liver, spleen and bone marrow for each time point. The total BNML cell population tends to grow to a plateau level of  $1.98 \times 10^{10}$  cells, the contributions of liver, spleen and bone marrow being  $1.68 \times 10^{10}$ ,  $1.58 \times 10^9$  and  $1.36 \times 10^9$  cells, respectively. The initial size of the population that starts growing on day zero appears to amount to about only 1.4% of the inoculum.

**2.2.2.5 In Conclusion** it can be said that unperturbed growth of the BNML population in the BN rat bone marrow, initiated by *i.v.* inoculation of BNML cells, is characterized by an exponential phase followed by a Gompertzian phase. This is also true for unperturbed growth in liver and spleen. Different parameter values apply for the different organs. Only a very small fraction of the inoculated cells grows out.

In the bone marrow, the effect of a single high dose of CFA is an instantaneous reduction in population size of some 6 decades (logs), after which a similar regrowth pattern follows. The regrowth rate after treatment may be slightly lower. The necessity of introducing a time delay during initial residual disease, or just the opposite, i.e., modeling cell import from elsewhere, has not yet clearly been demonstrated. Explanations for the discrepancy between the LCFU-S and bio-assay data must still be verified as well.

## 2.3 Cell Cycle Specificity of AMSA? (A Simulation Study)<sup>1</sup>

Within the framework of a larger study [Hagenbeek and Martens, 1986] a series of *in vivo* experiments involved testing the leukemic cell kill capability of the cytostatic drug AMSA (acridinyl anisidine). The drug, which intercalates into DNA [Jehn and Heinemann, 1991] and is nowadays clinically applied for reinduction of remission in acute nonlymphoblastic leukemia [Freund et al., 1991; Miller et al., 1991], was given either as a one-time dose or in a fractional way. The survival of the leukemic cells was measured some time later. The apparent non-constancy of the cell kill pattern in the case of fractionation gave reason to suspect that cell cycle phase specificity possibly plays a role. In other words, the drug may preferentially kill cells when they are at a certain stage on their maturation pathway. The present simulation study investigates this hypothesis.

### 2.3.1 METHODS

#### 2.3.1.1 The Rat Leukemia Model (BNML) (See Chapter 1.9.2)

Some time ago a myelocytic leukemia (BNML) was developed that is transplantable in the Brown Norway (BN) rat by cellular transfer. Its characteristics and those of human acute myeloid leukemia (AML) are similar to the extent that the former disease can be considered a very relevant laboratory model for the latter. It is, for example, very useful for evaluating chemotherapy regimens.

#### 2.3.1.2 Experimental Animals (See Chapter 1.9.1)

The experiments were conducted with male BN rats (Rijswijk inbred strain). Age varied between 13 and 16 weeks (240 g mean body weight).

Leukemia was induced by *i.v.* inoculation of  $10^7$  viable BNML cells on day zero. It was either allowed to develop without interference, yielding the death of the animals on day 22 (median), or drug therapy was given, starting on day 13.

#### 2.3.1.3 LCFU-S Assay (See Chapter 2.1.2.4)

The leukemic clonogenic cell assay is based on the fact that injecting graded numbers of leukemic cells into normal BN rats results in macroscopically visible

---

<sup>1</sup>Chapter 2.3 has also been published as:

Schultz FW, Martens ACM and Hagenbeek A (1991) Simulation studies on the cell cycle phase specificity of the cytostatic drug AMSA. In: Vichnevetsky R, Miller JJH (eds) IMACS'91, Proceedings of the 13th IMACS World Congress on Scientific Computation, 1991; Vol.3, Modelling and Simulation of Biomedical Systems, pp 1456-1457

cell colonies on the spleen surface some 19-20 days later. This relationship can be exploited reversedly. As each colony is assumed to be initiated by a single clonogenic leukemic cell, counting the number of colonies is a measure of the (unknown) number of clonogenic leukemic cells contained in and injected with, for instance, a bone marrow sample.

#### 2.3.1.4 Unperturbed Leukemia Growth

The unperturbed growth of the BNML cell population in femoral marrow has been analyzed before (Chapter 2.2 and [Schultz et al., 1987]). Assuming that the femoral marrow is representative for total bone marrow, the development of BNML after transfer of  $10^7$  viable cells can be characterized by an exponential phase (constant population doubling time,  $T_d = 0.78$  d) until the number of  $1.82 \times 10^8$  cells has been reached on day 8.6, followed by a Gompertz phase during which the doubling time increases (retardation constant,  $A = 0.44$  d<sup>-1</sup>) and the population grows to a plateau level of  $1.36 \times 10^9$  cells.

Knowing the growth curve parameters, the number of cells in the population at any time can be calculated.

#### 2.3.1.5 AMSA Administration

AMSA (acridinyl anisidide; Bristol Myers Co., New York) was administered *i.v.* as either a single 20 mg/kg dose on day 13 or as single doses of 5 mg/kg on four consecutive days (13, 14, 15 and 16). See Fig. 2.16.

#### 2.3.1.6 Log Cell Kill (LCK)

Based on his observations Skipper [1986a] formulated the now generally accepted "log cell kill principle". It says that a certain drug dose always kills a same fraction rather than a same number of cells. Thus, a dose causing U LCK reduces a population of C cells to  $C \cdot 10^{-U}$  cells. This is assumed to happen instantaneously in the case of a rapid *i.v.* injection.

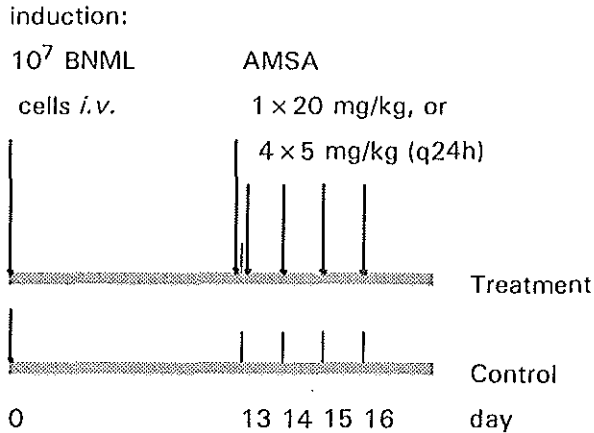
#### 2.3.1.7 Surviving Fraction

This variable is the ratio of the number of BNML cells found in an animal at a certain time after a treatment and the number of BNML cells found simultaneously in a similar animal that was not subjected to that treatment.

At present such cell numbers, and the surviving fraction, in the rat's femoral marrow were always assessed one day after each drug dose [Hagenbeek and Martens, 1986]. By considering (exponential) regrowth during the day after the drug dose while knowing how the leukemic cell population would have de-

Fig. 2.16 Scheme of the AMSA Administration Regimens (with definitions of surviving fraction and log cell kill)

EXPERIMENTAL SET-UP



n = 4 rats per group

at 24 h after each drug dose, for bone marrow:

LCFU-S assay  $\Rightarrow$  number of leukemic cells (L)

CFU-S assay  $\Rightarrow$  number of stem cells (S)

Surviving Fraction:  $SF = L_{\text{treatment}}/L_{\text{control}}, S_{\text{treatment}}/S_{\text{control}}$

Log Cell Kill:  $LCK = -^{10}\log(SF)$

veloped without treatment, the cell number ratio on the treatment day itself could be calculated and, subsequently, the LCK factor caused by the drug dose.

### 2.3.1.8 Leukemia Regrowth

It is assumed that BNML regrowth after drug therapy is not significantly different from the unperturbed growth pattern.

### 2.3.1.9 Mathematical Model for Simulating Leukemia (Re)Growth and Leukemic Cell Kill

Proliferating cells move through four phases that make up the cell cycle (Fig. 1.6). A newborn cell starts in phase  $G_1$ . When it duplicates its DNA it is in



phase S, after which it arrives in phase  $G_2$ . To complete the cycle, cell division takes place in phase M. Ideally, the cycle lasts a certain time,  $T_c$ , and each cell produces two daughters. This would yield exponential growth with (minimum) doubling time  $T_d = T_c$ .

The distribution of cell ages, between 0 and  $T_c$ , also would be exponential [Matthews, 1988], with twice as many new  $G_1$  cells as late M cells. In practice, cell cycle times are distributed. This can be expressed as B ( $1 \leq B \leq 2$ ) daughter cells per cell being born in  $T_c$ , or in the growth fraction,  $GF = B - 1$  [Steel, 1977]. Furthermore, cells may die, assumably anywhere during the cycle. The relative difference between the rate of cell production and the population's growth rate is expressed in the cell loss factor,  $\Phi$  [Steel, 1977; Schultz and Hagenbeek, 1991].

In the mathematical model (Fig. 1.7) the cell cycle is divided into 50 age compartments of  $0.02 \times T_c$  duration each. The  $G_1$ -phase comprises the first 3 compartments and compartments 4-38, 39-48 and 49-50 represent the S-,  $G_2$ - and M-phases, respectively. The number of compartments selected for each phase corresponds with the BNML phase durations, as established before [Martens et al., 1990]: 0.8, 10, 2.7 and 0.5 h for  $G_1$ , S,  $G_2$  and M, respectively. The total cycle time amounts to  $T_c = 14$  h.

First, the cells of the unperturbed day 13 population are put into the age compartments. This is done according to the theoretical distribution [Matthews, 1988] although growth on day 13 is no longer truly exponential. Then, a drug sensitive region is chosen, e.g., all S-phase compartments or only compartments corresponding to early, late or mid S, or to  $G_1$  and  $G_2$ , or to  $G_2$  and M, etc. Drug action is represented by deleting a fraction of the cells in each compartment (U LCK, where U is chosen from the range 0-3), and on top of this, by deleting a second fraction of the cells in the selected drug sensitive compartments only (V LCK, where V is chosen between 0 and 8).

Next, the remaining cells are moved from one compartment to the next with each time increment of  $0.02 \times T_c$ . Before moving the cells from the last compartment to the first one, a fraction  $p_2$  is deleted and of the remainder a fraction  $p_1$  is duplicated. It has been shown before [Schultz and Hagenbeek, 1991; also see Appendix B] that GF and  $\Phi$  can be expressed in terms of  $p_1$  and  $p_2$  and vice versa. So, either GF and  $\Phi$  or  $p_1$  and  $p_2$  determine, together with  $T_c$ , the dynamic behavior of the cell population. It was shown as well, that this model can be used to describe exponential and Gompertz growth by keeping  $p_2$  constant while varying  $p_1$  linearly with the total population size. In exponential phase, however,  $p_1$  is nearly constant. For the considered phase of exponential BNML regrowth the value of GF is estimated to be 80%. Then, given that  $T_c = 14$  h and  $T_d = 0.78$  d, it follows that  $\Phi = 12.4\%$ . This, in turn, yields  $p_1 = 0.8$  and  $p_2 = 0.07$ .

After letting the cells progress in this way for a simulated 24 h, the next

TABLE 2-4 CALCULATION OF THE LOG CELL KILL (LCK) FACTOR

| time (d)                      | nr. of BNML cells; no treatment <sup>1</sup> | surviving fraction, day T + 1 <sup>2</sup> | nr. of cells surviving on day T + 1, (obs.) <sup>3</sup> | on day T (calc.) <sup>4</sup> | LCK <sup>5</sup> |
|-------------------------------|--|--|--|-------------------------------|------------------|
| <u>dose: 1x 10 mg/kg i.v.</u> |  |  |  |                               |                  |
| 13                            | 1.0x10 <sup>9</sup>                          |  |  | 1.9x10 <sup>4</sup>           | 4.73             |
| 14                            | 1.1x10 <sup>9</sup>                          | 4.1x10 <sup>-5</sup>                       | 4.6x10 <sup>4</sup>                                      |                               |                  |
| <u>dose: 4x 5 mg/kg i.v.</u>  |  |  |  |                               |                  |
| 13                            | 1.0x10 <sup>9</sup>                          |  |  | 1.6x10 <sup>6</sup>           | 2.60             |
| 14                            | 1.1x10 <sup>9</sup>                          | 5.5x10 <sup>-3</sup>                       | 6.2x10 <sup>6</sup>                                      | 1.2x10 <sup>6</sup>           | 0.70             |
| 15                            | 1.2x10 <sup>9</sup>                          | 2.5 10 <sup>-3</sup>                       | 3.0x10 <sup>6</sup>                                      | 1.6x10 <sup>5</sup>           | 1.26             |
| 16                            | 1.3x10 <sup>9</sup>                          | 3.2x10 <sup>-4</sup>                       | 4.0x10 <sup>5</sup>                                      | 1.4x10 <sup>3</sup>           | 2.46             |
| 17                            | 1.3x10 <sup>9</sup>                          | 2.7x10 <sup>-4</sup>                       | 3.5x10 <sup>3</sup>                                      |                               |                  |

1. from fitted growth curve: unperturbed growth in bone marrow after 10<sup>7</sup> BNML cells *i.v.* on day 0 [Schultz et al., 1987]
2. observed (LCFU-S assays [Hagenbeek and Martens, 1986]); T = treatment time
3. from 1. and 2.
4. calculated from 3., assuming one day of exponential regrowth with doubling time  $T_d = 0.78$  d
5.  $-\log(\text{nr. of cells after}/\text{nr. of cells before treatment})$ ; variation of about 0.2 in LCK results from variation (SD) in observed surviving fraction (see 2.)

drug dose is given, again causing U and V LCK. The procedure is repeated until the last drug dose has been given.

Since track is being kept of the numbers of cells in each compartment, the total population size is known at every moment. The overall size reduction due to each drug dose, or LCK<sub>effective</sub>, can be calculated and compared to the LCK pattern derived from experiment. If identical patterns are found, the corresponding U and V values and any specific drug sensitive regions of the cell cycle will have been identified.

### 2.3.1.10 Software

Computer code for the simulation runs with the above mentioned model was written in TurboPascal 3.2 and implemented on a PC (IBM-XT compatible).

## 2.3.2 RESULTS

Calculated LCK values for the dose of 20 mg/kg and for each dose of 5 mg/kg are derived in Table 2-4. Figure 2.17 shows the corresponding (re)growth curves. Simulation runs with the model were not successful in reproducing the

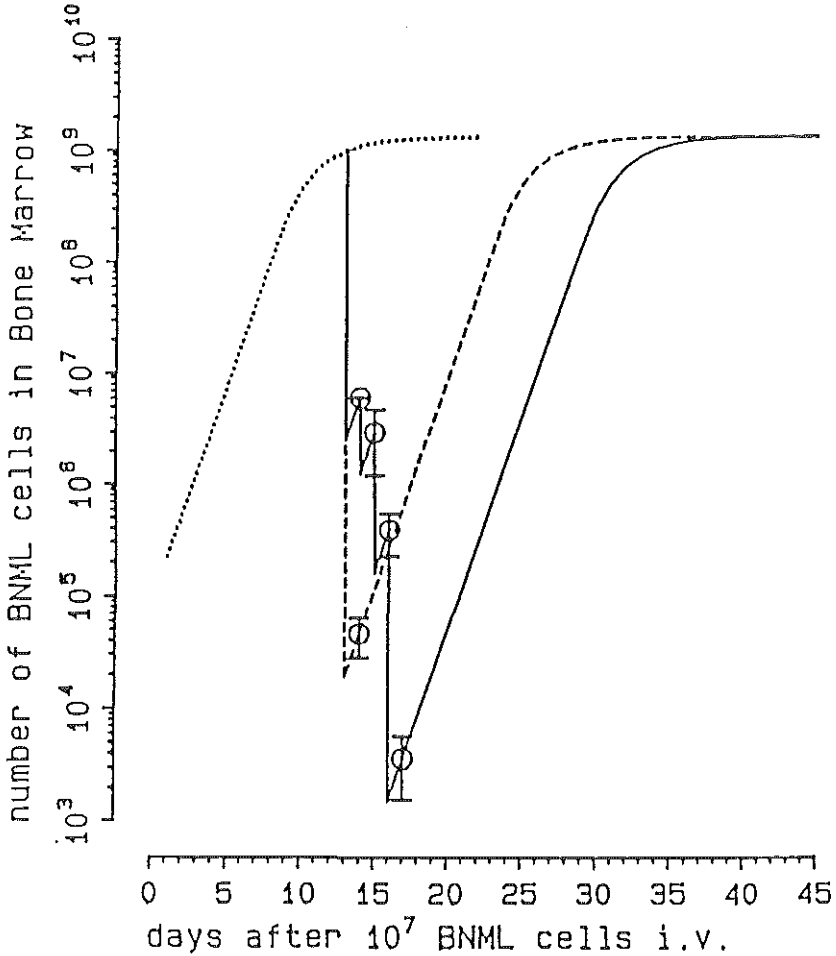
Fig. 2.17 Leukemia Development in Bone Marrow and AMSA Treatment

*dotted curve:* unperturbed growth [Schultz et al., 1987]

*dashed curve:* 1x20 mg/kg AMSA, day 13 (4.73 LCK)

*solid curve:* 4x5 mg/kg AMSA, days 13 (2.60 LCK), 14 (0.70), 15 (1.26) and 16 (2.46)

AMSA curves are based on LCFU-S survival datapoints (open circles, SD indicated [Hagenbeek and Martens, 1986]); regrowth identical with unperturbed growth; instantaneous drug action



observed LCK pattern. As example, for selected values of U and V, Fig. 2.18 illustrates a few results for the case of various fractions of S cells (early, late, all) being extra vulnerable to AMSA. None of the patterns agrees with the observed LCK sequence. Figure 2.19 shows the age distributions of the cells on

Fig. 2.18 Simulations of BNML Cell Kill (4x5 mg/kg AMSA)

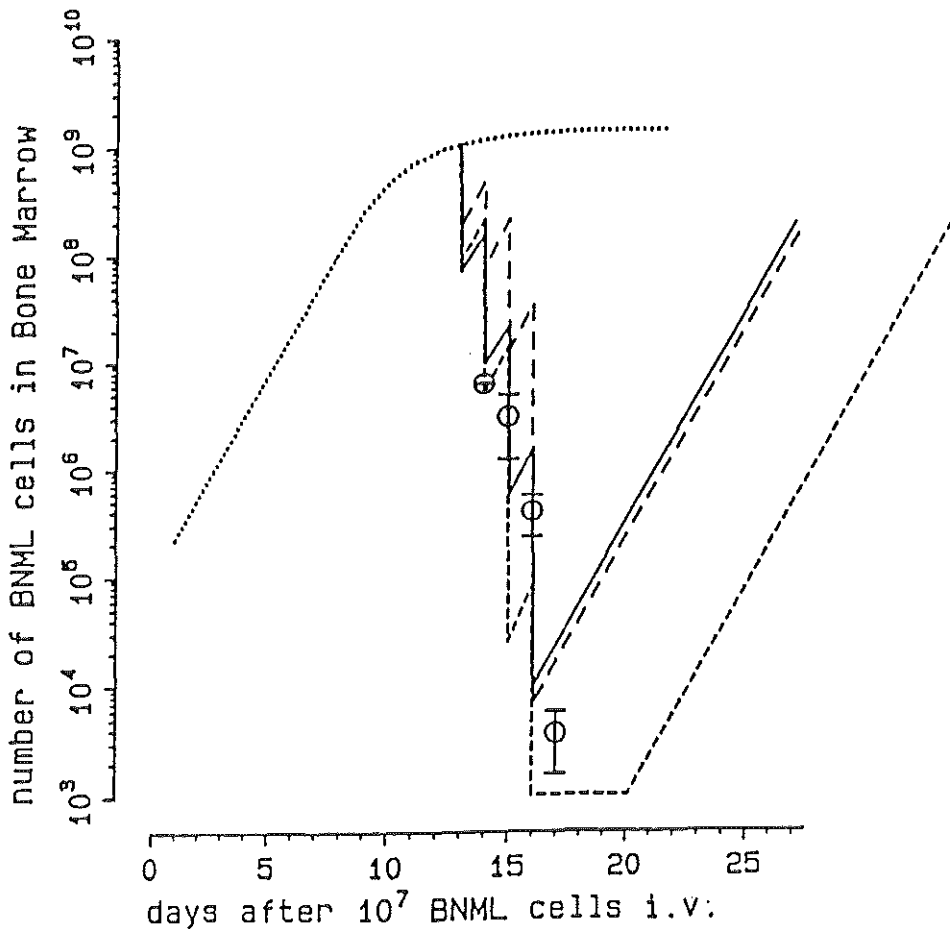
*dotted curve:* unperturbed growth [Schultz et al., 1987]

*short-dashed:*  $U=0.5$  LCK + for all S cells  $V=2.5$  LCK (age cmpts 4-38) yields  
1.08, 1.62, 2.75, 2.40 LCK<sub>off</sub>

*long-dashed:*  $U=0.5$  LCK + for early S cells  $V=4.0$  LCK (age cmpts 4-22) yields  
0.76, 0.80, 1.23, 3.72 LCK<sub>off</sub>

*solid curve:*  $U=1.0$  LCK + for late S cells  $V=2.0$  LCK (age cmpts 22-38) yields  
1.16, 1.21, 1.59, 2.21 LCK<sub>off</sub>

*from experiment:* 2.60, 0.70, 1.26, 2.46 LCK<sub>off</sub>



consecutive days, just before and after treatment, for the case of late S cells being sensitive. The effect of the drug dose on cells of the various ages as well as the effect of a day's regrowth can be seen. Figure 2.20 compares the ob-

Fig. 2.19 Age distributions of the BNML cell population on consecutive days just before (-) and just after (+) the dose of 5 mg/kg AMSA. Each dose causes  $U=1$  LCK in all compartments and an extra  $V=2$  LCK in compartments 22-38 (late S)

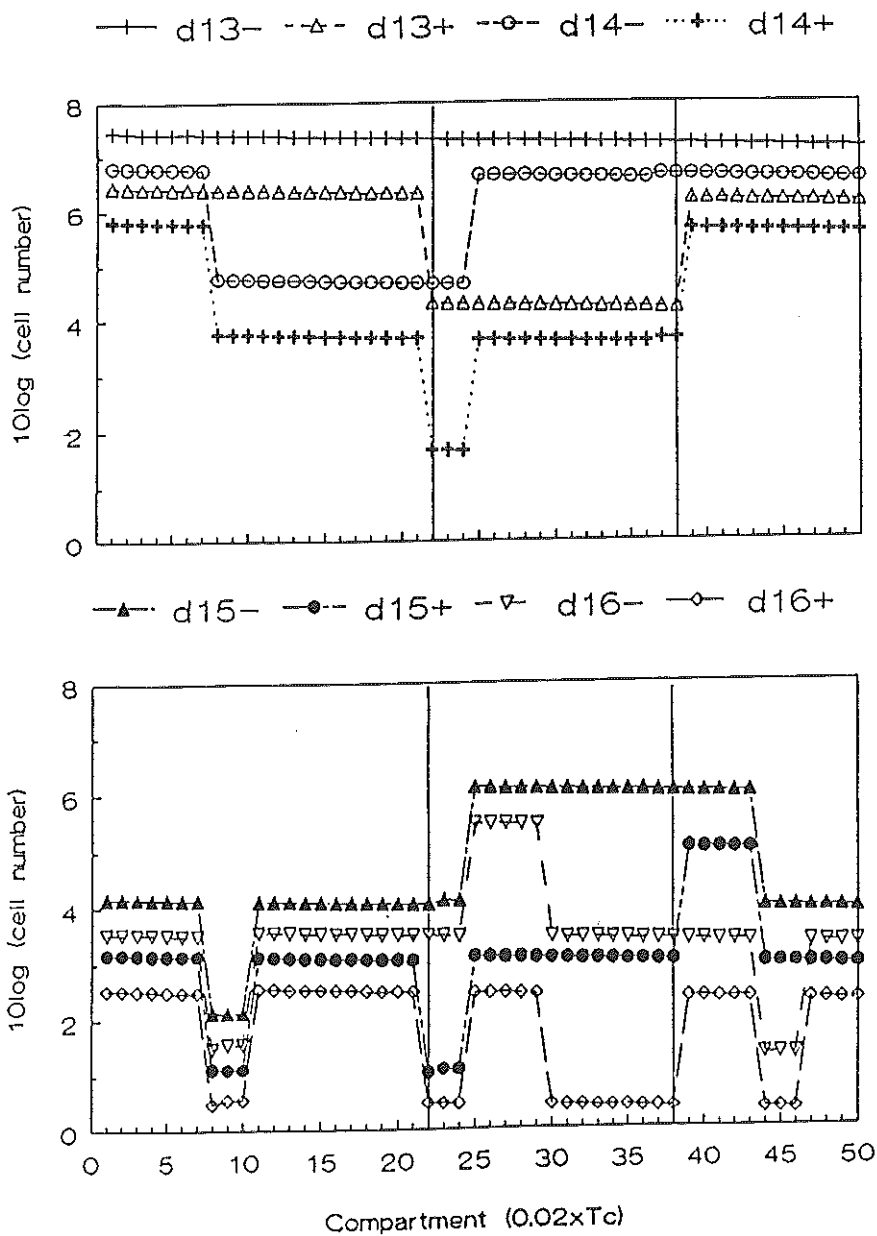
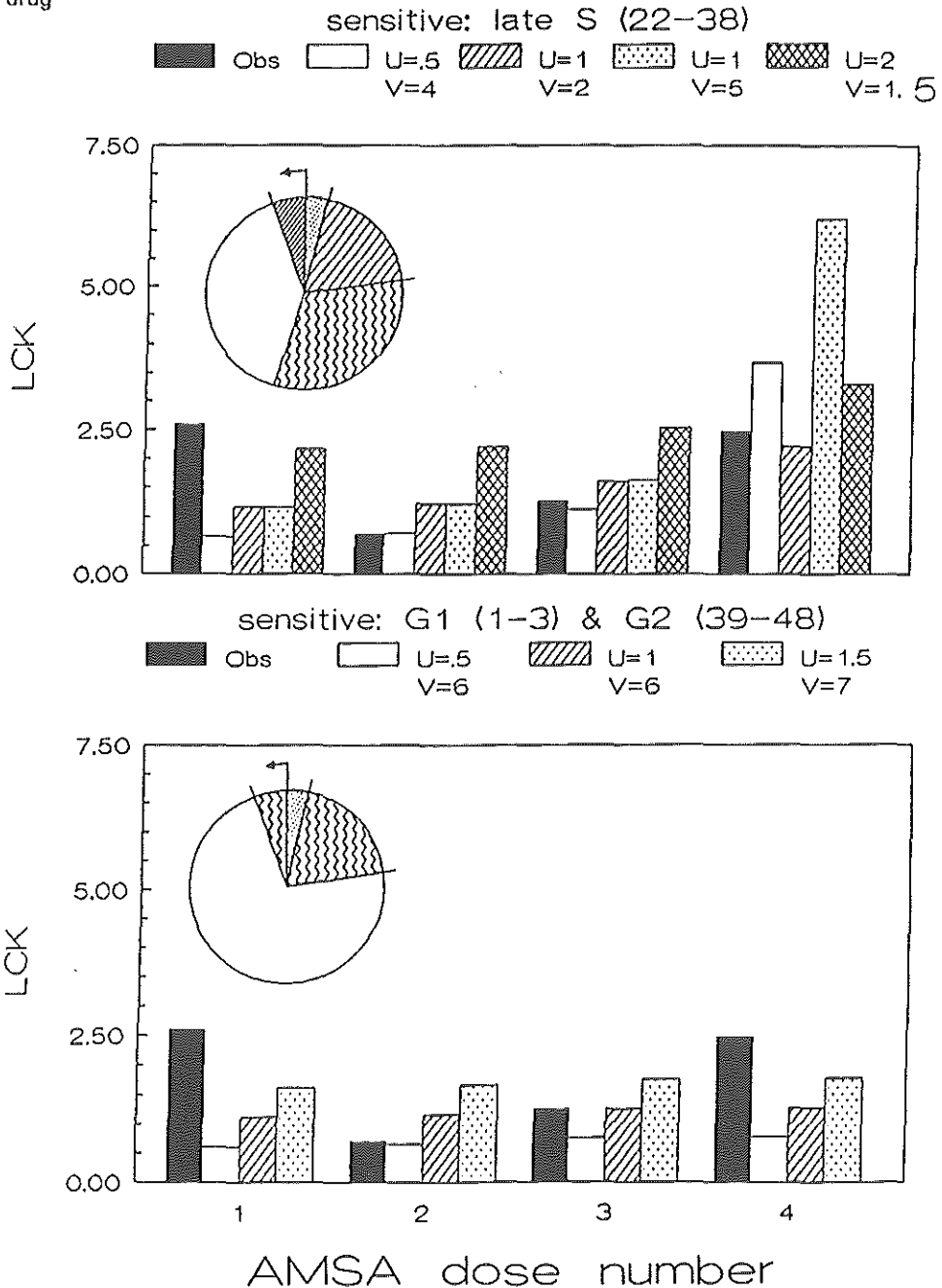


Fig. 2.20 Comparison of LCK effect of 4 doses of AMSA (5 mg/kg) for several combinations of U and V and either late S-cells or G<sub>1</sub> and G<sub>2</sub> cells being sensitive to the drug



served LCK pattern to those obtained by several simulations using various U and V values, for the case that either late S cells or G<sub>1</sub> and G<sub>2</sub> cells are drug sensitive. Again, no agreement is seen.

### 2.3.3 DISCUSSION

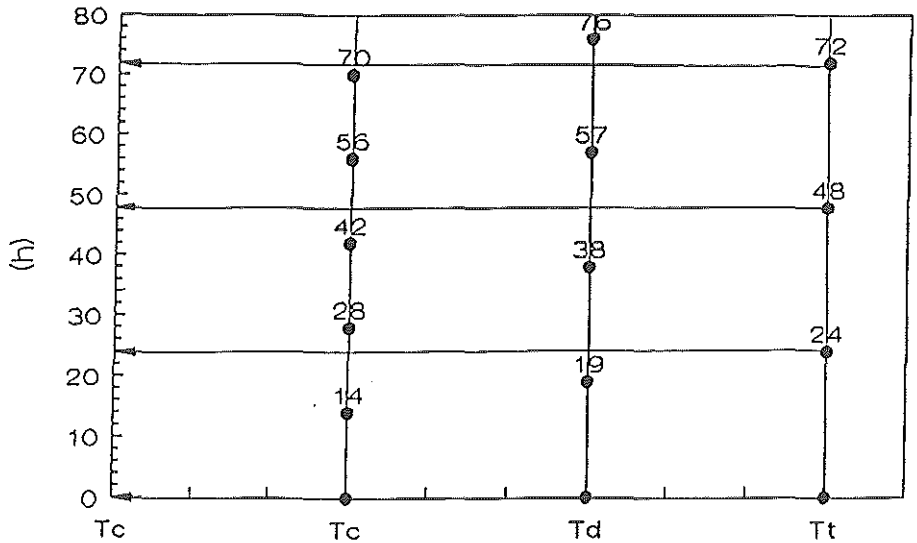
Effective LCK values for each dose of AMSA were calculated from observed surviving fractions, assuming that a) a previously found relationship between time and the size of an unperturbed BNML cell population in bone marrow holds true, b) the drug causes an instantaneous reduction in BNML cell population size, and c) the regrowth pattern after a drug dose is similar to that of unperturbed growth. In Fig. 2.17 it can be seen that under those circumstances three daily doses of 5 mg/kg are about as effective in killing BNML cells as the single 20 mg/kg dose. Thus, fractionation yields a 25% dose reduction for equal cell kill.

Striking, the LCK caused by the small drug dose is not constant. This seems to contradict Skipper's hypothesis. The first 5 mg/kg has much effect: the LCK is about half the LCK due to the 4x larger 20 mg/kg dose. The 2nd, 3d and 4th 5 mg/kg correspond to LCKs that are about 27%, 48% and 95% of the first LCK value, respectively. If the LCK would have continuously decreased the development of AMSA resistance might have been an explanation. The present decrease-increase pattern sooner suggests some cell cycle phase specific action of the drug. Especially when considering that the dosage interval (24 h) is almost as long as the population doubling time ( $T_d = 19$  h) and twice as long as the cell cycle time ( $T_c = 14$  h; Fig. 2.21). Therefore, the first dose (with large LCK) may have had a synchronizing effect, i.e., cells at sensitive stages of the cycle are killed by preference and cells in other positions are killed much less. Two cycle times later the majority of the surviving cells will have returned to their original insensitive positions. When now the second drug dose is given, it will obviously exert much less effect. Because the drug schedule and the cycle time are not perfectly in concert, and some cells cycle faster than others, desynchronization will occur and the original killing potential of the drug dose again becomes manifest later.

Furthermore, some experimental evidence has become available recently, that an AMSA derivative, amsacrine (m-AMSA), specifically kills promyelocytic leukemia cells in S-phase [Hotz et al., 1992; Gorczyca et al., 1993] and causes arrest in G<sub>2</sub>-phase [Del Bino et al., 1991; Rius et al., 1991].

Indeed it was demonstrated with the model, designed to test this hypothesis, that, in principle, non-constancy of the LCK by subsequent equal drug doses can be attributed to cell cycle phase specificity. However, results from systematically performed computer simulations show, in general, an increasing LCK pattern. The actually observed LCK pattern for AMSA could not be reproduced.

Fig. 2.21 Comparison of sequence of treatment times,  $T_t$  (24 h interval) to cell cycle time,  $T_c$  (14 h) and cell population doubling time,  $T_d$  (19 h in exponential growth phase) of BNML



Consequently, no cell cycle phases were identified as more sensitive to AMSA than others.

A weakness of the model is that no exact data is available on the regrowth doubling time, the initial age distribution, and the  $p_1, p_2$  values. Evaluation of how the varying of these parameters influences the simulation results should be a next step.



## 2.4 Drug Resistance

This chapter will deal with various aspects of drug resistance. An overview of relevant literature about investigations on (multiple) drug resistance in cancer chemotherapy is given first. Then, the development of drug resistant sublines from the BNML cell population is discussed. These subpopulations, either in pure form or to a variable extent diluted with the sensitive parent BNML cell line, can be used for *in vivo* experiments on cell kinetics of leukemia growth and response to chemotherapy. Finally, resistance to the cytostatic drug cyclophosphamide (CFA) is mathematically modeled and the implications of the model are investigated.

### 2.4.1 SUMMARY OF RELEVANT LITERATURE

#### 2.4.1.1 Evidence for Development of Drug Resistance

It is generally recognized that there is a degradation of the cell kill potential of a certain dose of a cytostatic drug when treatment is given at later stages in the development of a cancer. From experience with experimental cancer chemotherapy between 1955 and 1975 it appears that the neoplastic cell burden is inversely related to curability by any drug or combination of drugs [Skipper, 1986a]. The total tumor burden that still is curable, varies widely though, depending on the type of neoplasm and the chosen treatment regimen. For example, the apparently less effective result of identical treatment on larger sized human tumors was seen in choriocarcinoma [Goldstein, 1975], multiple myeloma [Sullivan and Salmon, 1972] and chronic lymphocytic leukemia [Rai et al., 1975]. The cause, as hypothesized already by DeVita [1983], may be found in the direct relationship between the neoplastic cell burden and the probability of the presence of permanently drug resistant cells.

For most anticancer drugs a given dose always kills a constant fraction rather than a constant number of cells (log cell kill concept; [Skipper et al., 1964; Wilcox, 1966]). This concept has been accepted by many. Hryniuk and Bush [1984] showed that there is a direct relationship between the observed temporary response rate of advanced breast cancer and the average relative dose of the individual drugs in combinations. Norton and Simon [1986], however, pointed out that, while the log cell kill concept implies that larger tumors should regress fastest, as a clinical observation intermediately sized tumors respond best to chemotherapy. This might be explained by the presence of genetically stable biochemical drug resistance in larger tumors. But another explanation is based on kinetic grounds, assuming that the cell kill might be affected by the growth fraction (Norton/Simon hypothesis, i.e., only proliferating cells, forming only a small part of a large tumor growing in Gompertzian fashion, are killed).

Other possible causes of heterogeneity in antitumor drug response can have their bases in the variation in the composition of equally sized tumors, in terms of subpopulations with various cellular characteristics (i.e., tumor heterogeneity; [Dexter and Leith, 1986]); in host factors like immune response and pharmacokinetics [Nooter et al., 1985, 1986]; and in the fact that some cytostatic agents exhibit very steep dose—response curves, which may cause large response fluctuations for relatively small changes in dose.

To a large extent, however, chemotherapeutic failure can be attributed to resistance phenomena, also in leukemias [e.g., Beck, 1983; Holmes et al., 1989]. The selection and overgrowth of resistant cells is a major factor limiting the clinical utility of cytostatic drugs in curative cancer chemotherapy. Therefore, the mechanisms of drug action and of resistance must be studied [Nelson, 1985; Griswold, 1986] to design new clinical treatment strategies that discourage the outgrowth of resistant cells. Chemotherapy regimens must be designed to kill both sensitive and resistant cells faster than they can regrow in the intervals between treatment, until all neoplastic stem cells have been eradicated [Skipper, 1986a]. In fact, the limitations of the effectiveness of the presently known cytostatic anticancer agents, at concentrations that can be safely employed, dictate the need for improved diagnostic tools, careful disease staging and the most appropriate treatment scheduling [Griswold, 1986] aimed at killing all tumor cells but also at the prevention of diversification, i.e., tumor cells growing increasingly aggressive and highly resistant [Nicolson and Lotan, 1986] that especially occurs with acute leukemias and small cell lung cancers [Curt et al., 1984]. It is recommended to study experimental systems that have obvious relevance to clinical treatment planning and drug usage [Curt et al., 1984].

#### 2.4.1.2 Classification of Different Types of Drug Resistance

DeVita [1983] distinguishes two types of resistance, both tumor mass dependent. They are *temporary* and *permanent* resistance. Temporary resistance is related to *environmental* influences, like the physical position of the cells that makes them less accessible to chemotherapy. Cells may grow in an anatomical sanctuary (e.g., the testes [Jackson et al., 1983]) or behind some pharmacological barrier (e.g., blood-brain barrier shielding the central nervous system from certain *i.v.* administered cytostatic drugs). It also may be related to the position of the cells in their cell cycle (*kinetic* resistance, e.g., resting cells are less sensitive than proliferating cells as hypothesized by Norton and Simon [1986]). Permanent resistance may be caused by some genetically stable mutation, which, in contrast to temporary resistance, is irreversible.

A second distinction may be *in vitro* or *in vivo* resistance with, perhaps, different underlying mechanisms.

Another distinction [Selby, 1984; Kaye and Merry, 1985; Skipper, 1986b]

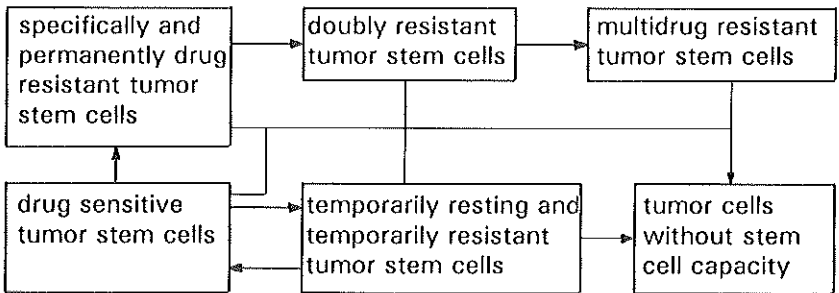
can be given as *intrinsic* or *natural* resistance versus *acquired* resistance. The former develops before any therapy has been given, e.g., by spontaneous mutation. Acquired drug resistance develops after the cell population has been challenged with the drug.

*Multidrug resistance* (MDR) is a term used for cells that exhibit resistance to a range of structurally unrelated cytostatic agents [Gerlach et al., 1986; Stark, 1986].

### 2.4.1.3 Mutation to Drug Resistance

Luria and Delbrück [1943] observed that bacteria have an inherent capacity to mutate toward resistance to agents they have never seen. There is a natural mutation frequency, and the size of the resistant population in different subcultures depends on the time point at which the mutation, i.e., a permanent genetic change, took place. Skipper [1983] considered the applicability of the mutation theory with respect to cancer chemotherapy. MacKillop [1986] drew attention to the fact that in relatively advanced malignancies sometimes primary failure of chemotherapy is observed, suggesting that at the initiation of treatment specifically and permanently drug resistant tumor cells must be present. In other words, drug resistant variant cells emerge and increase in number during the evolution of tumors, before treatment. Stephens et al. [1984] indeed identified a subpopulation of methyl-cyclohexyl-chlorethyl nitrosourea (MeCCNU) resistant cells in previously untreated Lewis lung tumors. Later [Stephens et al., 1986] it was shown that selective pressure of *in vivo* MeCCNU treatment could lead to an increasing growth advantage of resistant tumor cells. Shoemaker et al. [1983] also found that resistant cells in human cell lines (among which small cell lung cancer) may be present before the treatment.

Fig. 2.22 Scheme of emergence of multidrug resistant cells [Skipper, 1986a]



Drug resistant cell lines also can be made by challenging a drug sensitive parent cell line with small amounts of the drug. Schmid et al. [1976] found that repetitive passage of leukemia cells in animals treated with a 6-drug combination, delivered either simultaneously or in sequence, selected sublines that were specifically resistant against one, then two, three, four, five or six drugs in the combination. Skipper [1986a] suggests that in general the sensitive parent line mutates first to single drug resistant sublines that in turn mutate to double drug resistance and afterwards may mutate to multidrug resistance (Fig. 2.22).

The rate of mutation to resistance can be experimentally determined by exposing cell samples to a lethal dose of a drug at certain time intervals and subsequently measuring the uptake of bromodeoxyuridine (BrdUrd). Only resistant cells (survivors) will take up the BrdUrd, which can be quantified by flow cytometry after labeling with a specific fluorescent marker [DeFazio and Tattersall, 1985]. Flow cytometry is also useful for measuring other cell kinetic properties that are relevant to experimental chemotherapy, with emphasis on the mechanisms that lead to cell death in tumors [Tannock, 1986]. Changed cell kinetic properties might explain the apparent discrepancy noticed by Van Putten [1986] that transplanted cell lines in the mouse retain their drug sensitivity in spite of somatic mutation. The problem is solved if unchallenged sensitive cells grow faster than mutants, so that in an otherwise unperturbed situation the latter will soon be outnumbered. Zajicek [1986] remarks that acquired resistance must be cell type specific, because normal tissues, e.g., gastrointestinal mucosa or bone marrow, remain drug sensitive despite repeated chemotherapeutic treatment.

Another way of measuring the rate of resistance development is by examining growth delay and clonogenic cell survival, as described by McMillan et al. [1985] for MT murine mammary carcinoma treated with melphalan, cisplatinum and cyclophosphamide.

#### 2.4.1.4 Mechanisms of Drug Resistance

Carter [1984], comparing drug resistant cells to their drug sensitive counterparts, recognizes next to the already mentioned mechanisms of kinetic and pharmacological nature also a biochemical factor in the emergence of resistance. An inventory of possible mechanisms of drug resistance is given by Curt et al. [1984] and for acute leukemias in particular by Hall et al. [1989], e.g., defective drug transport (which may include reduced membrane permeability [Ling and Thompson, 1974] as well as changed drug uptake characteristics (active efflux) through the cell membrane [Colvin, 1984]), defective metabolism, enhanced DNA repair [Colvin, 1984], gene amplification, or altered target proteins. Mechanisms associated with different drugs are mentioned; e.g., increased deactivation of damaging cyclophosphamide metabolites by glutathione

S-transferase enzymes in a cyclophosphamide resistant Yoshida sarcoma cell line was reported [McGown and Fox, 1986], though in the BNML the detoxification enzyme aldehyde dehydrogenase seems more important [De Groot et al., 1992]. The role of these enzymes in cyclophosphamide resistance is reviewed by Morrow and Cowan [1990] and Waxman [1990].

Beck [1984] summarizes that MDR is related to a changed pharmacology (less drug accumulation, faster release, weaker binding) due to inherited membrane alterations (diminished permeability) based on chromosomal changes (gene amplification and increased transcription of the P170 glycoprotein [Kartner et al., 1983]). The presence of an MDR phenotype with an activated drug efflux pump has been demonstrated in many mammalian tumor cell lines, among which human ovarian tumors, carcinomas and sarcomas [Ling et al., 1983; Fojo et al., 1985; Gerlach et al., 1986; Pastan and Gottesman, 1987], as well as acute myeloblastic leukemia [Ma et al., 1987; Holmes et al., 1989]. Biomolecular studies [Roninson et al., 1984; Stark, 1986; Roninson, 1987; Goldstein et al., 1989; Hayes and Wolf, 1990; for leukemias, e.g., Tsuruo et al., 1987] have identified the responsible gene (MDR1) associated with the membrane protein that regulates transport of toxic compounds across the membrane (energy dependent efflux pump). The degree of resistance correlates with the increased expression, due to MDR1 gene amplification, of the P170 glycoprotein. Therefore, quantification of this membrane protein during some time may be another way of finding the rate of resistance development.

Reversal of the drug resistance may be achieved by using calcium channel blockers, e.g., verapamil or cyclosporin A, or other membrane transport modifiers that restore the intracellular drug accumulation by counteracting the efflux pump [Curt et al., 1984]. This is true, e.g., for acute myelocytic leukemia [Nooter et al., 1989, 1990a].

## 2.4.2 DEVELOPMENT OF DRUG RESISTANT BNML CELL LINES

The original BNML cell line is sensitive to the cytostatic agents cyclophosphamide (CFA; [Hagenbeek and Martens, 1982]), cytosine arabinoside (ARA-C; [Colly et al., 1984b]) and anthracyclines like doxorubicin (DOX; [Sonneveld et al., 1981]) and daunomycin (DAU; [Nooter et al., 1990b]). To study the phenomenon of drug resistance *in vivo* it was attempted to develop BNML sublines that are no longer sensitive to these drugs. The *in vivo* development of three BNML sublines, resistant to treatment with CFA, ARA-C and DAU, was successful [Hagenbeek et al., 1987b; Martens et al., 1991]. A fourth resistant variant, against the *orally* very potent drug Acetyldinaline, was developed recently [El-Beltagi et al., 1993]. The procedures used are very similar. As an example, the CFA case is briefly described below.

TABLE 2-5 RESPONSE OF THE PARENT BNML LINE AND THE CYCLOPHOSPHAMIDE RESISTANT LINE TO TREATMENT WITH CYCLOPHOSPHAMIDE, IFOSPHAMIDE AND MAPHOSPHAMIDE [Martens et al., 1991]

The response of the parent BNML cell line to treatment with three closely related drugs, cyclo-, ifo- and maphosphamide (ASTA Werke AG, Bielefeld, Germany) is compared with the response of the resistant line. The drug analogs were studied to provide additional information on possible cross resistance. The dose of each drug was chosen such that a significant antitumor effect could be expected. The resistant line appeared not to be sensitive to either cyclophosphamide (CFA) or ifosphamide (IFA) at the doses tested. Interestingly, maphosphamide (MFA) treatment resulted in slightly more than 1 log cell kill. A difference in drug activation—MFA is an already activated form of CFA, whereas the latter itself must be metabolized first—might be a factor responsible for the observed difference in response.

| TREATMENT                 | MdST<br>(days) | ILS<br>(days) | SURVIVING FRACTION<br>OF LEUKEMIC CELLS |
|---------------------------|----------------|---------------|---|
| <i>BNML PARENT LINE</i>   |                |               |   |
| EXP 1: CONTROLS           | 22.5           | --            | --                                      |
| CFA                       | 45             | 22.5          | $2.3 \times 10^{-6}$                    |
| IFA                       | 39             | 16.5          | $7.5 \times 10^{-5}$                    |
| EXP 2: CONTROLS           | 22             | --            | --                                      |
| MFA                       | 35             | 13            | $5.6 \times 10^{-4}$                    |
| <i>CFA RESISTANT LINE</i> |                |               |   |
| EXP 1: CONTROLS           | 19.5           | --            | --                                      |
| CFA                       | 20             | 0.5           | $7.5 \times 10^{-1}$                    |
| IFA                       | 19.5           | 0             | 1                                       |
| EXP 2: CONTROLS           | 20             | --            | --                                      |
| MFA                       | 25             | 5             | $7.5 \times 10^{-2}$                    |

MdST: median survival time after *i.v.* transfer of  $10^7$  BNML cells

CFA: 100 mg/kg *i.p.*

IFA: 200 mg/kg *i.p.* } at day 13 after leukemia transfer

MFA: 154 mg/kg *i.v.* } (n=5 rats/group)

ILS: increase in life span

The surviving fraction is deduced from the known relationship that 1 decade of tumor load reduction corresponds to 4 days ILS (Eq.(2.1))

#### 2.4.2.1 A Cyclophosphamide Resistant Subline

Leukemic BN rats (leukemia induction with  $10^7$  BNML cells injected *i.v.* at day zero) were treated repeatedly at intervals of two weeks with CFA doses of 100 mg/kg *i.p.* Under these conditions each dose induces a 5 log cell kill (LCK) in the drug sensitive leukemic cell population, as can be derived from known cell kinetic properties and observed survival times of the animals with respect to untreated controls. During the two week periods after treatment the surviving BNML cell population regrows until the next challenge with a CFA injection.

Fourteen days after the last of a total of seven injections with CFA the leukemia was transferred ( $10^7$  BNML cells) to a fresh group of rats. The new recipients received twice the periodical CFA treatment. At this stage the first signs of drug resistance became apparent. The leukemia was transplanted once more and again was treated with two doses of CFA at days 8 and 22 after the transfer. This procedure was continued until the subline of leukemic cells had been exposed 17 times to CFA drug doses of 100 mg/kg each.

From time to time the sensitivity of the subline for CFA was tested by scoring increased lifespan (ILS) of animals treated with 100 mg/kg CFA with respect to untreated controls. With an increasing number of exposures to the drug the resistance increased—as indicated by shortening ILS—until total resistance was attained (Table 2-5).

Many CFA resistant BNML cells were frozen and stored in liquid nitrogen, forming a stock supply for further experimentation.

#### 2.4.2.2 Mechanism of Cyclophosphamide Resistance

CFA requires bioactivation to become cytotoxic. Therefore, an *in vivo* approach was chosen, employing the CFA resistant and the sensitive parent BNML cell lines, to investigate the molecular mechanism(s) of CFA resistance [De Groot et al., 1994].

It appeared that the CFA-detoxifying enzyme, aldehyde dehydrogenase (ALDH) plays an important role. The level of ALDH activity in the drug resistant cells was about six times as high as in cells of the parent line. When ALDH was counteracted by pretreatment of rats, carrying resistant BNML cells, with the ALDH-inhibitor disulfiram (DSF), the *in vivo* cytotoxicity of CFA was fully restored. Log cell kill *in vitro* of CFA resistant BNML cells by maphosphamide (MFA, an activated CFA derivative) increased by 2-3 when these cells had been pretreated with DSF. This partial restoration of cytotoxic effect could be deduced from the differences in median survival times of groups of rats, injected with resistant BNML cells incubated with either MFA alone or MFA + DSF.

Another CFA-detoxification pathway involves glutathione (GSH) and glutathione dependent enzymes (GSTs). This pathway is probably much less important with respect to the CFA resistance phenomenon. Differences in GST contents of resistant and sensitive BNML cells are only moderate to small. Furthermore, GSH depletion by buthionine sulfoximine did not increase the sensitivity of CFA resistant cells.

### 2.4.3 MODELING CYCLOPHOSPHAMIDE RESISTANCE: SIMULATION STUDIES<sup>1</sup>

At the concentrations that may safely be employed, the presently known anticancer agents show a limited curative effectiveness [Griswold, 1986]. It is a well-known observation that identical treatment schedules, when applied in sequence, have less and less effect on the growth of the population of malignant cells. Assuming, as supported by an analysis of growth data in a case of lymphocytic leukemia [Schultz et al., 1989], that the intrinsic cell kinetic parameters themselves have not been altered by the treatment, a likely explanation may be that drug resistance has developed, causing the surviving fraction to grow larger in every successive treatment course. Figure 2.23 shows a nonlinear decrease with time in cell kill in the brown Norway rat acute myelocytic leukemia (BNML) by successive doses (100 mg/kg *i.p.*) of the drug cyclophosphamide (CFA).

It also is a clinical observation that a large tumor mass at diagnosis will respond relatively poorer than a small one to an identical treatment regimen. This may be so because of a large tumor containing relatively more resistant cells [DeVita, 1983]. This would imply that drug resistant mutant cells are developing "naturally" (in contrast to "acquiring" resistance afterwards because of the contact with the cytostatic agent [Selby, 1984]).

Thus, the failure of eradicating tumors with chemotherapy may, at least partly, be attributed to the development of (multi)drug resistance. To be able to formulate optimal therapy regimens it is necessary first to gain insight into the dynamics of cell population growth and the process of resistance development, under unperturbed as well as under chemotherapy conditions. Such insight may be gained in an efficient way through mathematical modeling and computer simulation [Garfinkel, 1983]. Based on available knowledge a model is postulated. Its implications are examined and, ideally, tested against observations from real-life experiments. The model and its underlying hypothetical mechanisms thus are validated; they are either accepted to (sufficiently) describe the real processes or it turns out that they should be rejected, which implies a redefinition of the model.

Below, a simple model is examined for its relevance with respect to the growth of the BNML leukemia that is treated with different single doses of CFA at several stages of the disease.

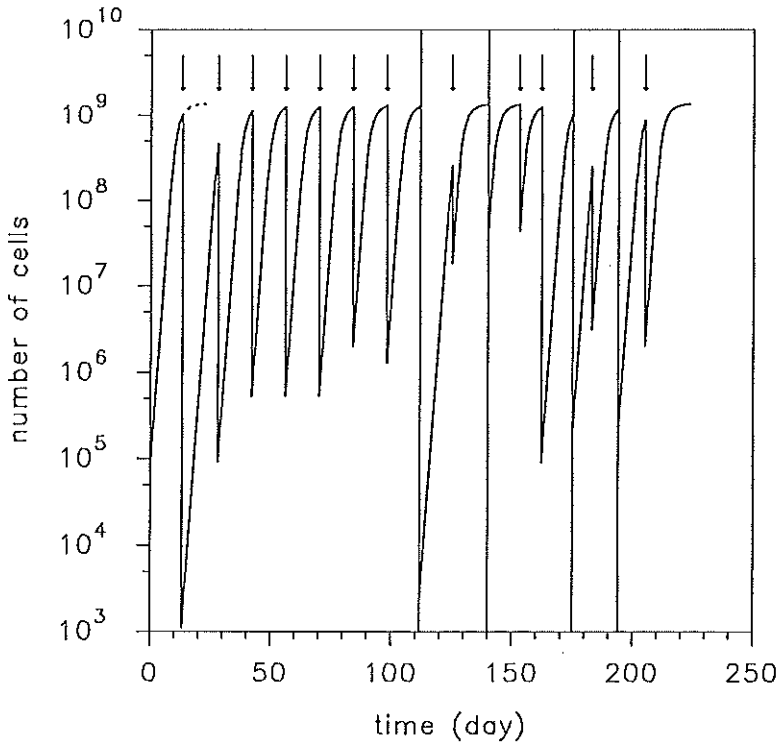
---

<sup>1</sup>Large parts of Chapter 2.4.3 have been published as:

Schultz FW, Martens ACM, de Vries AJ and Hagenbeek A (1989) Modeling cyclophosphamide resistance in the brown norway rat acute myelocytic leukemia: a first approach. In: Eisenfeld J, Levine DS (eds) Biomedical Modelling and Simulation, JC Baltzer AG Scientific Publishing Co., pp. 47-49



Fig. 2.23 Curves showing the BNML cell population in the bone marrow as function of time after *i.v.* transfer of  $10^7$  (CFA-sensitive) leukemic cells. Assuming "regrowth = growth", by backwards extrapolation of the unperturbed growth curve (Fig. 1.8) and matching observed MdST to cell number at death ( $=1.36 \times 10^9$ ), effect of consecutive CFA doses (arrows: 100 mg/kg *i.p.*) can be calculated. Also environmental influences, i.e., effect of retransplantation into a fresh recipient rat (bars) can be seen. Apparently, LCK is *not* constant, but variation is not a monotonous decrease either, as expected in case of developing drug resistance



### 2.4.3.1 LABORATORY EXPERIMENTS

The Rat Leukemia Model, BNML (See Chapter 1.9.2). BNML was chemically induced in a female Brown Norway (BN) rat and appeared to be cellular transferable within this strain, yielding a reproducible growth pattern. As few as 25 BNML cells inoculated *i.v.* will cause leukemia in 50% of the BN rats ( $ED_{50} = 25$ ). The principal target organs for the development of the malignant cell population are the bone marrow, the liver and the spleen. The characteristics of the disease show striking similarities with human acute myelocytic leukemia (AML). Therefore, experimental BNML studies bear great importance to

clinical treatment of AML.

**Unperturbed BNML Growth** (Chapter 2.2 and [Schultz et al., 1987]). The development of the leukemic cell population after *i.v.* inoculation of  $10^7$  BNML cells has been investigated in the major target organs (bone marrow, liver, spleen). The number of BNML cells present at various time points could be experimentally determined in several ways, ranging from counting cells—morphologically or after labeling with tumor cell specific fluorescent dyes—to more complex bio-assay methods. Figure 1.8 shows the datapoints for bone marrow, as well as the growth curve found by nonlinear least squares fitting. The curve consists of an exponential part (constant population doubling time,  $T_d$ ) that at a certain population size ( $2 \times 10^8$ ) changes into a Gompertz curve.  $T_d$  then increases exponentially and the population growth decelerates to a constant steady state plateau level. For liver and spleen similar growth curves were found. The size—time course of the total population is reflected by the sum of the three separate curves (Fig. 2.15). The animals die on day 22 (median value) with a tumor burden of some  $2 \times 10^{10}$  BNML cells.

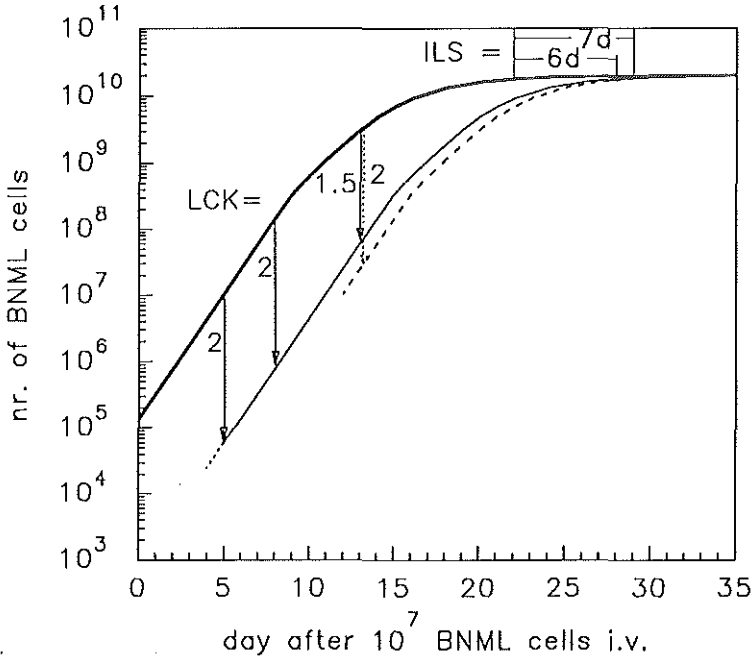
Of the  $10^7$  inoculated cells on day zero only some  $1.5 \times 10^5$  appear to grow out. This corresponds well with the fact that the  $ED_{50}$  amounts to 25 cells, i.e., a unit of 25 cells has a 50/50 chance to be "destroyed" (zero cells grow out; no tumor) or to yield 1 cell growing out (a tumor may develop from a single malignant cell!). When  $10^7$  cells are considered as  $4 \times 10^5$   $ED_{50}$  units, each with 50% chance to contribute 1 cell or 0 cells to grow out, then the expected initial population size becomes  $2 \times 10^5$ . Likewise, should  $10^6$  cells be inoculated, a factor of 10 reduction, then an expected number of  $2 \times 10^4$  would remain to grow out.

**CFA and an *in vivo* CFA Resistant BNML Cell Line.** A few remarks about the drug CFA are given in Chapter 2.2.1.2. The development of an *in vivo* CFA resistant BNML cell line is discussed in Chapter 2.4.2.1.

**The Log Cell Kill (LCK) Hypothesis.** In experimental cancer chemotherapy for most cytostatic agents an invariable dose (intensity) response is observed, from which the first order LCK hypothesis can be deduced [Skipper, 1986a]. It states that a same dose of a certain drug always kills a same fraction of a cell population rather than a same number of cells. Thus, a drug dose causing  $q$  LCK reduces a cell population of  $C$  cells to  $C \cdot 10^{-q}$  cells. This LCK hypothesis has been widely accepted and will be applied in this study to describe CFA action.

**Relation between Increase in Lifespan (ILS) and LCK.** By transferring various numbers of BNML cells (parent line) *i.v.* to BN rats and observing the subsequent median times of death (MdST) a dose—response relationship has

Fig. 2.24 Assuming the growth pattern after  $10^7$  BNML cells *i.v.* as standard (thick line; MdST = 22 d), a certain ILS can be achieved by a) another inoculation size or b) a single drug dose whose effect should decrease with increasing treatment time. E.g., 2 LCK on day 5 or 8, or 1.5 LCK on day 13 (thin line) yields ILS = 6 d (n.b., 2 LCK on day 13 yields ILS = 7 d (broken line)).



been established (Eq.(2.1); Fig. 2.2), from which it can be deduced that every 4 d ILS corresponds with a reduction by a factor of 10 (1 log) in the number of cells. This result can be generalized to:

$$q = 0.25 \times ILS, \tag{2.19}$$

relating the ILS (d) to q LCK (i.e., a reduction of  $10^{-q}$  in cell number).

Actually, a correction to Eq.(2.19) is needed, because of the Gompertzian nature of the BNML cell population's growth curve. It is assumed that the BNML growth pattern will always be identical (same constant  $T_d$  in exponential phase, followed by Gompertzian growth beyond a certain population size), both in the unperturbed situation and when treatment is given. Treatment just causes an instantaneous reduction in the cell number. Then, given one size—time datapoint the population size development can be (re)constructed by horizontal shift of the basic growth curve in Fig. 2.24, for which the total BNML popula-

tion growth curve of Fig. 2.15 is taken.

Starting with inoculating  $10^7$  BNML cells an MdST of 22 d is observed. To postpone the MdST to day 28 (ILS = 6 d; shift the basic growth curve 6 days to the right), a 2 LCK drug dose should be given on day 5, or 1.95 LCK on day 8 or only 1.44 LCK on day 13. A drug dose of 2 LCK on day 13 would yield 7 d ILS. For 12 d ILS a drug dose of 4, 3.9 or 3.5 LCK would be required on days 5, 8 or 13, respectively. Thus, for achieving a same ILS, late treatment requires less LCK (a smaller dose of the drug) than early treatment, because of the Gompertzian nature of population growth in later stages. The other way around, giving a certain drug dose (i.e., a certain LCK) late rather than early would yield a larger ILS. This effect is more pronounced when treatment is given later and/or the LCK induced is smaller.

An unsolved discrepancy is the fact that 6 d ILS after 2 LCK on day 5 does not quite agree with the 8 d ILS expected, when  $10^5$  (of which  $1.5 \times 10^3$  should grow out) BNML cells are inoculated on day zero.

**Dose—effect Measurements with CFA.** BN rats were inoculated *i.v.* with  $10^7$  BNML cells of the parent cell line and treated on different days (5, 8, 9, 11, 13) with single *i.p.* doses of CFA of various magnitudes (10, 60, 80, 100, 120, 140, 160 mg/kg). The times of death due to leukemia were measured and the MdSTs were calculated and converted to LCK values as described above (Eq.(2.19)). The results are shown in Fig. 2.25. No correction for the deviation due to Gompertzian growth was made, as the assumption on identical growth patterns is uncertain. So for smaller doses at later time points the shown net log cell kill,  $q$ , might be calculated somewhat too high.

Nowrousian and Schmidt [1984] experimentally found a linear relationship between CFA dose and the logarithm of the surviving fraction (SF) of pluripotential (CFU-S) and granulocytic (CFU-C) progenitor cells in mouse bone marrow. Thus,

$$\log(SF) = -h \cdot dose. \quad (2.20)$$

where  $h$  is a positive constant. Let the dose correspond to  $q$  LCK, then

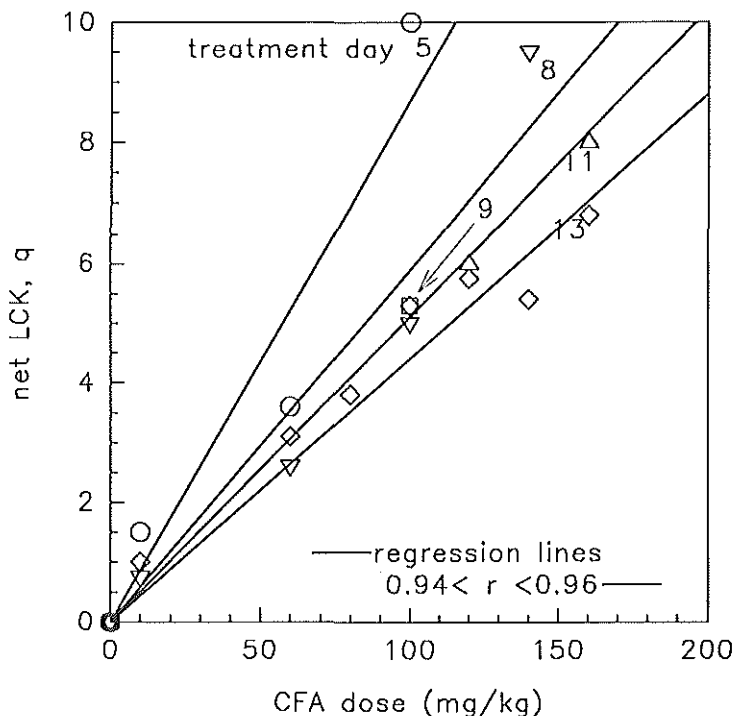
$$SF = C^+ / C^- = 10^{-q}, \quad (2.21)$$

where  $C^+$  and  $C^-$  are the numbers of cells just after and before the drug administration, respectively. Substituting Eq.(2.21) into Eq.(2.20) yields

$$q = h \cdot dose. \quad (2.22)$$

Knowing that the dose—LCK relation is linear, regression lines (going through

Fig. 2.25 Observed dose—effect relations for CFA: net LCK,  $q$  (instantaneous total population size reduction factor) for different CFA doses given at different treatment times. Linear regression through the origin yields lines whose slopes decrease with increasing treatment time: a certain dose given later, has less effect



the origin, zero dose has zero effect) can be drawn. The slope of the regression lines appears to decrease with increasing treatment time; *contrary to what would be expected from the growth dynamics, a certain dose, given later, shows less effect (lower  $q$ )*. Is this due to relatively more CFA resistant cells occurring in a well-developed tumor, or are the assumptions on the growth dynamics non-valid?

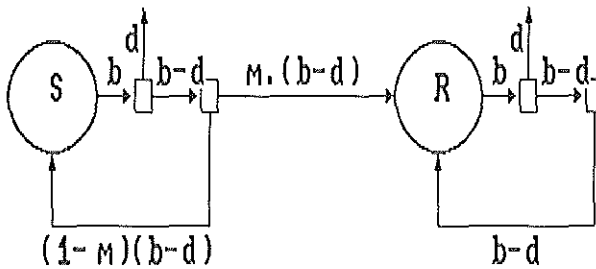
**Experiments with Mixed Populations of BNML Cells from the Parent Line and the CFA Resistant Line.** Mixtures of different ratios of cells of the parent BNML cell line and the *in vivo* developed CFA resistant cell line were inoculated *i.v.* into BN rats in quantities of  $10^7$  cells. Death due to leukemia was awaited without further treatment (control) or with an *i.p.* injection of 100 mg/kg CFA on day 13. MdSTs and ILSs were calculated (Table 2-6). As could be expected, starting with more initial CFA resistant cells results in earlier death and less effect of the same drug dose.

TABLE 2-6 MdST AND ILS AS FUNCTION OF INITIAL S/C RATIO

| percent $S_0$ | $S_0/C_0$ | MdST | ILS |         |
|---------------|-----------|------|-----|---------|
| 100           | 1.00      | 22   | -   | control |
|               |           | 44   | 22  | CFA     |
| 99            | 0.99      | nd   | nd  | control |
|               |           | 39.5 | nd  | CFA     |
| 90            | 0.9       | nd   | nd  | control |
|               |           | 34   | nd  | CFA     |
| 0             | 0         | 29   | -   | control |
|               |           | 29   | 0   | CFA     |

$10^7$  cells *i.v.* on day zero; CFA: 100 mg/kg cyclophosphamide *i.p.* on day 13; MdST: median survival time; ILS: increase in lifespan; S: drug sensitive cells; C: total nr of cells

Fig. 2.26 Scheme of the proliferation model (birth/death rates:  $b/d$ ) for drug sensitive (S) and drug resistant (R) cells, and mutation (rate  $m$ ) from S to R.



### 2.4.3.2 MODELING AND SIMULATION

The Model for Resistance Development and the Growth of CFA-Sensitive and -Resistant Subpopulations. The relations between the two proliferating subpopulations are schematically shown in Fig. 2.26. Looking at the sensitive cells (CFA/S), per unit of time and per cell,  $b$  cells are born and  $d$  cells disappear (e.g., natural death). Of the difference, a constant fraction  $m$  (spontaneous mutation rate) changes into resistant cells (CFA/R); the rest returns to the CFA/S compartment. The CFA/R population also has birth and loss rates, assumably equal to  $b$  and  $d$ . The growth rates of the subpopulations can be written as (a dot above a symbol denotes the time-derivative):

$$\begin{pmatrix} \dot{S} \\ \dot{R} \end{pmatrix} = (b-d) \cdot \begin{pmatrix} 1-m & 0 \\ m & 1 \end{pmatrix} \cdot \begin{pmatrix} S \\ R \end{pmatrix} \quad \text{cells/unit time.} \quad (2.23)$$

Starting from initial population sizes,  $S_0$  and  $R_0$ , the development of the population with time,  $t$ , is obtained from integration of Eq.(2.23) under the assumption that  $b$  and  $d$  are constant ( $b - d = k > 0$ ):

$$\begin{pmatrix} S \\ R \end{pmatrix} = \begin{pmatrix} S_0 & 0 \\ -S_0 & S_0 + R_0 \end{pmatrix} \cdot \begin{pmatrix} \exp\{k \cdot (1-m) \cdot t\} \\ \exp\{k \cdot t\} \end{pmatrix}. \quad (2.24)$$

$$\text{The total population,} \quad C = S + R. \quad (2.25)$$

If  $m \rightarrow 0$  then the subpopulations grow independently in identical ways. If  $m \rightarrow 1$  then all cells produced become resistant:  $S = S_0$  remains constant and  $R = C_0 \cdot \exp\{k \cdot t\} - S_0$  grows.

The assumption of  $b$  and  $d$  being constant is valid in exponential growth when the population doubling time,  $T_d$ , is constant ( $k = \ln 2 / T_d$ ). Real populations will grow (nearly) exponentially during some time span and then  $T_d$  starts to increase more and more until the population size reaches a steady state plateau level (see Fig. 1.8, where actual growth data are fit with contiguous exponential and Gompertz curves). Computer simulations (not further discussed here) revealed that this type of growth can be the result of a constant  $d$ , while  $b$  decreases linearly with population size. The linear decrease, however, is such, that the value of  $b$  can be considered a constant for a long period.

**Development of the R/S Ratio.** Let the initial population consist of a mixture of CFA/S and CFA/R cells, i.e.,

$$S_0 = f \cdot C_0 \quad \text{and} \quad R_0 = (1-f) \cdot C_0; \quad 0 \leq f \leq 1. \quad (2.26)$$

The R/S ratio at time  $t$  follows from Eq.(2.26) and Eq.(2.24), if  $f \neq 0$  then:

$$R/S = \exp\{k \cdot m \cdot t\} / f - 1, \quad (2.27)$$

with time-derivative

$$(R/S) = (k \cdot m/f) \cdot \exp\{k \cdot m \cdot t\}, \quad (2.28)$$

which both are  $\geq 0$  always. So, if  $m > 0$  then the R/S ratio increases with time. However, substitution of some probable values, i.e.,  $k = 0.65/d$  ( $T_d = 1$  d);  $f = 0.9$ ;  $m = 10^{-4}$  (100 times the estimated somatic mutation rate of 1 cell in a million [Goldie and Coldman, 1985]), reveals that R/S increases from 0.11 to only 0.12 in over 100 d. Thus, practically spoken the R/S ratio remains fairly constant.

Because  $C = R + S$ , Eq.(2.27) can be rewritten as:

$$S/C = f \cdot \exp\{-k \cdot m \cdot t\}. \quad (2.29)$$

Net Log Cell Kill by a Single Drug Dose. Starting with  $C_0$  cells at time zero, of which  $f \cdot C_0$  are sensitive, at time  $t^-$  just before treatment the population has grown to  $C^-$  cells. The numbers of S and R cells then amount to:

$$S^- = C^- \cdot f \cdot \exp\{-k \cdot m \cdot t^-\}; \quad R^- = C^- - S^-. \quad (2.30)$$

Treatment consists of a single drug dose causing  $p$  log cell kill ( $p$  LCK) to sensitive cells, i.e., the drug sensitive population is instantaneously reduced by a factor of  $10^{-p}$ . This value  $p$  can be considered as a potential LCK, i.e., it is the maximum attainable value. As resistant cells may be present, which will not be affected, in general the net LCK,  $q$ , will be smaller. So, just after treatment at time  $t^+$  the population will be reduced to:

$$S^+ = S^- \cdot 10^{-p}; \quad R^+ = R^- = C^- - R^-; \quad (2.31)$$

$$C^+ = S^+ + R^+ = C^- \cdot [1 + (10^{-p} - 1) \cdot f \cdot \exp\{-k \cdot m \cdot t\}]. \quad (2.32)$$

$$\text{By definition, } C^+ = C^- \cdot 10^{-q}. \quad (2.33)$$

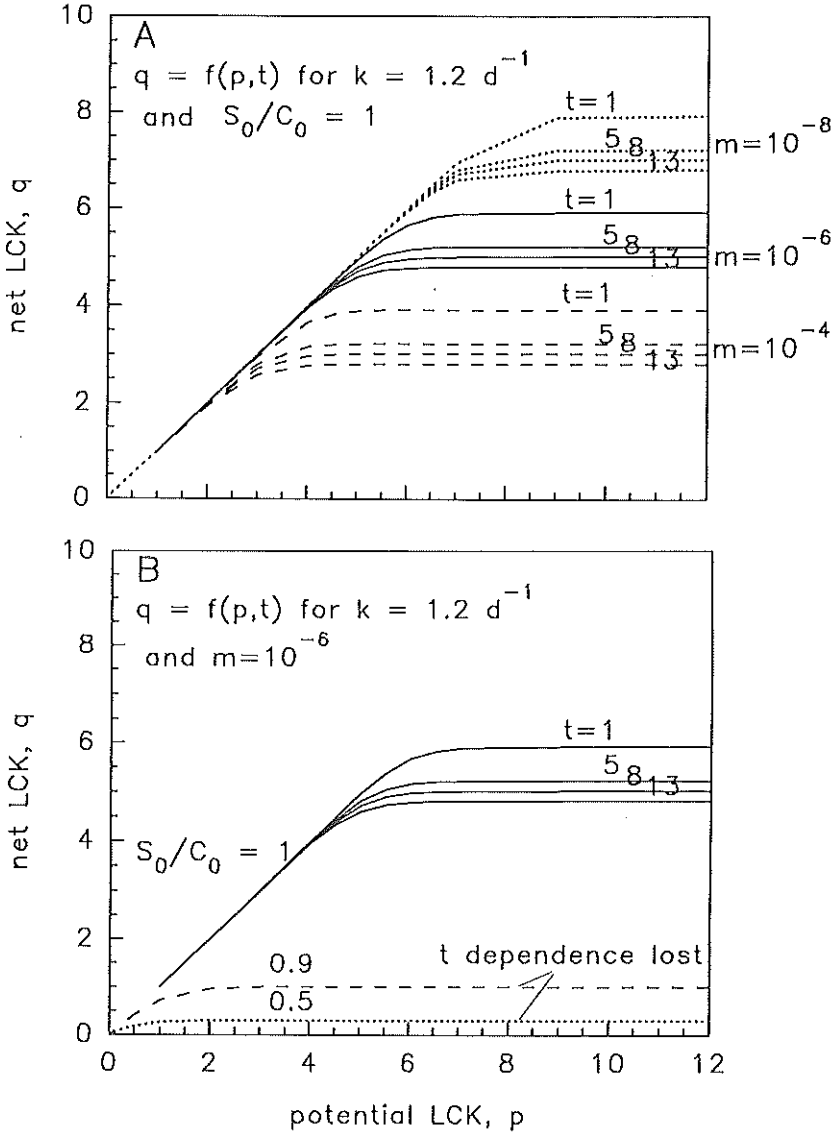
From Eq.(2.32) and Eq.(2.33) the net LCK becomes:

$$q = -\log[1 + (10^{-p} - 1) \cdot f \cdot \exp\{-k \cdot m \cdot t\}]. \quad (2.34)$$

Again it follows that  $q < p$  if  $m > 0$ . Furthermore, for constant  $p$  the value of  $q$  will decrease with time; in other words, due to the phenomenon of resistance development a same drug dose given at a later time point will kill a smaller



Fig. 2.27A/B Relation between net LCK,  $q$ , and potential LCK,  $p$ , for different treatment times  $t$  (days 1, 5, 8 or 13), mutation rates  $m$  ( $10^{-4}$ ,  $10^{-6}$  or  $10^{-8}$ ) and  $f$ -factor (1.0, 0.9 or 0.5: initially, all cells are sensitive (A) or there are 10% or 50% resistant cells (B), respectively).  $q$  reflects the total population reduction,  $C^+ = C \cdot 10^{-q}$  and  $p$  reflects the reduction in sensitive cells,  $S^+ = S \cdot 10^{-p}$  ( $+/-$  denotes population size just after/before drug administration)



fraction of the total population than when given earlier.

The relation between  $q$  and  $p$ , according to the model, is shown in Figs 2.27A and 2.27B for various treatment times, mutation rates and  $f$ -fractions, chosen from probably realistic ranges. If  $f = 1$ , starting with all cells being drug sensitive,  $p$  and  $q$  remain approximately equal over a large range (larger when  $m$  is smaller, or, when  $m$  is constant, larger when treatment is given earlier). Then, a rather abrupt conversion occurs and  $q$  tends to stay constant with increasing  $p$ . *This suggests not only that treatment should be given as early as possible, but also that increasing the drug dose above a certain level would be rather useless. The net LCK is no longer increased by increasing the potential LCK.*

If the  $f$ -factor decreases, i.e., if the initial population already partly consists of resistant cells (10% or 50% shown in Fig. 2.27B), the conversion level decreases considerably while no treatment time dependency is seen anymore. Then there is even less point in giving high dose treatment.

### 2.4.3.3 VALIDATION

How well does the model describe the tumor growth including formation of drug resistant cells?

Comparing Figs 2.25 and 2.27 the linear  $q$ ,CFA-dose relationship is not quite reflected by the model derived  $q$ , $p$  relationship, unless perhaps the mutation rate is very low ( $<10^{-9}$ ) and the population does not initially contain drug resistant cells ( $S/C=1$ ; if resistant cells are present initially, according to the model high values of  $q$  cannot be achieved at all). Still, the observed decrease of  $q$  with treatment time cannot be explained quantitatively by the model, as is shown in Table 2-7. Perhaps  $m$  is not a constant but increases progressively with population size?

Evaluation of the experiments with mixed cell types ( $R/S= 0/100, 1/99, 10/90, 0/100$ ) being inoculated shows that a considerable net LCK can be achieved, if it is assumed a) that growth curves before and after treatment with CFA have identical shapes, b) that these shapes are not influenced by the population's  $R/S$  composition, and c) that death from leukemia occurs at a constant BNML cell burden ( $1.8 \times 10^9$  cells in the bone marrow). Backwards extrapolation of the growth curves (Fig. 2.7) from the time of death until just after treatment time (day 13), and comparing the calculated cell number with the number of cells just before treatment (unperturbed growth curve) yields the sought net LCK values (Table 2-8).

Except for the case of 1% resistant cells the predicted near-constancy of the  $R/S$  ratio is confirmed. Splitting the population size just before treatment in an  $S$  and  $R$  component based on the initial  $R/S$  ratio, and reducing only the  $S$  cells by a factor of  $10^{-4.8}$  while leaving the  $R$  cells unaffected, nearly yields the required

TABLE 2-7 OBSERVED AND MODEL DERIVED DECREASE IN NET LCK, q, AS FUNCTION OF TREATMENT TIME AND MUTATION RATE, FOR CFA = 100 mg/kg AND  $S_0/C_0 = 1$

| treatment time (d) | observed q | model derived q |           |            |
|--------------------|------------|-----------------|-----------|------------|
|                    |            | $m = 10^{-8}$   | $10^{-9}$ | $10^{-10}$ |
| 1                  | -          | 8.0             | 9.0       | 10.0       |
| 5                  | 8.99       | 7.3             | 8.3       | 9.3        |
| 8                  | 5.99       | 7.1             | 8.1       | 9.1        |
| 11                 | 5.0        | 7.0             | 8.0       | 9.0        |
| 13                 | 4.49       | 6.9             | 7.9       | 8.9        |

TABLE 2-8 CALCULATION OF NET LCK, q, AS FUNCTION OF INITIAL R/S MIXTURE

| initial R/S ratio                       | 0/100             | 1/99              | 10/90             | 100/0             |
|---|-------------------|-------------------|-------------------|-------------------|
| death ( $1.8 \times 10^9$ cells) on day | 44                | 39.5              | 34                | 29                |
| nr of cells after treatment*            | $1.6 \times 10^4$ | $8.4 \times 10^5$ | $1.1 \times 10^8$ | $1.0 \times 10^9$ |
| nr of cells before treatment*           | $1.0 \times 10^9$ | $1.0 \times 10^9$ | $1.0 \times 10^9$ | $1.0 \times 10^9$ |
| LCK (q)                                 | 4.8               | 3.1               | 0.9               | 0.0               |

day 0:  $10^7$  BNML *i.v.*; \*: day 13, 100 mg/kg CFA *i.p.*; "regrowth=growth"

TABLE 2-9 BNML CELL POPULATION SIZE AFTER CFA-TREATMENT FOR VARIOUS R/S MIXTURES; CALCULATION AND OBSERVATION

|   |                   |                   |                   |                   |
|---|-------------------|-------------------|-------------------|-------------------|
| nr of cells <i>before</i> treatment     |                   |                   |                   |                   |
| total <sup>a</sup>                      | $1.0 \times 10^9$ | $1.0 \times 10^9$ | $1.0 \times 10^9$ | $1.0 \times 10^9$ |
| assume: R <sup>-</sup> /S <sup>-</sup>  | 0/100             | 1/99              | 10/90             | 100/0             |
| ⇒ resistant, R <sup>-</sup>             | 0                 | $1.0 \times 10^7$ | $1.0 \times 10^8$ | $1.0 \times 10^9$ |
| ⇒ sensitive, S <sup>-</sup>             | $1.0 \times 10^9$ | $9.9 \times 10^8$ | $9.0 \times 10^8$ | 0                 |
| nr of cells <i>after</i> treatment      |                   |                   |                   |                   |
| sensitive <sup>b</sup> , S <sup>+</sup> | $1.6 \times 10^4$ | $1.6 \times 10^4$ | $1.4 \times 10^4$ | 0                 |
| resistant <sup>c</sup> , R <sup>+</sup> | 0                 | $1.0 \times 10^7$ | $1.0 \times 10^8$ | $1.0 \times 10^9$ |
| total                                   | $1.6 \times 10^4$ | $1.0 \times 10^7$ | $1.0 \times 10^8$ | $1.0 \times 10^9$ |
| total, observed <sup>d</sup>            | $1.6 \times 10^4$ | $8.4 \times 10^5$ | $1.1 \times 10^8$ | $1.0 \times 10^9$ |

<sup>a</sup>unperturbed growth, day 13 after  $10^7$  BNML cells *i.v.*

<sup>b</sup>100 mg/kg CFA causes 4.8 LCK to sensitive cells (see Table 2-8), so,  $S^+ = S^- \cdot 10^{-4.8}$

<sup>c</sup>CFA leaves resistant cells unaffected, thus,  $R^+ = R^-$

<sup>d</sup>from MdST and "regrowth=growth", see Table 2-8

number of cells after treatment again (Table 2-9).

By exploring a few properties of a simple model for cell proliferation and formation of drug resistant cells it appeared that results of laboratory experiments could be simulated nicely to some extent, although by this first approach a few observed phenomena could not yet be explained to satisfaction. Therefore, the model must be refined and more detailed investigations, in particular on the validity of its assumptions, will be necessary.

For example, four days ILS for a factor of 10 smaller inoculum of BNML cells implies a population doubling time of 1.2 d for the concept of the inoculum being equivalent to a number of  $ED_{50}$  units to hold. Fitting the growth data, however, results in some 0.8 d. Are these values, thus the mentioned concept, really incompatible, or is the difference not significant and merely due to biological variation?

To help finding answers to questions like this, the sensitivity of the model behavior to changes in the values of variables should be looked into. This way it should be possible to discriminate between absolute model errors (corresponding to established biological impossibilities) and less severe inaccuracies originating from uncertainties in the biological data.

## 2.5 REFERENCES

- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Autom Contr*, AC19:716-723
- Beck WT (1983) Vinca alkaloid-resistant phenotype in cultured human leukemic lymphoblasts. *Cancer Treatm Rep* 67:875-882
- Beck WT (1984) Mechanisms of multiple drug resistance. *Cell Tissue Kinet* 17:667
- Bevington PhR (1969) Data reduction and error analysis. McGraw Hill, New York USA.
- Carter SK (1984) Some thoughts on resistance to cancer chemotherapy. *Cancer Treatm Rev* 11 (supplement A):3-7
- Colly LP, Van Bekkum DW, Hagenbeek A (1984a) Enhanced tumor load reduction after chemotherapy induced recruitment and synchronization in a slowly growing rat leukemia model (BNML) for human acute myelocytic leukemia. *Leuk Res* 8:953-963
- Colly LP, Van Bekkum DW, Hagenbeek A (1984b) Cell kinetic studies after high dose Ara-C and Adriamycin treatment in a slowly growing rat leukemia model (BNML) for human acute myelocytic leukemia. *Leuk Res* 8:945-952
- Colvin OM (1984) Mechanisms of resistance to alkylating agents. *Cell Tissue Kinet* 17:667
- Curt GA, Clendenin NJ, Chabner BA (1984) Drug resistance in cancer. *Cancer Treatm Rep* 68:87-99
- DeFazio A, Tattersall MH (1985) Rapid fluorometric detection of drug resistant tumour cells. *Br J Cancer* 52:553-564
- De Groot CJ, Martens ACM, Hagenbeek A (1992) On the role of aldehyde dehydrogenase in a cyclophosphamide resistant variant of Brown Norway rat acute myelocytic leukemia. In: Hiddemann W, Büchner T, Plunkett W, Keating M, Wörmann B, Andreeff M (eds) *Acute Leukemias. Pharmacokinetics and Management of Relapsed and Refractory Disease*. Springer Verlag, Berlin, Haematology and Blood Transfusion 34:60-64

- De Groot CJ, Martens ACM, Hagenbeek A (1994) Aldehyde dehydrogenase involvement in a variant of the BN rat acute myelocytic leukemia (BNML) that acquired cyclophosphamide resistance *in vivo*. *Cancer Res*, *in press*
- Del Bino G, Skierski JS, Darzynkiewicz Z (1991) The concentration dependent diversity of effects of DNA topoisomerase I and II inhibitors on the cell cycle of HL-60 cells. *Exp Cell Res* 195:485-491
- DeVita VT jr (1983) The relationship between tumor mass and resistance to chemotherapy; implications for surgical adjuvant treatment of cancer. *Cancer* 51:1209-1220
- Dewys WD (1972) A quantitative model for the study of the growth and treatment of a tumor and its metastases with correlation to proliferative state and sensitivity for cyclophosphamide. *Cancer Res* 32:367-373
- Dexter DL, Leith JT (1986) Tumor heterogeneity and drug resistance. *J Clin Oncol* 4:244-257
- Donelli MG, D'Incalci M, Garattini S (1984) Pharmacokinetic studies of anticancer drugs in tumor bearing animals. *Cancer Treatm Rep* 68:381-400
- Dubin N (1976) A stochastic model for immunological feedback in carcinogenesis; analysis and approximations. In: *Lecture Notes in Biomath.* 9, Springer Verlag, New York USA
- El-Beltagi HM, Martens ACM, Lelieveld P, Haroun EA, Hagenbeek A (1993) Acetyldinaline: a new oral cytostatic drug with impressive differential activity against leukemic cells and normal stem cells. Preclinical studies in a relevant rat model for human acute myelocytic leukemia. *Cancer Res* 53:3008-3014
- Fojo A, Akiyama S, Gottesman MM, Pastan I (1985) Reduced drug accumulation in multiple drug-resistant human KB carcinoma cell lines. *Cancer Res* 45:3002-3007
- Freund M, Gillers S, Hinrichs F, Baars A, Meran J, Korfer A, Link H, Poliwoda H (1991) Five-day 4'-(9-acridinylamino)methanesulphon-m-anisidide and intermediate dose cytosine arabinoside in high risk relapsing or refractory acute myeloid leukemia. *J Cancer Res Clin Oncol* 117:489-492
- Friedman OM, Myles A and Colvin M (1979) Cyclophosphamide and related phosphoramidate mustards: current status and future prospects. *Adv Cancer Chemother* 1:143-204
- Garfinkel D (1983) Modeling of inherently complex biological systems: problems, strategies, methods. *Math Biosci* 72:131-139
- Gerlach JH, Kartner N, Bell DR, Ling V (1986) Multidrug resistance. *Cancer Surv* 5:25-46
- Goldie JH and Coldman AJ (1985) A model for tumor response to chemotherapy: an integration of the stem cell and somatic mutation hypotheses. *Cancer Invest* 3:553-564
- Goldstein DP (1975) Gestational trophoblastic neoplasms. In: Greenspan EM (ed.) *Clinical cancer chemotherapy*. Raven Press, New York, 257-284
- Goldstein LJ, Galski H, Fojo A, Willingham N, Lai SL, Gazdar A, Pirker R, Green A, Crist W, Brodeur GM, Lieber M, Cossman J, Gottesman MM, Pastan I (1989) Expression of a multidrug resistance gene in human cancers. *J Natl Cancer Inst* 81:116-124
- Gorzycza W, Gong J, Ardelt B, Traganos F, Darzynkiewicz Z (1993) The cell cycle related differences in susceptibility of HL-60 cells to apoptosis induced by various antitumour agents. *Cancer Res* 53:3186-3192
- Griswold DP jr (1986) Body burden of cancer in relationship to therapeutic outcome: consideration of preclinical evidence. *Cancer Treatm Rep* 70:81-85
- Hagenbeek A, Martens ACM (1981) The proliferative state of normal hemopoietic stem cells during the progression of leukemia; studies in the BN acute myelocytic leukemia. *Leuk Res* 5:141-147
- Hagenbeek A, Martens ACM (1982) High dose cyclophosphamide treatment of acute myelocytic leukemia; studies in the BNML rat model. *Eur J Cancer Clin Onc* 18:763-769
- Hagenbeek A, Martens ACM (1985) Reinfusion of leukemic cells with the autologous marrow graft: preclinical studies on lodging and regrowth of leukemia. *Leuk Res* 9:1389-1395

- Hagenbeek A, Martens ACM (1986) AMSA: In vivo log cell kill for leukemic clonogenic cells versus toxicity for normal hemopoietic stem cells in a rat model for human acute myelocytic leukemia (BNML). *Eur J Cancer Clin Oncol* 22:1255-1258
- Hagenbeek A, Martens ACM (1987a) On the fate of leukemic cells infused with the autologous marrow graft. In: Büchner T, Schellong G, Hiddemann W, Urbanitz D, Ritter J (eds) *Acute Leukemias. Prognostic factors and treatment strategies*. Springer Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, pp:553-558
- Hagenbeek A, Martens ACM, Colly LP (1987b) *In vivo* development of cytosine arabinoside resistance in the BN acute myelocytic leukemia. *Semin Oncol* 14-2 (suppl. 1):202-206
- Hagenbeek A (1992) Minimal residual disease in leukemia: state of the art 1991. *Leukemia* 6, Suppl 2:12-16
- Hall A, Cattan AR, Proctor SJ (1989) Mechanisms of drug resistance in acute leukaemia. *Leuk Res* 13:351-356
- Hanson F, Tier C (1982) A stochastic model of tumor growth. *Math Biosci* 61:73-100
- Hayes JD, Wolf CR (1990) Molecular mechanisms of drug resistance. *Biochem J* 272:281-295
- Hewitt HB (1958) Studies of the dissemination and quantitative transplantation of a lymphocytic leukemia of CBA mice. *Br J Cancer* 12:378-401
- Holmes J, Jacobs A, Carter G, Janowska-Wieczorek A, Padua RA (1989) Multidrug resistance in haemopoietic cell lines, myelodysplastic syndromes and acute myeloblastic leukemia. *Br J Haematol* 72:40-44
- Hotz MA, Del Bino G, Lassota P, Traganos F, Darzynkiewicz Z (1992) Cytostatic and cytotoxic effects of fostriecin on human promyelocytic HL-60 and lymphocytic MOLT-4 leukemic cells. *Cancer Res* 52:1530-1535
- Houghton PJ, Tew KD, Taylor DM (1976) Some studies on the distribution and effects of cyclophosphamide in normal and neoplastic tissues. *Cancer Treatm Rep* 60:459-464
- Hryniuk W, Bush H (1984) The importance of dose intensity in chemotherapy of metastatic breast cancer. *J Clin Oncol* 2:1281-1288
- Iversen S, Arley N (1950) On the mechanism of experimental carcinogenesis. *Acta Path Microbiol Scand* 27:773-803
- Jackson H, Bock M, Jackson NC, Lendon M (1983) The testis - a protected environment for leukaemic cells against cyclophosphamide in a mouse model. *Cancer Chemother Pharmacol* 11:200-202
- Jehn U, Heinemann V (1991) New drugs in the treatment of acute and chronic leukemia with some emphasis on m-AMSA. *Anticancer Res* 11:705-711
- Kartner N, Riordan JR, Ling V (1983) Cell surface P-glycoprotein associated with multidrug resistance in mammalian cells. *Science* 221:1285-1288
- Kaye S, Merry S (1985) Tumor cell resistance to anthracyclines - a review. *Cancer Chemother Pharmacol* 14:96-103
- Lahiri SK, Keizer HJ, Van Putten LM (1970) The efficiency of the assay for haemopoietic colony forming cells. *Cell Tissue Kinet* 3:355
- Laird AK (1964) The dynamics of tumor growth. *Br J Cancer* 28:490-502
- Ling V, Thompson LH (1974) Reduced permeability in CHO cells as a mechanism of resistance to colchicine. *J Cell Physiol* 83:103-116
- Ling V, Kartner N, Sudo T, Siminovitch L, Riordan JR (1983) Multidrug resistant phenotype in Chinese hamster ovary cells. *Cancer Treatm Rep* 67:869-874
- Luria SE, Delbrück M (1943) Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* 28:491-511
- Ma DDF, Davey RA, Harman DH, Isbister JP, Scurr RD, Mackertich SM, Dowden G, Bell DR (1987) Detection of a multiple resistant phenotype in acute nonlymphoblastic leukemia. *Lancet* i:135-137

- MacKillop WJ (1986) Intrinsic versus acquired drug resistance. *Cancer Treatm Rep* 70:817-818
- Martens ACM, Johnson RJ, Kaizer H, Hagenbeek A (1984) Characteristics of a monoclonal antibody (RM124) against acute myelocytic leukemia cells. *Exp Hematol* 12:667-671
- Martens ACM, Hagenbeek A (1985) Detection of minimal disease in acute leukemia using flow cytometry; studies in a rat model for human acute leukemia. *Cytometry* 6:342-347
- Martens ACM, Van Bekkum DW, Hagenbeek A (1990) The BN acute myelocytic leukemia (BNML). A rat model for studying human acute myelocytic leukemia. *Leukemia* 4:241-257
- Martens ACM, De Groot CJ, Hagenbeek A (1991) Development and characterization of a cyclophosphamide resistant variant of the BNML rat model for acute myelocytic leukemia (AML). *Eur J Cancer* 27:161-166
- Matthews JH (1988) Cell cycle characteristics in acute nonlymphocytic leukemia. *Blood* 71:532-534
- McGown AT, Fox BW (1986) A proposed mechanism of resistance to cyclophosphamide and phosphoramidate mustard in a Yoshida cell line in vitro. *Cancer Chemother Pharmacol* 17:223-226
- McMillan TJ, Stephens TC, Steel GG (1985) Development of drug resistance in a murine mammary tumour. *Br J Cancer* 52:823-832
- Miller LP, Pyesmany AF, Wolff LJ, Rogers PL, Siegel SE, Wells RJ, Buckley JD, Hammond GD (1991) Successful reinduction therapy with ansacrine and cycloytidin in acute nonlymphoblastic leukemia in children. *Cancer* 67:2235-2240
- Morrow CS, Cowan KH (1990) Glutathione S-transferases and drug resistance. *Cancer Cells* 2:15-22
- Nelson JA (1985) Bases for heterogeneity in antitumor drug response. *The Cancer Bull* 37:192-195
- Neyman J, Scott E (1967) Statistical aspects of the problems of carcinogenesis. In: Fifth Berkeley Symposium on Math. Statistics and Probab. Univ. of Calif. Press, Berkeley USA, 745-776
- Nicolson GL, Lotan R (1986) Preventing diversification of malignant tumor cells during therapy. *Clin Exp Metastasis* 4:231-235
- Nooter K, Sonneveld P, Deurloo J, Oostrum R, Martens ACM, Hagenbeek A (1985) Differences in the pharmacokinetics of daunomycin in normal and leukemic rats. *Cancer Res* 45:4020-4026
- Nooter K, Sonneveld P, Deurloo J, Oostrum R, Schultz FW, Martens ACM, Hagenbeek A (1986) Tissue distribution and myelotoxicity of daunomycin in the rat: rapid bolus injection versus continuous infusion. *Eur J Cancer Clin Oncol* 22:801-806
- Nooter K, Oostrum R, Jonker R, Van Dekken H, Stokdijk W, Van den Engh G (1989) Effect of cyclosporin A on daunorubicin accumulation in multidrug resistant P388 leukemia cells measured by real-time flow cytometry. *Cancer Chemother Pharmacol* 23:296-300
- Nooter K, Sonneveld P, Oostrum R, Herveijer H, Hagenbeek A, Valerio D (1990a) Overexpression of the *mdr1* gene in blast cells from patients with acute myeloblastic leukemia is associated with decreased anthracycline accumulation that can be restored by cyclosporin-A. *Int J Cancer* 45:263-268
- Nooter K, De Vries A, Martens ACM, Hagenbeek A (1990b) Effect of cyclophosphamide pretreatment on daunorubicin in rat acute leukaemia model. *Eur J Cancer* 26:729-732
- Norton L, Simon R (1977) Tumor size, sensivity to chemotherapy and design of treatment schedules. *Cancer Treatm Rep* 61:1307-1317
- Norton L, Simon R (1986) The Norton-Simon hypothesis revisited. *Cancer Treatm Rep* 70:163-169
- Nowrousian MR, Schmidt CG (1984) Comparative effects of ASTA Z 7557 (INN mafosfamide) and cyclophosphamide on hematopoiesis in mice. *Invest New Drugs* 2:207-213
- Pastan I, Gottesman M (1987) Multiple drug resistance in human cancer. *New England J Med*

316:1388-1393

- Pohl J: ASTA Werke AG, Bielefeld, BRD (personal communication, 1986)
- Prajneshu A (1979) A stochastic Gompertz model with heredity effect. *J Math Biol* 8:189-196
- Rai KR, Sawitsky A, Cronkite EP, Chanana AD, Levy RV, Pasternak BS (1975) Clinical staging of chronic lymphocytic leukemia. *Blood* 46:219-234
- Rius C, Zorilla AR, Cabanas C, Mata F, Bernabeu C, Aller P (1991) Differentiation of human promyelocytic leukemia U-937 cells with DNA topoisomerase II inhibitors: induction of vimentin gene expression. *Mol Pharmacol* 39:442-448
- Roninson IBB, Abelson HT, Housman DE, Howell N, Varshavsky A (1984) Amplification of specific DNA sequences correlates with multidrug resistance in Chinese hamster cells. *Nature* 309:626-628
- Roninson IBB (1987) Molecular mechanism of multidrug resistance in tumor cells. *Clin Physiol Biochem* 5:140-151
- Schmid FA, Hutchison DJ, Otter Gm, Stock C (1976) Development of resistance to combinations of six antimetabolites in mice with L1210 leukemia. *Cancer Treatm Rep* 60:23-27
- Schultz FW, Martens ACM, Hagenbeek A (1987) Computer simulation of the progression of an acute myelocytic leukemia in the Brown Norway rat. *Comput Math Applic* 14:751-761
- Schultz FW, van Dongen JJM, Hählen K and Hagenbeek A (1989) Time-history of the malignant population in the peripheral blood of a T-cell acute lymphoblastic leukemia patient: a pilot study. *Computers Math Applic* 18:929-936
- Schultz FW, Hagenbeek A (1991) Simulation studies on the regrowth of acute myeloid leukemia after autologous bone marrow transplantation. In: Arino O, Axelrod D, Kimmel M (eds). *Proceedings of the Second International Conference on Mathematical Population Dynamics*, Rutgers University, New Brunswick, NJ, 17-20 May 1989. Marcel Dekker Inc, New York, pp 689-709
- Selby P (1984) Acquired resistance to cancer chemotherapy. *Br Med J* 288:1252-1253
- Shoemaker RH, Curt GA, Carney DN (1983) Evidence for multidrug resistant cells in human tumor cell populations. *Cancer Treatm Rep* 67:883-888
- Simpson-Herren L, Lloyd HH (1970) Kinetic parameters and growth curves for experimental tumor systems. *Cancer Chemother Rep* 54:143-174
- Skipper HE, Schabel FM, Wilcox WS (1964) Experimental evaluation of potential anticancer agents XII; on the criteria and kinetics associated with curability of experimental leukemia. *Cancer Chemother Rep* 35:1-111
- Skipper HE (1983) The forty-year-old mutation theory of Luria and Delbrück and its pertinence to cancer chemotherapy. *Adv Cancer Res* 40:331-363
- Skipper HE (1986a) Laboratory models: some historical perspective. *Cancer Treatm Rep* 70:3-7
- Skipper HE (1986b) Some thoughts on intrinsic versus acquired drug resistance in cancers that are classified as responsive, refractory or very refractory to chemotherapy. Booklet 10, Southern Research Institute, Birmingham, Alabama
- Smith CE, Tuckwell HC (1974) Some stochastic growth processes. In: *Math. Problems in Biol., Victoria Conf., Lecture Notes in Biomath.* 2, Springer Verlag, New York USA, p211-225
- Sonneveld P, Mulder JA, Van Bekkum DW (1981) Cytotoxicity of doxorubicin for normal hematopoietic and acute myelocytic leukemia cells of the rats. *Cancer Chemother Pharmacol* 5:167-173
- Stark GR (1986) Cancer chemotherapy. Progress in understanding multidrug resistance. *Nature* 324:407-408
- Steel GG (1977) *Growth kinetics of tumors*. Clarendon, Oxford UK
- Stephens TC, Adams K, Peacock JH (1984) Identification of a subpopulation of MeCCNU resistant cells in previously untreated Lewis lung tumours. *Br J Cancer* 50:77-83
- Stephens TC, Adams K, Peacock JH (1986) Emergence of nitrosourea resistant sublines of Lewis



- lung tumour following MeCCNU treatment in vivo. *Br J Cancer* 53:237-245
- Sullivan PW, Salmon SE (1972) Kinetics of tumor growth and regression in IgG multiple myeloma. *J Clin Invest* 51:1697-1708
- Swan GW (1977) Some current mathematical topics in cancer research. Univ Microfilms Inter, Ann Arbor USA
- Tannock IF (1986) Experimental chemotherapy and concepts related to the cell cycle. *Int J Radiat Biol* 49:335-355
- Tsuruo T, Sugimoto Y, Hamada H, Roninson I, Okumura M, Adachi K, Morishima Y, Ohno R (1987) Detection of multidrug resistance markers, P-glycoprotein and *mdr1* mRNA, in human leukemic cells. *Jap J Cancer Res* 78:1415-1419
- Van Bekkum DW (1977) The appearance of the multipotential hemopoietic stem cell. In: Baum SJ, Ledney GD (eds) *Experimental hematology today 1977*. Springer Verlag, New York, pp:3-10
- Van Putten LM (1986) Of tumours in mice and man, the different roles of somatic mutation in treatment failure. *Eur J Cancer Clin Oncol* 22:753-755
- Waxman DJ (1990) Glutathione S-transferases: role in alkylating agent resistance and possible target for modulation chemotherapy - a review. *Cancer Res* 50:6449-6454
- Wette R, Katz IN, Robin EY (1974) Stochastic processes for solid tumor kinetics. *Math Biosci* 19:231-255
- Whittemore A, Keller JB (1978) Quantitative theories of carcinogenesis. *SIAM Rev.* 20:1-30
- Wilcox WS (1966) The last surviving cancer cell: the chances of killing it. *Cancer Chemother Rep* 50:541-542
- Zajicek G (1986) Resistance to cancer chemotherapy. *Med Hypotheses* 19:103-112



## Chapter 3

# Cell Population Dynamics of Childhood T-Cell Acute Lymphoblastic Leukemia<sup>1</sup>

The average adult leukemia patient will carry at time of diagnosis a leukemic cell load of some  $10^{12}$  cells (approx. 1 kg). With modern (remission induction) chemotherapy most patients (children 95%, adults 80%, [Smith et al., 1986]) will soon enter an apparent disease free state (complete remission), i.e., the symptoms disappear and in standard tests no leukemic cells can be found anymore. However, that the number of malignant cells has disappeared below the clinical level of detection does not mean that all clonogenic leukemic cells have been eradicated. This is revealed by the fact that in many patients (children 30-40%, adults 70-80%) the disease will return eventually (relapse), due to re-growth of surviving cells. The clinical detection level is now about 1 percent, i.e., the occurrence of less than 1 malignant cell in 100 normal cells will go unnoticed [Hagenbeek and Martens, 1985]. In other words, an apparently healthy person having 1 kg of marrow may carry an undetectable malignant cell burden (minimal residual disease) of some 0.01 kg, corresponding with  $10^{10}$  leukemic cells. The clinician's problem is how to anticipate an imminent relapse and how to tune his (maintenance) chemotherapy schedule to prevent or delay it, bearing in mind that patients should be spared the unnecessary toxicity that is inherent to intensive chemotherapy. So, in fact this makes up an optimization problem.

Therefore, it is most important to gather information on the time-history of the malignant population during and after treatment. Before treatment optimization can be dealt with, it must be known how the malignant population develops, how it reacts—in terms of surviving cell numbers—to the drugs administered, and how it will regrow after the chemotherapy.

In the present study datapoints were gathered by means of immunological marker analysis [Van Dongen et al., 1986] revealing at certain time instants during therapy the size of the population of malignant cells in the peripheral blood (PB) and the bone marrow (BM) of several children with T-cell acute lymphoblastic leukemia (T-ALL). The raw PB data are examined and compared. Through simulation and curve fitting it is tested whether certain models for population growth and drug action are applicable. A particular case is studied in detail. The time-history of this population is regarded as the output of a dynamic

---

<sup>1</sup>To a large extent the contents of this chapter were published in: Schultz FW, Van Dongen JJ, Hähnen K and Hagenbeek A: Time-history of the malignant population in the peripheral blood of a child with T-cell acute lymphoblastic leukemia: a pilot study. *Comput Math Applic* (1989) 18:929-936

system that responds to an input signal consisting of the clinically applied chemotherapy schedule. The relation between in- and output, as well as the unperturbed output behavior, must be identified. In the first instance, (partially) making use of already accumulated knowledge, it is tried to describe the evolution of the population in mathematical terms, in such way that the observed datapoints can be explained.

### 3.1 METHODS AND MATERIALS

#### 3.1.1 Detection and Quantification of T-ALL Cells by Immunological Marker Analysis

The malignant cells in nearly all cases of T-ALL express the nuclear enzyme terminal deoxynucleotidyl transferase (TdT) and T-cell markers such as the CD5 antigen on the membrane. Normally, cells with the CD5<sup>+</sup>/TdT<sup>+</sup> phenotype are found in the thymus only [Van Dongen et al., 1985]. The presence of these cells on locations outside the thymus is therefore indicative of the presence of T-ALL. By use of the CD5<sup>+</sup>/TdT<sup>+</sup> double immunofluorescence (IF) staining technique, it is possible to detect one CD5<sup>+</sup>/TdT<sup>+</sup> cell among 10,000 or even 100,000 normal cells, as has been established in a series of previous dilution experiments [Van Dongen et al., 1986].

During remission induction (RI) and follow-up of several T-ALL patients many peripheral blood and bone marrow samples were taken and analyzed for the presence of residual T-ALL cells by use of the CD5/TdT double IF staining. For this purpose the mononuclear cells (MNC) were isolated from the obtained PB and BM samples by ficoll density centrifugation (Ficoll-Paque; density 1.077 g/cm<sup>3</sup>). The MNC were incubated with the anti-CD5 monoclonal antibody (McAb) Leu-1 (Becton Dickinson, Sunnyvale, CA) and subsequently with a tetramethylrhodamine-isothiocyanate (TRITC) conjugated goat anti-mouse immunoglobulin (Ig) antiserum. Afterwards, at least two cyto-centrifuge preparations per sample were made, each containing at least 25,000 cells. These preparations were fixed in methanol and subjected to an indirect staining for TdT by use of a goat anti-TdT antiserum (Supertechs, Bethesda, MD) and a fluorescein-isothiocyanate (FITC) conjugated goat anti-rabbit Ig antiserum as a second step reagent. The cyto-centrifuge preparations were evaluated on Zeiss fluorescence microscopes, equipped with phase-contrast facilities and filter combinations for the selective visualization of FITC and TRITC. The CD5<sup>+</sup>/TdT<sup>+</sup> cells (i.e., T-ALL cells) present were quantified as percentages double positive cells per MNC, always by first determining the fraction of TdT<sup>+</sup> cells per MNC, followed by counting those TdT<sup>+</sup> cells that were also positive for the second marker. For this purpose at least 1000 TdT<sup>+</sup> cells were evaluated each time.

TABLE 3-1 T-ALL PATIENT INFORMATION

| patient | sex | birth date | patient | sex | birth date |   |        |
|---------|-----|------------|---------|-----|------------|---|--------|
| 1       | MZ  | f          | 190777  | 5   | ER         | m | 010281 |
| 2       | IA  | f          | 131279  | 6   | KG         | f | 140876 |
| 3       | RS  | m          | 190772  | 7   | CDJ        | f | 170375 |
| 4       | PM  | m          | 010980  |     |            |   |        |

TABLE 3-2 QUALITATIVE DESCRIPTION OF COURSE OF DISEASE BASED ON 'RAW DATA'

| patient | T-ALL cells/mm <sup>3</sup> of PB at start of therapy | time until remission (wks) | time until relapse (wks) | IF neg. period (wks) | IF pos. at (wk) |
|---------|---|----------------------------|--------------------------|----------------------|-----------------|
| MZ      | 1.48x10 <sup>8</sup>                                  | 3                          | 126                      | 9-98                 | 111             |
| IA      | 1.46x10 <sup>8</sup>                                  | 3                          | 59                       | 13-34                | 58              |
|         | 3.08x10 <sup>5</sup>                                  | 1                          | 14                       | remains positive     |                 |
| RS      | 1.33x10 <sup>8</sup>                                  | 3                          | *1                       | 6-20                 | 27 <sup>x</sup> |
| PM      | 3.95x10 <sup>7</sup>                                  | 1                          | *2                       | 3-30 <sup>+</sup>    | -               |
| ER      | 4.8 x10 <sup>6</sup>                                  | 1                          | *3                       | 7-12 <sup>+</sup>    | -               |
| KG      | 4.9 x10 <sup>5</sup>                                  | 1                          | *4                       | 10-13 <sup>+</sup>   | -               |
| CDJ     | 2.0 x10 <sup>5</sup>                                  | 1                          | 24                       | remains positive     |                 |
|         | 9.14x10 <sup>2</sup>                                  | 1                          | 23                       | 4-12                 | 16              |
|         | 2.57x10 <sup>2</sup>                                  | 1                          | 14                       | remains positive     |                 |

\*i: still in remission after 143 wks (i=1), 30 wks (i=2), 12 wks (i=3), 13 wks (i=4)

+ : latest observation; still IF negative (i.e., nothing detected by immunofluorescence)

x: alternatingly positive and negative with approx. 21 wk period

remission: < 10<sup>5</sup> malignant cells/mm<sup>3</sup> PB (< 10<sup>2</sup> for CDJ)

relapse: at 10<sup>5</sup> malignant cells/mm<sup>3</sup> PB (10<sup>2</sup> for CDJ)

For the mathematical analysis of the data the percentages of T-ALL cells per MNC were recalculated to the absolute number of T-ALL cells per mm<sup>3</sup> of PB or BM. It is a small complication that within the MNC population after ficoll density centrifugation still up to 15-20% myeloid cells are present. Those myeloid cells can easily be detected, since they strongly express the CD15

antigen, which can be marked by use of the McAb VIM-D5. Therefore, the percentage of T-ALL cells per MNC (= Z) was first corrected for the presence of CD15<sup>+</sup> cells per MNC (= X) to obtain the percentage of T-ALL cells per "true" MNC (i.e., MNC not polluted with myeloid cells). The percentage of myeloid cells per total leukocytes (= Y) was obtained from differential counts of the PB or BM. This percentage was used to convert the percentage of T-ALL cells per "true" MNC to the percentage of T-ALL cells per total leukocyte population (=  $Z \cdot (100 - X)^{-1} \cdot (100 - Y)$ ). Subsequently, the total number of T-ALL cells per mm<sup>3</sup> was calculated by use of the BM cell count or the white blood cell count (WBC) per mm<sup>3</sup> (= C).

These calculations can be summarized in the following formula:

$$\text{nr. of T-ALL} = Z \cdot 10^{-2} \cdot (100 - X)^{-1} \cdot (100 - Y) \cdot C, \quad (3.1)$$

where nr. of T-ALL denotes the total number of T-ALL cells per mm<sup>3</sup> PB or BM.

### 3.1.2 Patients

A total of seven patients has been followed with respect to the time-course of their disease, starting at diagnosis of T-ALL (Tables 3-1 and 3-2). The initial leukemic cell load varied from  $1.5 \times 10^8$  to  $2 \times 10^5$  T-ALL cells per mm<sup>3</sup> in peripheral blood. In all patients remission was induced within 1 to 3 weeks (patients with high leukemic cell loads took longer). Four patients (RS, PM, ER, KG) remained in first remission without detectable residual disease. Follow-up times, however, are still rather short for them (12-30 weeks), except for patient RS (143 weeks) in whose PB by use of IF techniques low levels of T-ALL cells were found regularly after week 27.

Two patients, MZ and IA, relapsed after 126 and 59 weeks, respectively. This could be foretold a few weeks in advance, in view of the increase of the malignant cell load as seen with the IF method. A second RI therapy in patient IA apparently was successful within 1 week and lasted until week 14. The remission was not as deep as the first one, as is shown not only by the shorter duration of the remission, but also by the continuous presence of T-ALL cells during this period (according to the IF method). The last patient, CDJ, relapsed twice. As the data for this patient are the most extensive, she will be discussed in detail below.

### 3.1.3 Case Report for Patient CDJ

After T-ALL had been diagnosed the 8 year old girl CDJ was monitored during a period of about 500 days. During this follow-up period the patient had two

relapses of the disease. High intensity chemotherapy to obtain remission induction was applied three times, starting on days 0 (diagnosis), 198 (first relapse) and 365 (second relapse). Such RI therapy was given during 3 weeks, always employing a same multi-drug combination according to a standard RI protocol. The first two RI therapies were successful and were therefore followed by a constant low intensity continuous maintenance therapy (CMT) schedule consisting of 6-mercaptopurine/methotrexate and, during the first remission, of two additional cycles of prednisone treatment, starting on days 121 and 177. The second relapse did not respond sufficiently to chemotherapy and the patient died.

### 3.1.4 Mathematical Analysis

Only time courses of T-ALL cell population sizes in PB have been analyzed because in this compartment the observations were the most abundant.

Four phases of analysis can be distinguished:

- 1- Inspection of raw data, i.e., the patterns of the observed numbers of T-ALL cells per  $\text{mm}^3$  in PB plotted semi-logarithmically against time.
- 2- Log-linear regression of cell numbers,  $C$ , on time,  $t$ , to find halftimes and doubling times ( $T_{1/2}$  and  $T_2$ , respectively) at various stages of therapy and follow-up, between times  $t$  and  $t_0$  when growth or decline of the malignant cell population is presumably exponential:

$$C(t) = C(t_0) \cdot \exp\left\{\frac{\ln(2)}{T} \cdot (t - t_0)\right\}, \quad (3.2)$$

where  $T = -T_{1/2}$  or  $T = T_2$ .

- 3- Simulation studies; to estimate the cell kill effect of daily drug doses, under the assumption that a same drug dose kills a same fraction of cells (log cell kill hypothesis [Skipper et al., 1964]; see below, section 3.1.4.4) and exponential regrowth with doubling times as determined under 2.
- 4- Curve fitting; malignant cell population growth and the influence of therapy can be modeled by some mathematical function that, after successful fitting to the datapoints, yields an empirical description of the development of this cell population.

**3.1.4.1 Inspection of Raw Data.** By simply plotting the observations on semi-log paper a first impression is obtained about how the malignant cell population evolves with time. It allows any correspondence—or noncorrespondence—with known events, e.g., start and end of a therapy course, to be noticed at a glance.

The estimated accuracy of the cell population size as observed with immunofluorescence in general is better than 10 percent. However, when the

uncorrected percentage of T-ALL cells per MNC ( $Z$  in Eq.(3.1)) is less than 0.1, the uncertainty may rise up to 200%. For some datapoints the degree of contamination of the MNC population with myeloid cells (percentage  $X$  in Eq.(3.1)) could not be determined. An average value for  $X$  was substituted then. Therefore, the datapoints concerned are slightly less certain. Further remarks on the accuracy of the observations can be found in Appendix C.

**3.1.4.2 Log-linear Regression.** From the inspection of raw data (section 3.1.4.1) it appears that the T-ALL cell populations in the periods just after start of RI therapy and just before and after relapse evolve almost according to exponential growth. Application of log-linear regression to the datapoints concerned therefore seems appropriate to estimate the corresponding halftimes and doubling times.

**3.1.4.3 Simulations.** For various patients (PM, ER, KG) the influence of the periodically administered drug doses was evaluated by modeling a constant fraction of T-ALL cells being killed by each separate dose. Dealing with equal doses and constant time intervals (mostly daily administrations) the cell population decreases exponentially and a halftime can be calculated. The killing effect of each drug dose, expressed as a number of logs (log cell kill, LCK; see section 3.1.4.4), can be estimated by matching the simulation results to the actually observed population size—time datapoints.

**3.1.4.4 Curve fitting.** A very general growth equation describing the time course of a cell population of size  $C(t)$  at time  $t$  is given by:

$$C(t + \Delta t) = C(t) + \Delta t \cdot \frac{dC(t)}{dt}. \quad (3.3)$$

In the present case the time increment  $\Delta t$  was chosen 0.01 d in a series of computer simulations. The growth rate  $dC/dt$  was derived from the experimental data, bearing in mind some already available knowledge about tumor growth [Steel, 1977], both under unperturbed circumstances and under influence of chemotherapy, which will be briefly elucidated below.

*Unperturbed growth.* The unperturbed development of a population of leukemia cells is often characterized by exponential growth, i.e., the population doubling time,  $T_2$ , is constant and on a log-linear plot a straight line is seen. In particular this is commonly true in the early stage of the disease. At a later stage the doubling time may decrease more and more with time, and the population evolves to a steady state plateau phase (constant size). For instance, in the Brown Norway rat acute myelocytic leukemia—whose characteristics correspond closely



to the human acute myeloid leukemia [Van Bekkum and Hagenbeek, 1977]—such behavior is observed. After an exponential growth phase  $T_2$  starts decreasing exponentially, which can be described with a Gompertz function [Schultz et al., 1986].

In general the increase of the population size per unit time can be related to the current population size as follows:

$$\frac{dC(t)}{dt} = GF \cdot C(t), \tag{3.4}$$

where  $GF = \ln 2/T_2$  is constant in the exponential growth phase, and for Gompertzian growth  $GF = A \cdot \ln(C_{\max}/C)$ , thus depends on the current and the maximum population size (plateau phase), as well as on a retardation constant,  $A$  [ $\text{time}^{-1}$ ].

This growth fraction  $GF$  can be regarded as that part of the population that doubles its size during time interval  $dt$ . It is related to biological variables like the generation time of the cell line (time from birth of a cell until cell division), the ratio of resting and actively proliferating cells, the loss of cells from the population, and environmental factors.

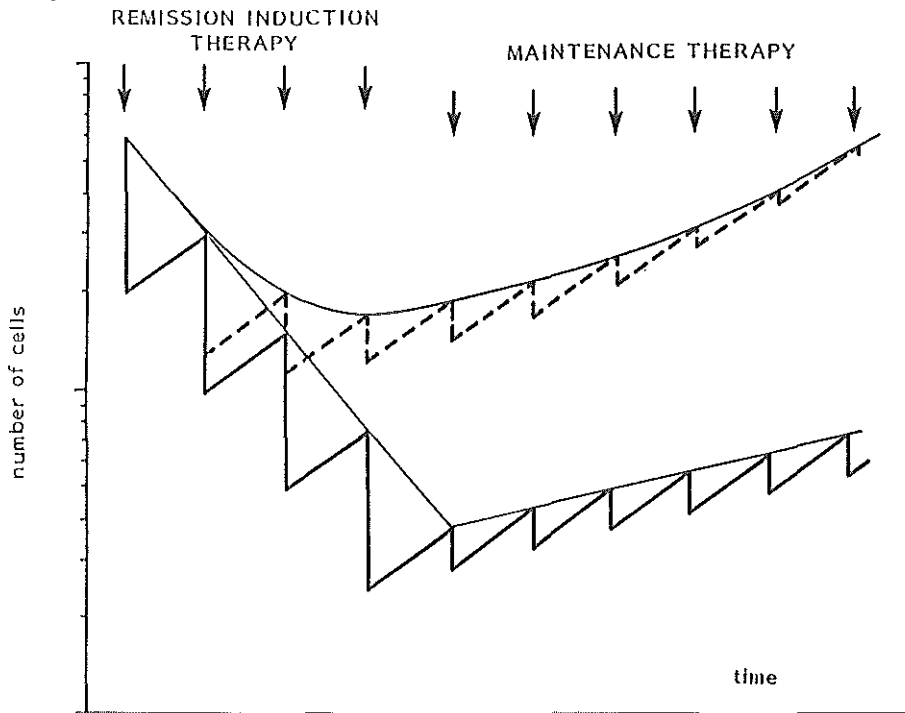
*Influence of chemotherapy.* Based on experiments measuring the increased lifespan of tumor bearing laboratory animals due to chemotherapy the "log cell kill" hypothesis [Skipper et al., 1964] says that a given dose of a given cytostatic agent always kills a constant fraction of cells, rather than a constant absolute number. After this instantaneous reduction in cell number the drug effect has disappeared and the surviving population regrows in a way similar to the unperturbed growth. This hypothesis has been widely accepted for describing chemotherapeutic impact on the size of a cell population (e.g., [Birkhead and Gregory, 1984]). Thus, a drug dose is said to have an effect of  $B$  log cell kill if the population sizes just before (time  $t^-$ ) and just after (time  $t^+$ ) the administration are related by:

$$C(t^+) = C(t^-) \cdot 10^{-B}. \tag{3.5}$$

If a series of equally spaced constant drug doses are administered to a population in exponential growth, and assuming that the doubling time during regrowth will be identical with that of the original unperturbed population, then, on a log-linear plot, the population's time-history will be a regular saw-tooth shaped curve. Because each sequential dose will kill a constant fraction of cells this curve will progress along a straight line, whose slope (either up- or downward) depends on the log cell kill factor per dose and on the value of  $T_2$ .

The time-history of the average population size—i.e.,  $C(t) =$

Fig. 3.1 Qualitative influence of remission induction and maintenance chemotherapy (high and low intensity, respectively) on the size of a cell population under a) *solid line*: the "log cell kill" hypothesis or b) *broken line*: development of therapy resistance. The thin solid lines show a "smoothed" time-history (connecting datapoints just before each new drug administration)

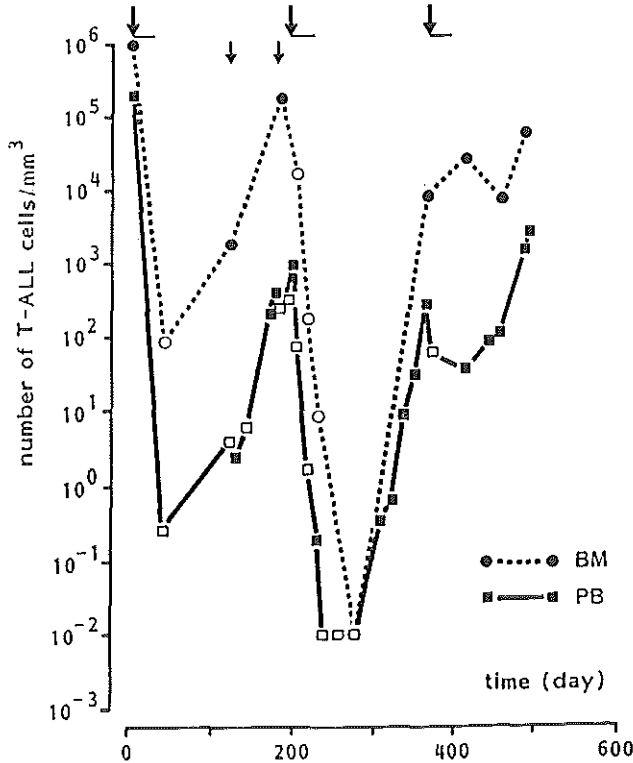


$\frac{1}{2}\{C(t) + C(t^+)\}$ —can then be described with the equation:

$$\frac{dC(t)}{dt} = GF \cdot \{1 - K \cdot L(t)\} \cdot C(t), \quad (3.6)$$

where—as in [Norton and Simon, 1977]— $K$  is a scaling constant and  $L(t)$  a function that describes the therapy level, in this case having a constant value as well. Substitution of Eq.(3.6) into Eq.(3.3) will yield a "smoothed" time-history of the population's size (not showing every single saw-tooth) that can be used even when—like in the present case—the drug doses are not evenly spaced, nor constant with respect to the log cell kill factor. If  $K \cdot L(t)$  is less than one the population will increase (exponentially) in size despite the therapy, if it is greater than one the malignant population will become (exponentially) smaller. The latter situation will correspond with RI therapy, the former situation may exist during less severe CMT that eventually results in a relapse as soon as the

Fig. 3.2 Patient CDJ: Experimental datapoints showing population size at various times; open symbols are less certain points (see text, section 3.2.1); start and duration of remission induction therapy is shown (thick arrows); thin arrows indicate additional treatment. BM: bone marrow; PB: peripheral blood



population size exceeds a minimum value (clinical detection level) again (see Fig. 3.1). Of course, deviations from the exponential curves may occur if, for any reason (e.g., spontaneous mutation [Goldie and Coldman, 1979], cell kinetic changes [Norton and Simon, 1986], or—in this case with one PB compartment not very likely—pharmacological changes [Jackson et al., 1983]) drug insensitive (sub)populations start to develop. The constant log cell kill hypothesis is valid for drug sensitive cells only; gradually arising drug resistant cells will lower the achieved log cell kill factor more and more.

*Application to data of patient CDJ.* In the present case no tendency toward Gompertzian growth was observed. Inspection of the datapoints suggested exponential regrowth under CMT, identically for all three cycles, and, at least initially, an exponential decrease in population size during RI therapy. This decrease seems to continue later at a lower rate (Fig. 3.2).

For the reconstruction of the (smoothed) time-history of the population by computer simulation the influence of chemotherapy was split into two parts:

- 1) as at a first glance there were no signs in the observed datapoints that would suggest the formation of therapy resistant subpopulations, a constant influence corresponding with CMT is assumed for the total treatment interval.  $K \cdot L(t)$  in Eq.(3.6) is smaller than one, and  $GF_r \cdot \{1 - K \cdot L(t)\}$  is a positive constant,  $GF_r$ .
- 2) the initial high intensity chemotherapy corresponding with RI therapy—from which the already accounted for influence due to modeling CMT as in 1) is subtracted—is modeled to have an ever decaying influence according to:

$$K \cdot L(t) = H_0 \cdot \exp(-R \cdot t), \quad (3.7)$$

where  $R$  is a measure of the rate of decay from the initial therapy level  $H_0$ . For, at first sight, phenomena like therapy resistance formation may play a role here.

Thus, the resulting growth equation can be written as:

$$C(t + \Delta t) = C(t) + GF_r \cdot \{1 - H_0 \cdot \exp(-R \cdot t)\} \cdot \Delta t. \quad (3.8)$$

The value of  $GF_r$  was estimated from a log-linear regression analysis of the experimental PB datapoints for CMT. The values of  $H_0$  and  $R$  were estimated for the three treatment cycles by fitting the simulated curve of Eq.(3.8) to the observed datapoints, using a standard routine for least squares fitting to a non-linear function with parabolic expansion of chi-square [Bevington, 1969].

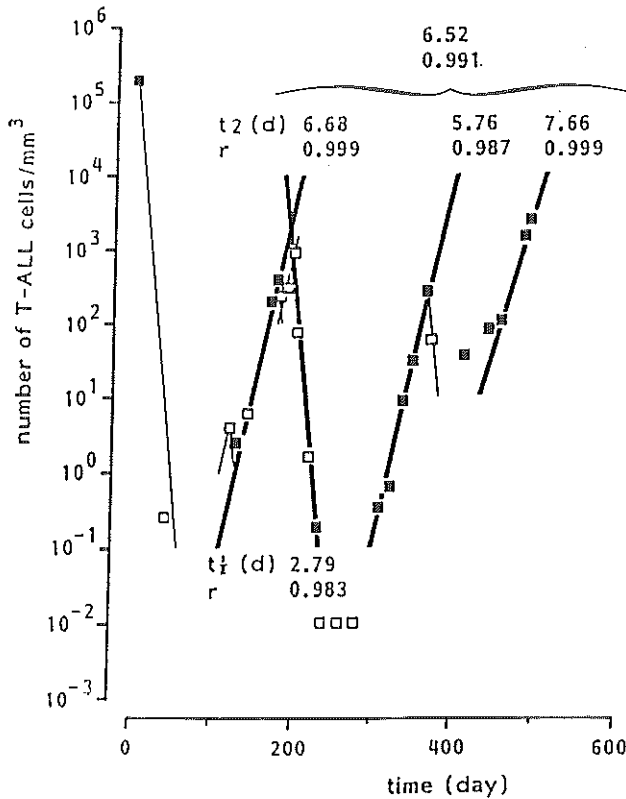
## 3.2 RESULTS AND DISCUSSION

### 3.2.1 PB Data of Patient CDJ

Figure 3.2 shows the *datapoints* observed after immunological marker analysis. It should be noted that for several datapoints the degree of contamination of the MNC population with myeloid cells (percentage  $X$  in Eq.(3.1)) could not be determined. An average value for  $X$  was substituted then, yielding datapoints that are slightly less certain. The fate of the ALL cells in the bone marrow seem to be reflected in the blood. The patterns of PB and BM data look alike, except perhaps at the end of the time interval (day 400). Being the most abundant, the PB data were used for further mathematical analysis.

*Log-linear regression analysis* in the regrowth phases, when influence of RI therapy has virtually vanished and only CMT influence remains, revealed that the population doubling time remained approximately constant,  $T_2 = 6.7, 5.8$  and  $7.7$  d, respectively, for the three consecutive cycles. (N.B., in the first cycle the datapoints between the additional prednisone treatments were taken).

Fig. 3.3 Results of regression analyses performed on the peripheral blood (PB) data-points of patient CDJ: halftime ( $t_{1/2}$ ), doubling times ( $t_2$ ) and correlation coefficients ( $r$ ) are shown. As in Fig. 3.2, open symbols are less certain observations



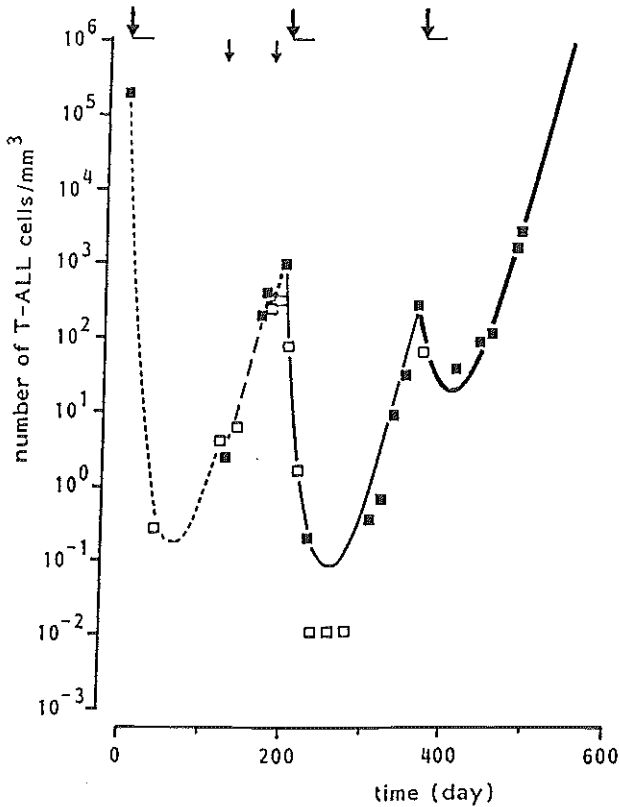
These differences are not significant and after pooling of the data the mean doubling time amounted to 6.5 d (correlation coefficient  $r = 0.991$ ), which corresponds with a constant  $GF_r = \ln 2/T_2 = 0.106 \text{ d}^{-1}$  (Fig. 3.3).

This indicates that the growth kinetic properties of the regrowing population will not have changed during the courses of treatment, and that the population remains in exponential growth phase even under CMT, without any evidence that therapy resistant cells are being formed here.

In the same way an estimate of the halftime was obtained in the declining phase of the growth curve (Fig. 3.3).

With the thus derived estimate for  $GF_r$  the remaining parameters describing the therapy level,  $H_0$  and  $R$  in Eq.(3.8), were found for each treatment cycle by fitting Eq.(3.8) to the corresponding datapoints, starting from  $2 \times 10^5$  T-ALL cells/ $\text{mm}^3$  blood on day zero. The resulting values are listed in Table 3-3. The

Fig. 3.4 Patient CDJ: Computer simulation of the time-history of the population after the start of each treatment (thick arrows: remission induction therapy or thin arrows: additional prednisone treatment), i.e., least squares fit of Eq.(3.8) to the datapoints (open symbols are less certain observations, see text (section 3.2.1)); parameters in Table 3-3



resulting fitted curves representing the time-history of the T-ALL population in PB are shown in Fig. 3.4.

As can be seen in Fig. 3.4 the fitted curves predict a minimum of about 0.1 T-ALL cell per  $\text{mm}^3$  of peripheral blood in the range of day 230-280. The datapoints observed in this time interval—indicated in Fig. 3.4, although not been used for curve fitting purposes—have been arbitrarily set to 0.01, but in this range the actually observed numbers amount to zero. These observations do not contradict the model output. A typical sample size for the immunological marker analysis in this region is 10 ml blood, containing 3000 cells per  $\text{mm}^3$ . According to the model the frequency of the leukemic cells then is  $0.1/3000 = 3.3 \times 10^{-5}$ . This value is within the margin of the stated sensitivity of this detection method, i.e., 1 leukemic cell in  $10^4$  to  $10^5$  normal cells will be detectable

TABLE 3-3 PATIENT CDJ: RESULTS OF CURVE FITTING TO PB DATA

| start of<br>therapy<br>d | $GF_r = 0.106 \text{ d}^{-1}$ |        |                                      |         | $T_2 = 6.52 \text{ d}$ |                                  |      |       |
|--------------------------|-------------------------------|--------|--------------------------------------|---------|------------------------|----------------------------------|------|-------|
|                          | $H_0 \pm \text{sd}$           |        | $R \pm \text{sd}$<br>$\text{d}^{-1}$ |         | AUC<br>d               | LCK (dC/dt)/C<br>$\text{d}^{-1}$ |      |       |
| * 0                      | 6.972                         | 0.264  | 0.0304                               | 0.0013  | 229.34                 | 6                                | ---- | -0.63 |
| 121                      | 1.503                         | 0.005  | 0.1137                               | 0.0003  | 13.22                  | -                                | 0.09 | -0.05 |
| 177                      | 2.856                         | 0.004  | 0.2256                               | 0.0004  | 12.66                  | -                                | 0.99 | -0.19 |
| *198                     | 5.086                         | 0.008  | 0.0277                               | <0.0001 | 183.61                 | 4                                | 0.10 | -0.43 |
| *365                     | 2.386                         | <0.001 | 0.0207                               | <0.0001 | 115.27                 | 1                                | 0.10 | -0.15 |
|                          |                               |        |                                      |         |                        |                                  | B    | A     |

\* = remission induction therapy

$GF_r, T_2$ : growth fraction and doubling time in maintenance therapy phase

$H_0, R$ : initial therapy level and decay rate

AUC: area under therapy level curve LCK: log cell kill

(dC/dt)/C: relative growth rate, just (B)efore or (A)fter start of therapy

[Van Dongen et al., 1985]. Furthermore, a simple calculation using Poisson statistics reveals that, if the chance that a cell is leukemic equals  $3.3 \times 10^{-5}$ , the probabilities of finding zero and of finding one such cell in a sample of 30,000 (10 ml blood) will be equally large, namely 37%.

That the effectivity of the RI part of the therapy is reduced in every new treatment cycle can be deduced from (Table 3-3):

A) the initial therapy level  $H_0$  decreases with the cycle number; the rate of decay  $R$  does not change very much. The halftime of this decay amounts to some 23 d, which is nearly as long as the duration of a RI therapy protocol. N.B., for the prednisone treatments just the reverse is observed, i.e.,  $H_0$  increases to double the value and so does  $R$ . However, this may be an artifact due to too few datapoints in these regions. For instance, for the first prednisone treatment an almost equally good fit could be obtained by assuming  $H_0 = 335.6$  and  $R = 20.86$ , i.e., a very steep decay from a very high initial level. So, to evaluate the most appropriate parameter values more datapoints are necessary;

B) the area under the  $L(t)$  curve, which may serve as a measure of the therapy efficiency and can be calculated by integration of Eq.(3.7) from time zero until infinity (yielding  $H_0/R$ ), goes down with the cycle number. N.B., for the prednisone cases the area under the curve goes down a little as well;

C) by looking at the ratio of the population size at the start of each RI and the minimum population size achieved in that cycle, decreasing overall log cell kill values can be derived;

D) the relative growth rate just after the start of treatment shows ever decreasing values, showing that the decline of the population size is less and less steep. Just before treatment the relative growth rate always shows an approximately constant value, indicating that the state of the growing population is the same each time. Prednisone treatments again show a deviant behavior.

### 3.2.2 PB Data of Other Patients

Figures 3.5 through 3.12 show the datapoints—in general the datapoints were acquired just *before* therapy was applied—simulations and fitted growth curves concerning the development of the T-ALL cell populations in peripheral blood. An overview of the characteristics of these time histories is given in Table 3-4 ('raw' data), Table 3-5 (regression lines), Table 3-6 (simulations) and Table 3-7 (curve fitting).

Three patients started with equally high leukemic cell loads at diagnosis ( $\approx 1.4 \times 10^8/\text{mm}^3$ ). The other patients had a lower (up to 3 logs) initial leukemic cell load (Fig. 3.13). First remission (tumor load  $< 10^5$  cells/ $\text{mm}^3$ ) was successfully induced in all patients, within 1 to 3 weeks. During at least some time the leukemic cell load dropped below detectable levels, which is here about 3 leukemic cells per  $\text{mm}^3$  blood. (N.B., assuming 5 l peripheral blood per patient, the theoretical minimal malignant cell frequency that can lead to relapse is  $1/5000 = 2 \times 10^{-4}$ ). Three patients relapsed (MZ, IA, CDJ), two of which were given a second RI therapy (IA, CDJ). Their malignant cell loads then were 1.5-2.5 logs below the levels at first diagnosis. Both relapsed again and for one patient (CDJ) third RI therapy was started, leukemic cell load now being 3 logs below that at first diagnosis.

For patient ER three distinct phases during first RI treatment can be distinguished, with starting malignant cell loads at two and three logs, respectively, below that of diagnosis. Patient KG shows two distinct phase during first RI, the second phase starting with a leukemic cell load of 1.5 logs below that of diagnosis.

The rates at which the patients went into remission are not very different; the population halftime is about 2 days (varies "interpatient" between 0.5 and 3 days). In this respect, per patient, there is hardly any difference between first or second RI period, nor—where applicable—per phase of the first RI period. More, but certainly no spectacular, variation is seen in the doubling time of the regrowing cell populations (Fig. 3.13). Thus, the time histories of the T-ALL cell populations in PB are characterized by stages of exponential decline and exponential regrowth.

In a few patients (PM, ER, KG) the influence of the daily drug doses could be simulated by letting them cause a constant LCK/dose. By choosing a constant, "average", population doubling time of 6.52 d (3.25 d in one case), and



Fig. 3.5 T-ALL cell population in PB of patient MZ. Datapoints are shown with estimated accuracies; values below  $0.01 \text{ ml}^{-1}$  have been drawn on the 0.01 level. Halftimes ( $T_h$  [d]) and/or doubling times ( $T_d$  [d]) between certain datapoints are indicated. The best-fitted curve has an unrealistic minimum of  $2.06 \times 10^{-14}$  T-ALL cells/ $\text{mm}^3$  on day 275

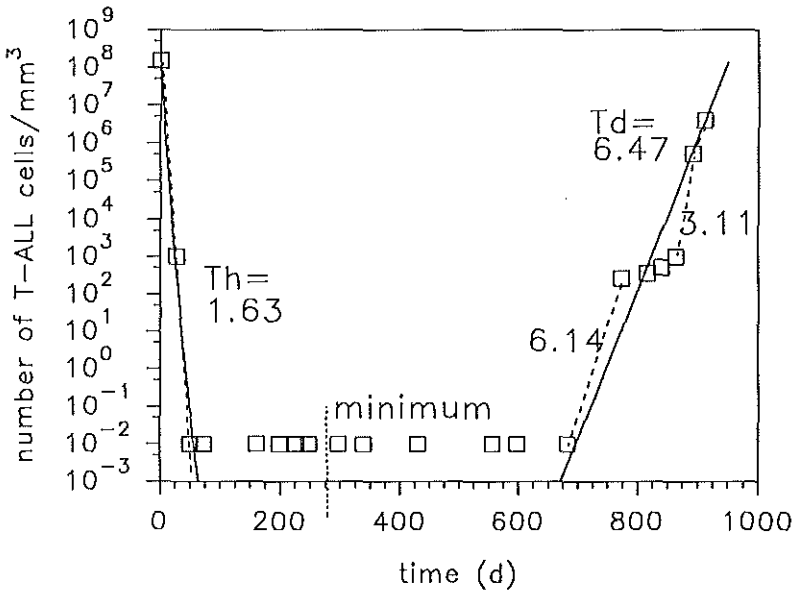


Fig. 3.6 T-ALL cell population in PB of patient IA. Best-fitted curves are drawn for both remission (induction) periods. See also legend of Fig. 3.5

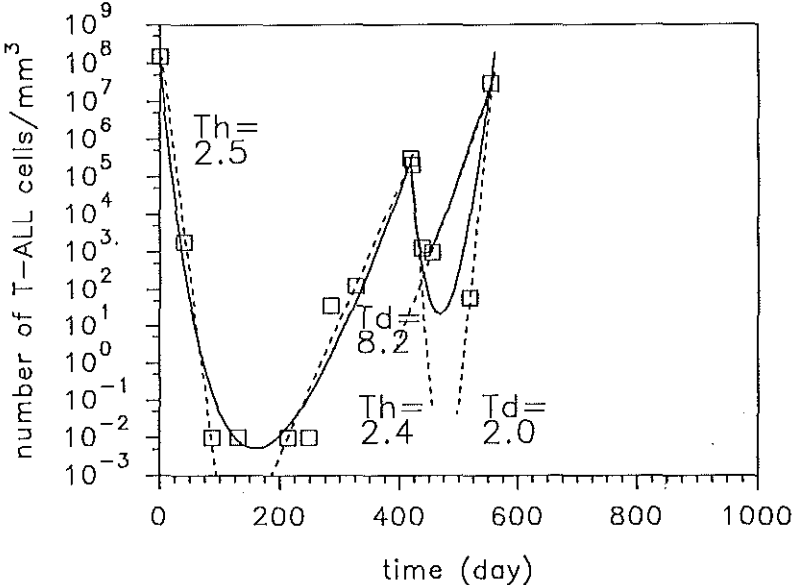


Fig. 3.7 T-ALL cell population in PB of patient RS. Large fluctuations in population size are being 'smoothed' by the fitted curve. See also legend of Fig. 3.5

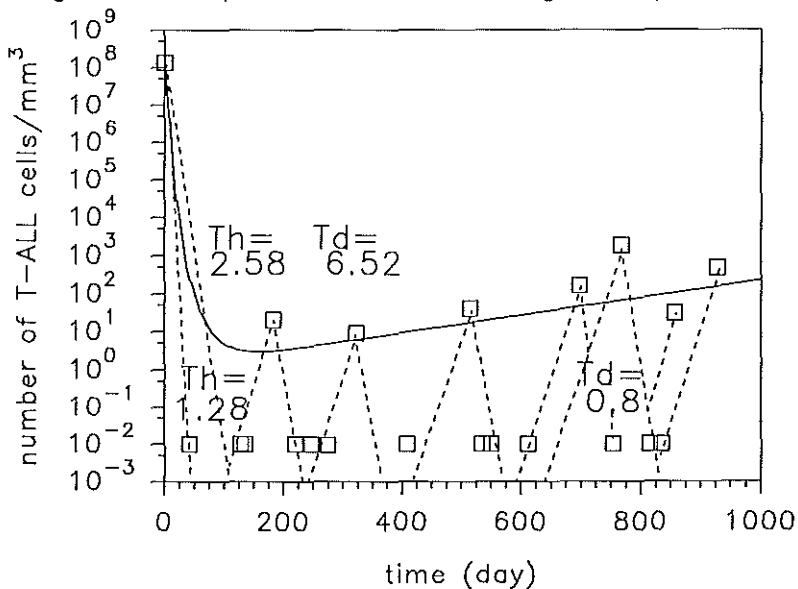


Fig. 3.8 T-ALL cell population in PB of patient PM. The fitted curve has an unrealistic minimum of  $2.4 \times 10^{-33}$  T-ALL cells/mm<sup>3</sup> on day 314 and predicts the return to the  $10^5$  cells level to happen at day 1232. See also legend of Fig. 3.5

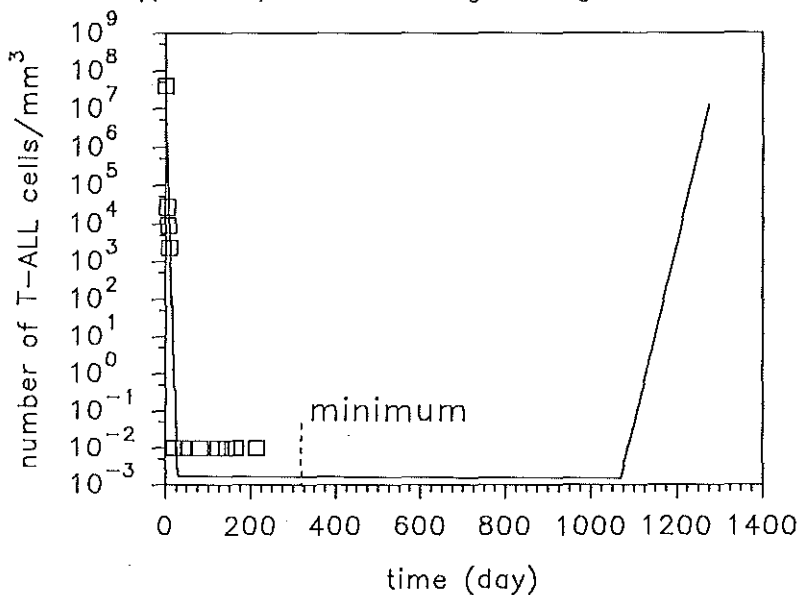


Fig. 3.9 T-ALL cell population in PB of patient PM, focussed on the first 100 days. Fitted curve A predicts eradication of the malignancy (but is physically impossible as it implies ever increasing therapy influence), curve B predicts a too early return to the  $10^5$  cells level at day 100. Curve C represents an exponential decrease. All curves yield equal goodness of fit values (tcc). The simulation (D), assuming that each daily drug dose causes 0.57 LCK and regrowth doubling time is 6.52 d, follows the datapoints rather well. See also legend of Fig. 3.5

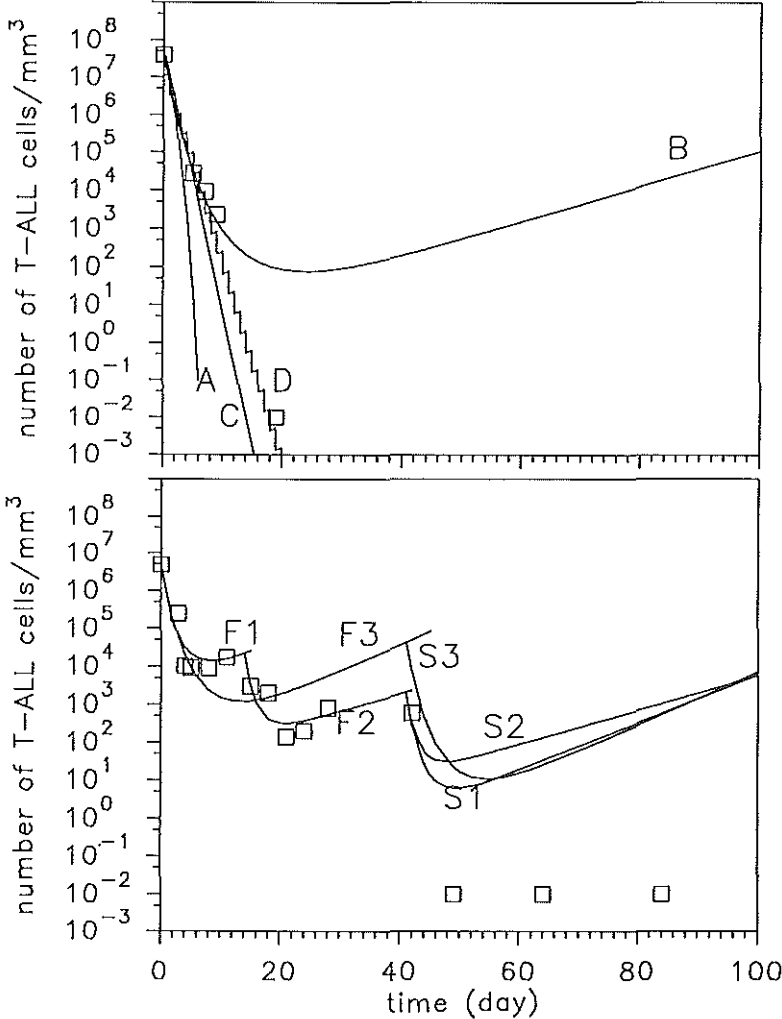


Fig. 3.10 T-ALL cell population in PB of patient ER during the first 100 days. The period is divided in two parts: before and after day 40. Points before day 40 were fitted (Fit 03) and the same curve was used after day 40 (Sim 03). Splitting the first period once more, at day 14, both parts were fitted (Fit 01 and Fit 02, respectively) and either curve was used to simulated the period after day 40 (Sim 01, respectively, Sim 02). All curves predict about  $10^4$  T-ALL cells per ml on day 100

Fig. 3.11 T-ALL cell population in PB of patient ER during the first 100 days. A simulation with daily drug doses on days 0-6, 14-20 and 42+ causing 0.47 LCK each. Regrowth doubling time is set to 3.25 d (instead of 6.52 d)

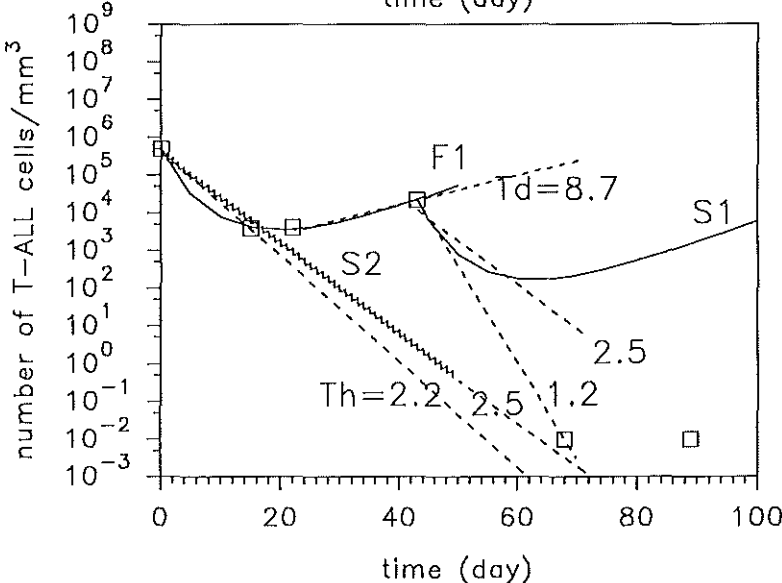
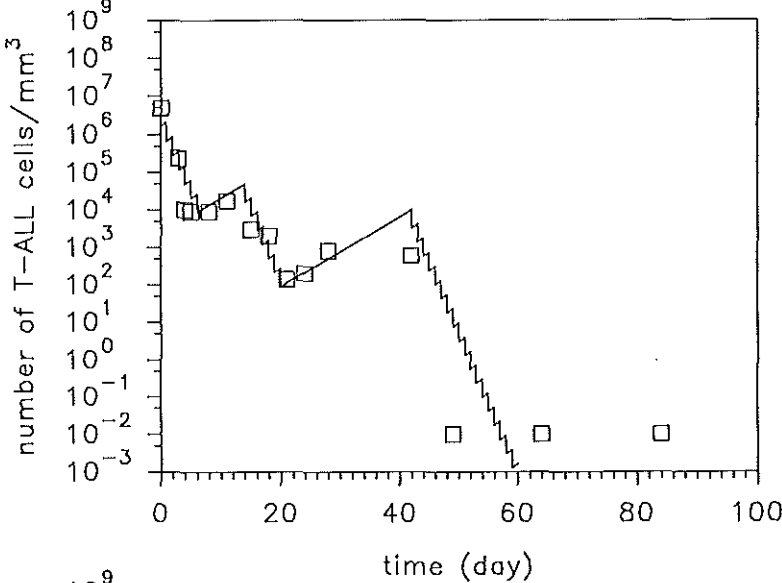


Fig. 3.12 T-ALL cell population in PB of patient KG during the first 100 days. Observations may be explained with a simulation with daily drug doses causing 0.18 LCK each, discontinued between days 21 and 42, while regrowth doubling time is 6.52 d. Observations up to day 42 were well fitted; using the same curve after day 42 predicts relapse ( $10^5$  cells) much too soon, around day 122

TABLE 3-4 OVERVIEW OF ANALYSIS RESULTS 'RAW DATA'

| patient              | MZ      | IA                 | RS      | PM      | ER      | KG      | CDJ  |
|----------------------|---------|--------------------|---------|---------|---------|---------|------|
| <i>initially</i>     |         |                    |         |         |         |         |      |
| $T_0(d)$             | 0       | 0                  | 0       | 0       | 0       | 0       | 0    |
| $C_0 \times 10^{-6}$ | 148.5   | 145.7              | 133.2   | 39.5    | 4.8     | 0.49    | 0.20 |
| <i>R1</i>            |         |                    |         |         |         |         |      |
| $T_{1/2}(d)$         | 1.6     | 2.5                | 1.3-2.6 | 0.5-1.3 | 0.5-1.4 | 1.2-2.2 | 2.8  |
| $T_2(d)$             | 3.1-6.5 | 8.2                | 0.8-6.5 | ?       | 2.0-5.8 | <8.7?   | 6.7  |
| $T_{e5}(d)$          | 857-867 | 377                | -       | -       | -       | -       | 260  |
| <i>R2</i>            |         |                    |         |         |         |         |      |
| $C_0$                |         | $3.08 \times 10^5$ |         |         |         |         | 914  |
| $T_0(d)$             |         | 418                |         |         |         |         | 198  |
| $T_{1/2}(d)$         |         | 2.4                |         |         |         |         | 2.8  |
| $T_2(d)$             |         | 2.0?               |         |         |         |         | 5.8  |
| $T_{e5}(d)$          |         | 113                |         |         |         |         | 247  |
| <i>R3</i>            |         |                    |         |         |         |         |      |
| $C_0$                |         |                    |         |         |         |         | 257  |
| $T_0(d)$             |         |                    |         |         |         |         | 365  |
| $T_{1/2}(d)$         |         |                    |         |         |         |         | 2.8  |
| $T_2(d)$             |         |                    |         |         |         |         | 7.7  |
| $T_{e5}(d)$          |         |                    |         |         |         |         | 187  |

- $R_i$  =  $i^{\text{th}}$  remission (induction) period  
 $T_0$  = time of starting therapy;  $C_0$  = number of cells per ml Peripheral Blood at start of therapy  
 $T_{1/2}$  = estimated halftime of cell population;  $T_2$  = estimated doubling time of cell population  
 $T_{e5}$  = estimated time of reaching the  $10^5$  cells level

TABLE 3-5 OVERVIEW OF ANALYSIS RESULTS 'LOG-LINEAR REGRESSIONS'

| patient              | MZ | IA | RS                   | PM    | ER    | KG    | CDJ   |
|----------------------|----|----|----------------------|-------|-------|-------|-------|
| <i>R1</i>            |    |    |                      |       |       |       |       |
| A                    |    |    | 1.62                 | 7.35  | 6.75  | 6.43  | 12.78 |
| B(d <sup>-1</sup> )  |    |    | $4.6 \times 10^{-3}$ | -0.48 | -0.58 | -0.19 | -0.23 |
| r                    |    |    | 0.68                 | -0.10 | -0.97 | -0.96 | -0.95 |
| $T_{1/2}(d)$         |    |    | -                    | 0.705 | 0.518 | 1.56  | 2.96  |
| $T_2(d)$             |    |    | 150.3                | -     | -     | -     | -     |
| for time interval(d) |    |    | 40-920               | 0-9   | 0-5   | 11-21 | 0-21  |

regression line:  $\ln(C) = A + B \cdot t$ , time t in days; r: correlation coefficient

TABLE 3-6 OVERVIEW OF ANALYSIS RESULTS 'SIMULATIONS'

| patient                 | MZ | IA | RS | PM   | ER                    | KG   | CDJ          |
|-------------------------|----|----|----|------|-----------------------|------|--------------|
| simulation              |    |    |    | D    | C                     | A    | B            |
| R1 treatm. duration (d) |    |    |    | 0-21 | 0-6/<br>14-20/<br>42+ | 0+   | 0-21/<br>43+ |
| LCK/dose                |    |    |    | 0.57 | 0.48                  | 0.17 | 0.17         |
| T <sub>1/2</sub> (d)    |    |    |    | 0.58 | 0.78                  | 2.5  | 2.5          |
| T <sub>2</sub> (d)      |    |    |    | 6.52 | 3.25                  | 6.52 | 6.52         |

- R1: first remission (induction) period  
 LCK/dose: calculated Log Cell Kill per daily drug dose  
 T<sub>1/2</sub>: observed value of the population halftime  
 T<sub>2</sub>: set value of the population doubling time  
 Simulation: the population grows exponentially with doubling time T<sub>2</sub>; a drug dose reduces the population instantaneously by a factor of 10<sup>LCK</sup>; this yields a net exponential decrease in cell numbers with halftime T<sub>1/2</sub>

taking the rate of decline (halftime) into consideration, the log cell kill per dose needed to "match" the datapoints varies between 0.2 and 0.6 for different patients (Table 3-6).

Another way of looking at the influence of therapy is by fitting a growth curve (Eq.(3.8)) to the datapoints. This allows estimating the variables H<sub>0</sub> and R, respectively, a measure of the initial impact of the therapy and the rate at which this impact decays. A low H<sub>0</sub> value means a low initial therapy impact, a low R means that the initial impact decays but slowly. Both values low, thus, means a small but long lasting effect. Overall effect can be evaluated by looking at the area under the curve (AUC), which is the integral of H<sub>0</sub> · exp(-R · t). Values of H<sub>0</sub>, R and AUC are shown in Fig. 3.14, per patient, for the best fitting curve. Large variations occur both in H<sub>0</sub> and in R. A factor of about 85 between highest and lowest AUC is seen. When looking at subsequent RI therapies in a same patient, it is noticed that the AUC always decreases, although only a little bit sometimes. By using the AUC, the clinical notion that "A same therapy given later is less effective" can thus be quantified.

### 3.3 CONCLUDING REMARKS

In conclusion, from this pilot study it can be deduced for *patient CDJ* that:

A) assuming a constant influence of low intensity maintenance therapy the otherwise "unperturbed" population in this T-ALL patient grows exponentially

TABLE 3-7 OVERVIEW OF ANALYSIS RESULTS 'CURVE FITTING'

| patient              | MZ       | IA                  | RS        | PM                 | ER                  | KG                 | CDJ    |      |      |
|----------------------|----------|---------------------|-----------|--------------------|---------------------|--------------------|--------|------|------|
| <i>initially</i>     |          |                     |           |                    |                     |                    |        |      |      |
| $T_0(d)$             | 0        | 0                   | 0         | 0                  | 0                   | 0                  | 0      |      |      |
| $C_0 \times 10^{-6}$ | 148.5    | 145.7               | 133.2     | 39.5               | 4.8                 | 0.49               | 0.20   |      |      |
| curve fit            | <u>A</u> | <u>01</u>           | <u>01</u> | <u>C</u> <u>0A</u> | <u>01</u> <u>03</u> | <u>01</u>          |        |      |      |
| $R1$ start(d)        | 0        | 0                   | 0         | 0                  | 0                   | 0                  | 0      |      |      |
| $H_0$                | 6.53     | 4.80                | 105.38    | 14.52              | 9                   | 18.0               | 513.81 | 6.42 | 6.97 |
| $R(d^{-1})$          | 0.007    | 0.01                | 0.03      | 0                  | .007                | 0.35               | 0.2    | 0.1  | 0.03 |
| $H_0/R = AUC$        | 933      | 486                 | 4215      | -                  | 1286                | 52                 | 70     | 67   | 229  |
| $T_{1/2}(d)$         | -        | -                   | -         | 0.48               | 0                   | -                  | -      | -    | -    |
| $T_2(d)$             | 7.34     | 6.52                | 130.89    | 6.52               | 6.2                 | 4.83               | 4.35   | 5.27 | 6.52 |
| tcc                  | >.99     | >.99                | .04       | >.99               | >.99                | >.99               | .99    | >.99 |      |
| $R2$ start(d)        |          | <u>02</u> <u>03</u> |           |                    | <u>02</u>           |                    |        |      |      |
| C                    |          | 418 418             |           |                    | 14                  | 43                 | 198    |      |      |
| $H_0$                |          | $3.08 \times 10^5$  |           |                    | $2.21 \times 10^4$  | $2.29 \times 10^4$ | 914    |      |      |
| $R(d^{-1})$          |          | 7.43 2.32           |           |                    | 24.27               | 6.42               | 5.09   |      |      |
| $H_0/R = AUC$        |          | 0.08 0.02           |           |                    | 0.49                | 0.1                | 0.03   |      |      |
| $T_{1/2}(d)$         |          | 93 135              |           |                    | 50                  | 67                 | 184    |      |      |
| $T_2(d)$             |          | -                   |           |                    | -                   | -                  | -      |      |      |
| tcc                  |          | 6.52 2.02           |           |                    | 6.71                | 5.27               | 6.52   |      |      |
|                      |          | >.99 >.99           |           |                    | .89                 | >.99               |        |      |      |
| $R3$ start(d)        |          |                     |           |                    | 41                  | 41                 | 365    |      |      |
| C                    |          |                     |           |                    | $2.12 \times 10^3$  |                    | 257    |      |      |
| $H_0$                |          |                     |           |                    | 18.05               | 24.27              | 2.39   |      |      |
| $R(d^{-1})$          |          |                     |           |                    | 0.35                | 0.49               | 0.02   |      |      |
| $H_0/R = AUC$        |          |                     |           |                    | 52                  | 50                 | 115    |      |      |
| $T_{1/2}(d)$         |          |                     |           |                    | -                   | -                  | -      |      |      |
| $T_2(d)$             |          |                     |           |                    | 4.83                | 6.71               | 6.52   |      |      |
| tcc                  |          |                     |           |                    |                     |                    |        |      |      |

$R_i$  =  $i^{th}$  remission (induction) period;  $T_0$  = time of starting therapy;  $C_0$  = number of cells per ml Peripheral Blood at start of therapy;  $T_{1/2}$  = estimated halftime of cell population;  $T_2$  = estimated doubling time of cell population;  $H_0, R$ : initial therapy level and decay rate; fitted curve: Eq.(3.8); total correlation coefficient,  $tcc = \{1 - \Sigma (C_{fit} - C_{obs})^2 / \Sigma C_{obs}^2\}^{1/2}$  is a measure of goodness of fit (tends to 1 for perfect fit)

with a mean doubling time of 6.5 d, no matter the preceding treatment history; B) intensive remission induction therapy can be modeled with an exponentially decaying therapy level  $L(t)$  that, in a series of equal treatments, depends on the sequence number. As it increases the efficiency of the remission induction therapy decreases, which is apparent from the facts that 1)  $L(t)$  shows a less steep decay from a lower initial value; 2) the area under the curve decreases; and 3) the induced overall log cell kill decreases. This indicates the development of therapy resistance.

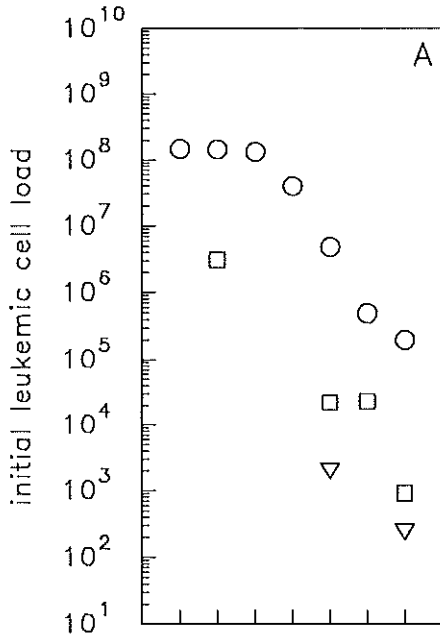
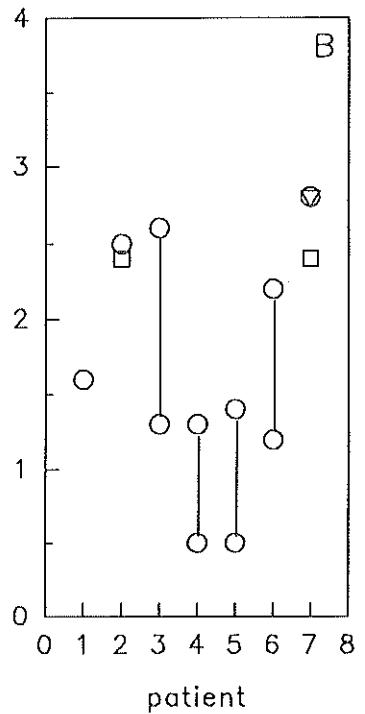
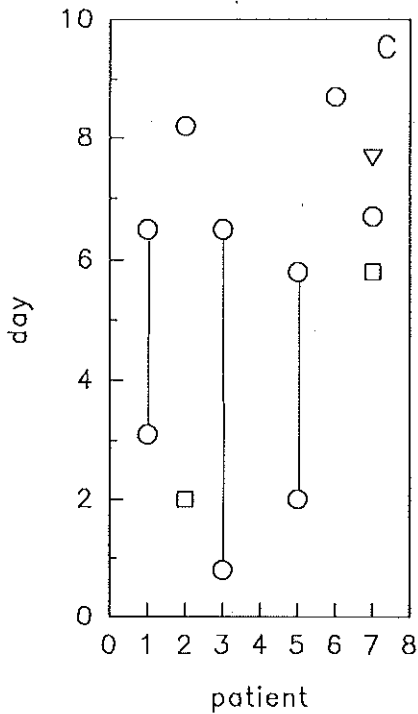


Fig. 3.13 A) Malignant cell load at start of therapy for each of the seven patients:  $\circ$  1st RI,  $\square$  2nd RI or 2nd phase of 1st RI,  $\nabla$  3d RI or 3d phase of 1st RI; B) halftimes of T-ALL cell disappearance, estimated from 'raw' data; C) doubling times of T-ALL cell population, estimated from 'raw' data





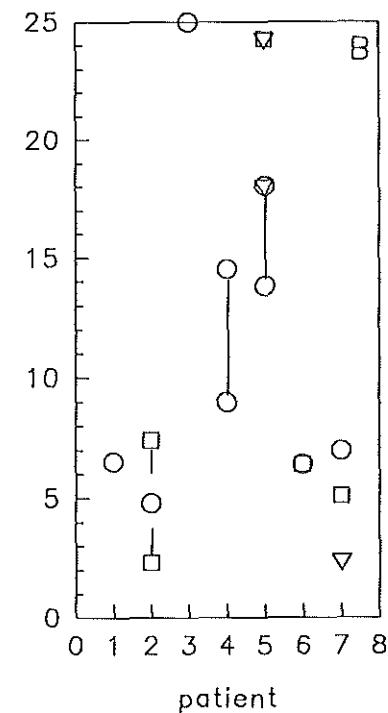
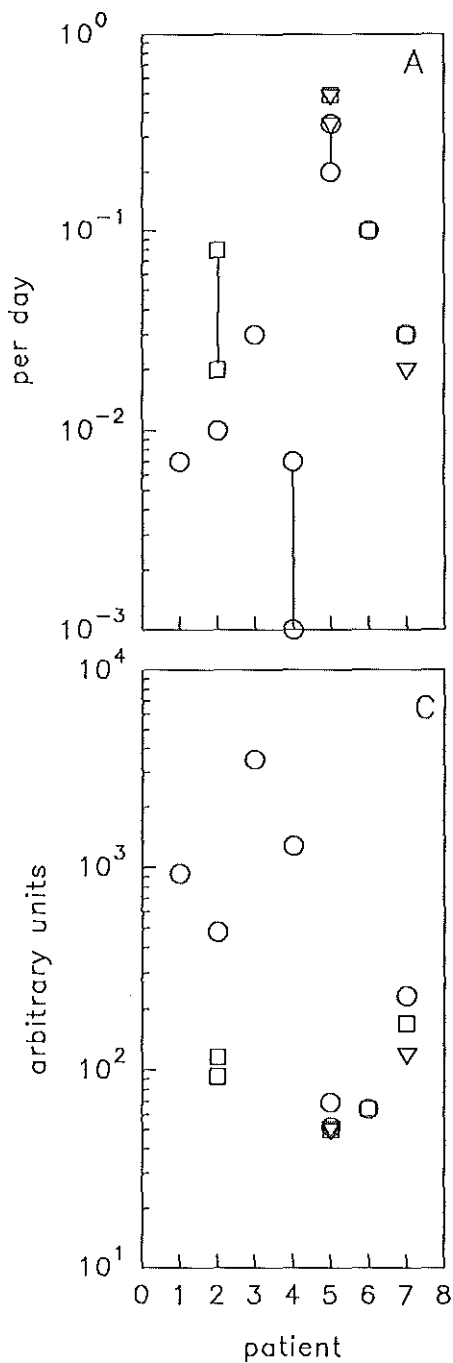


Fig. 3.14 Therapy influence measure, obtained from best fits of Eq.(3.8) to PB datapoints for each of the seven patients:  $\circ$  1st RI,  $\square$  2nd RI or 2nd phase of 1st RI,  $\nabla$  3d RI or 3d phase of 1st RI; A) value of R (rate of decay of initial therapy influence parameter  $H_0$ ); B) value of  $H_0$ ; C) area under the curve (AUC), i.e., time integral of  $K \cdot L(t) = H_0 \cdot \exp(-R \cdot t)$

For *the other patients* much fewer datapoints are available. Some of them have not yet been followed long enough. Although it seems that the above conclusions can—to some extent—be generalized to apply to most, if not all patients, variation in some parameter values is large between individuals.

In a study that involved a large number (158) of newly diagnosed ALL patients [Rautonen et al., 1988] it was found that the rate at which blast cells disappear from the peripheral blood is a predictive factor for the patient's response to therapy. Slow disappearance ( $> 10$  d) indicates a poor prognosis. According to this blast cell clearance rate criterion the seven patients would be in the promising prognosis category. With so few patients in the present study, however, it is not useful to try and find other categories—patients who have their values of certain parameters between this and that boundary—and determine some parameter profile upon which a sound prognosis for therapy outcome can be based.

Still, it can be said that the present results show that mathematical analysis of data obtained by a sensitive method for the detection of low numbers of malignant cells yields valuable information about the growth characteristics of the cell population, as well as about its sensitivity to chemotherapy. The mathematical analysis suggests further experimental investigations on the emergence of a drug resistant subpopulation. It would be interesting to make *in vitro* cultures of T-ALL cells [Lange, 1989] taken from the population at various instants during the treatments and see whether A) the unperturbed growth kinetics (doubling time) indeed remain the same, and B) the response to administration of the various drugs changes. Another approach would be to look for a possible amplification of multi-drug resistance (MDR) genes increasing with time. This would be expressed as an ever increasing amount of MDR gene products, to be detected by immunocytochemical assays. For instance, Ma et al. [1987] in this way showed a correlation between the amount of P-170 plasma membrane glycoproteins and the presence of MDR cells in two cases of acute non-lymphoblastic leukemia.

### 3.4 REFERENCES

- Bevington PhR (1969) Data reduction and error analysis. McGraw-Hill, New York  
Birkhead BG and Gregory WM (1984) A mathematical model of the effects of drug resistance in cancer chemotherapy. *Math Biosci*, 72:59-69  
Goldie JH and Coldman AJ (1979) A mathematical model for relating the drug sensitivity of tumors to their spontaneous mutation rate. *Cancer Treatment Rep*, 63:1727-1733  
Hagenbeek A and Martens ACM (1985) Detection of minimal residual disease in acute leukemia:

- possibilities and limitations. *Eur J Cancer Clin Oncol*, 21:389-395
- Jackson H, Bock M, Jackson NC and Lendon M (1983) The testis: a protected environment for leukemic cells against cyclophosphamide in a mouse model. *Cancer Chemother Pharmacol*, 11:200-202
- Lange BJ (1989) Growth of human leukemia cells in vitro. In: Baserga R (ed.), *Cell growth and division, a practical approach*. IRL Press, Oxford, 61-80
- Ma DDF, Davey RA, Harman DH, Isbister JP, Surr RD, Mackertich SM, Dowden G and Bell DR (1987) Detection of a multidrug resistant phenotype in acute non-lymphoblastic leukemia. *Lancet*, i:135-137
- Norton L and Simon R (1977) Tumor size, sensitivity to chemotherapy and design of treatment schedules. *Cancer Treatment Rep*, 60:1307-1317
- Norton L and Simon R (1986) The Norton-Simon hypothesis revisited. *Cancer Treatment Rep*, 70:163-169
- Rautonen J, Hovi L, Sümes MA (1988) Slow disappearance of peripheral blast cells: an independent risk factor indicating poor prognosis in children with acute lymphoblastic leukemia. *Blood*, 71:989-991
- Schultz FW, Martens ACM and Hagenbeek A (1986) A mathematical model for leukemia (re-)growth in the rat. In: Eisenfeld J and Witten M (eds.), *IMACS Transactions on Scientific Computing 1985, 5, Modelling of Biomedical Systems*. North Holland, Amsterdam, 41-49
- Skipper HE, Schabel FM and Wilcox WS (1964) Experimental evaluation of potential anticancer agents XII: on the criteria and kinetics associated with curability of experimental leukemia. *Cancer Chemother Rep*, 35:1-111
- Smith RG, Hetherington ML, Huntsman PR and Buchanan GR (1986) Surveillance of terminal deoxynucleotyl transferase-positive cells in peripheral blood of patients with acute lymphoblastic leukemia. In: Hagenbeek A and Löwenberg B (eds.), *Minimal residual disease in acute leukemia 1986*. Martinus Nijhoff Publishers, Dordrecht, 134-140
- Steel GG (1977) *Growth kinetics of tumours*. Clarendon Press, Oxford
- Van Bekkum DW and Hagenbeek A (1977) Relevance of the BN leukemia as a model for human acute myeloid leukemia. *Blood Cells*, 3:565-579
- Van Dongen JJM, Hooijkaas H, Comans-Bitter M, Hählen K, De Klein A, Van Zanen GE, Van't Veer MB, Abels J and Benner R (1985) Human bone marrow cells positive for terminal deoxynucleotidyl transferase (TdT), HLA-DR, and a T cell marker may represent prothymocytes. *J Immunol*, 135:3144-3150
- Van Dongen JJM, Hooijkaas H, Adriaansen H, Hählen K and Van Zanen GE (1986) Detection of minimal residual acute lymphoblastic leukemia by immunological marker analysis: possibilities and limitations. In: Hagenbeek A and Löwenberg B (eds.), *Minimal residual disease in acute leukemia 1986*. Martinus Nijhoff Publishers, Dordrecht, 113-133



## Chapter 4

### Pharmacokinetics<sup>1</sup>

#### [Identification of the Dynamic System of *in vivo* Distribution and Metabolism of the Cytostatic Drug Daunomycin in the Brown Norway Rat]

Daunomycin (DAU) is an anthracycline antibiotic that is commonly used for chemotherapy of various types of cancer [Arcamone, 1981; Gottlieb et al., 1983]. For most anthracyclines the clinically applied doses not only reduce the neoplastic burden, they also cause adverse effects on normal tissues. For example, cases of cumulative cardiotoxicity [Blum and Carter, 1974; Villani et al., 1985] and severe myelosuppression [O'Bryan et al., 1977] have been frequently reported and often these adverse effects are dose-limiting. Therefore, with respect to the administration of DAU an optimization problem arises. Those administration regimens should be chosen that guarantee a maximum of chemotherapeutic effectiveness, but also a minimum risk of unwanted side effects on healthy tissues. In other words, sufficiently high DAU concentration must be present for sufficiently long time at the site of the tumor, while exposure of healthy tissues should be kept as small as possible. What should be considered sufficient in this respect is a subject of another study, namely, on the response of the various cells and tissues to a certain concentration of the drug (pharmacodynamics [Testa, 1987]; this may include 'counteraction measures' by the disturbed tissue system, i.e., development of drug tolerance [Peper et al., 1987]). Only from the combination of pharmacokinetics and pharmacodynamics will be learned what therapeutic and/or toxic effects can be expected.

Thus, before the problem of optimal drug administration can be dealt with, one necessary thing to be obtained is detailed information on the pharmacokinetics of the drug. It should be predictable with good accuracy, what the concentration—time courses will be in the organs of interest for a given input of the drug. Often it is not only the administered (or parent) drug that plays a role, but frequently the formation of metabolites and their distribution kinetics must be considered as well [Carson et al., 1981].

Drug concentrations in most organs cannot be measured directly without destructive methods. Monitoring plasma levels alone, which would cause little trouble in the clinic, unfortunately does not yield sufficient information [McVie, 1984; Van Rossum et al., 1988] without the pharmacokinetic characteristics of the drug being known. In other words, the dynamic behavior of the system of drug distribution and elimination in response to some drug administration

---

<sup>1</sup>Parts of Chapter 4 have been published before in various conference proceedings [Schultz et al., 1985, 1987, 1988a, 1988b]

schedule must be explored. To this purpose the application of general systems theory and mathematical modeling methods, as developed in the technical sciences, may serve as a convenient tool [e.g., Schmidt, 1982]. In particular, the system identification 'recipe' for matching a theoretical—descriptive—model to an actual physical system, distinguishes three phases. Phase I consists of drafting a plausible mathematical model of the system, i.e., the description of the system's processes by mathematical formulae. All *à priori* knowledge on and experience with the system's properties should be considered. Then, in phase II, parameter estimation, the chosen model is optimized by finding specific values for the parameters in the formulae. By substituting certain parameter values in the formulae a model response is calculated. The model response is compared to the actually observed system behavior. The parameter values then are adjusted, repeatedly if necessary, until the calculated model response corresponds best with the observations. In the last phase, III, a decision must be made whether the optimized model yields sufficient resemblance to the system. If deviations between the system's behavior and the optimized response of the current model remain too large, a more appropriate model must be found. Perhaps a minor refinement or may be an extensive adjustment will be necessary. This means going back to phase I.

To illustrate this with an example, suppose that it has been observed that every quarter of an hour the concentration of some compound is half what it was the last time; say 20 at time zero, 10 after 15 min and 5 after 30 min. As a model for this decrease linear decay may be assumed. This means a description of the concentration—time history,  $C$  as function of  $t$ , by a straight line. In general, the class of straight line models is characterized by the formula:  $C = a \cdot t + b$ , where  $a$  and  $b$  are the parameters (slope and intercept). Here, the optimum values of  $a = -0.5$  and  $b = 19.2$  are found, yielding the smallest (sum of squared) deviations between the line and the observations. Still, deviations are present and, from experience, it is known that exponential decay is a better model for this  $C-t$  relationship. This class of models is described by  $C = C_0 \cdot \exp(-b \cdot t)$ , where parameter  $C_0$  is the initial concentration and parameter  $b$  a measure of the decay rate. Indeed, by optimizing,  $C_0 = 20$  and  $b = 0.046$  make deviations between the model and the observations disappear. Thus, the last model is superior to the first.

Once the pharmacokinetic processes have been identified, the subsequent ability to predict the system behavior in response to arbitrary inputs (drug administration regimens) will improve controlled chemotherapy, because of the possibility to optimize the drug input for desired effects [Swan, 1985].

The present chapter will be about the system identification of DAU pharmacokinetics in the rat. Identification means finding the system's behavior from input/output measurements (drug administration regimen/drug concentrations as function of time and location). As indicated above, for reasons of

observability it will be necessary to make use of *in vivo* experiments with laboratory animals. For ethical and economical reasons the number of animals used must be kept as low as possible. Therefore, the quantity of data minimally needed to produce reliable results should be established. Several computer simulation studies to serve this purpose are reported below.

The main purpose of present study is to look into the fate of DAU in the Brown Norway (BN) rat. Although chemotherapy is usually applied to sick people, and in the BN rat a myelocytic leukemia (BNML) can be induced to mimic human disease, for the time being only healthy rats are considered. The influence of the presence of a tumor load may be incorporated at later (future) stages.

DAU was administered quickly as a single intravenous dose, which can be considered a pulse input. *In vivo* concentration—time datapoints in many organs could be obtained by sacrificing the animals after certain time intervals. The relation between these in- and output data was analyzed by means of mathematical multicompartement models, based on a model previously used to describe adriamycin pharmacokinetics [Sonneveld and Mulder, 1981; Sonneveld, 1980]. New in the present model is that, under the constraint of first order kinetics to describe the *in vivo* distribution and elimination processes, metabolism is allowed at either a concentration dependent or a constant rate (as DAU metabolizes to the compound daunomycinol (DOL)). Furthermore, various model structures were assumed, i.e., several anatomically and physiologically possible pathways for drug transport via plasma to and from the different organs—that are represented by the compartments—were considered. These models were tested for their adequacy in explaining the actual observed concentration—time datapoints, using numerical optimization algorithms to estimate the corresponding transfer rate constants.

This modeling approach further differs from the work by most other investigators in that physiological meaningful models are used, i.e., the compartments and their interconnections represent actual anatomically well defined regions. Furthermore, the models are large (up to 11 organ or tissue regions, each containing two compartments for the compounds DAU and DOL, respectively), allowing much detail in the drug distribution patterns. In most pharmacokinetic studies non-physiological models with only two [Johansen et al., 1984] or three compartments [Eksborg et al., 1986], representing plasma and tissues (with fast and slow drug exchange), are considered because of the ease with which the corresponding model equations can be solved analytically [Wagner, 1975], as well as their extensively investigated properties [Garrett, 1980; Godfrey, 1983]. They also can yield information that is useful for another purpose [Metzler, 1971], e.g., even the comparison of routes of administration [Collins et al., 1980], but of course they do not provide much vital, detailed information. Finally, in contrast to other physiological models, e.g., flowlimited

models [Bischoff and Dedrick, 1968; Chan et al., 1978; Molino et al., 1986], no *à priori* values for physiological variables other than compartment volumes need to be supplied. Such variables (e.g., blood flow velocities, membrane characteristics) are often difficult to assess experimentally and in general show large interindividual fluctuations. Such variability is here considered implicitly in the transfer rates estimated as model parameters. These are considered as random variables, having a distribution with a mean value and a variance, rather than as unknown constants.

Before elaborating on the large-scale models for DAU-pharmacokinetics, first the tools used in mathematical modeling are explained below. As further preparatory work, different approaches to existing (classical) solution methods are elucidated. They were tested for their performance with respect to speed of computation and accuracy of results. This enabled the selection of the best performing optimization method. It should be kept in mind that without the availability of efficient algorithms (i.e., fast, accurate, economical in occupation of computer memory space) the identification of large systems—involving the estimation of many parameters, the simultaneous solution of many differential equations, the manifold calculation of concentration—time histories—becomes a tedious affair. Such efficiency requirements also are important, if eventually—as desired—similar identification and (especially, drug administration) optimization procedures are to be transferred to computer systems for clinical use.

## 4.1 MATERIALS AND METHODS

### 4.1.1 Experimental Data Acquisition

To provide the necessary input/output data, *in vivo* experiments were conducted as described by Nooter et al. [1986]. DAU (supplied by Farmitalia, Milano, Italy) was dissolved in 0.5 ml saline and administered, under light aether anesthesia, *i.v.* as a rapid bolus injection into a tail vein of 12-week-old female Brown Norway rats weighing 165 g (Rijswijk inbred strain). The dose was 7.5 mg/kg, which is comparable to the clinically accepted dose of 40 mg/m<sup>2</sup> in man [Freireich et al., 1966]. The rats had free access to food and water. Urine was collected by a non-invasive method. After certain time intervals groups of 4-5 rats were sacrificed by cervical dislocation under aether anesthesia. Plasma was obtained from aortic blood samples, EDTA was added to prevent coagulation. Organs of interest were removed and frozen with liquid nitrogen. The material was stored at -20 °C until further processing.

Wet organ weights were measured in forty rats. The organ volumes could be calculated, assuming a specific density of 1.0.

In follow-up experiments plasma and bile samples were collected at various times after the same way of drug administration. A surgically placed



canule enabled the draining of the gall bladder.

Cumulative amounts of DAU and its major metabolite, DOL, in urine and bile and their concentrations in plasma and tissues were determined by straight phase high pressure liquid chromatography according to a standard procedure [Baurain et al., 1979]. Other metabolites were found only in negligible quantities.

## 4.1.2 Multicompartment Models

**4.1.2.1 Model Structures.** The body fluids and organs that were shown to be involved in the disposition of DAU in the rat can be represented by separate compartments in a model. The amount of drug any time present within a compartment is considered to be uniformly distributed. Three processes must be considered: the distribution of the parent drug DAU, the formation of its metabolite DOL and the distribution of DOL. The metabolite was found in the same tissues and fluids as the parent drug, which may be due to local formation and/or to distribution following formation elsewhere. So, two compartments are necessary for each organ; one for DAU to occupy, and its double for DOL. The compartments must be interlinked by anatomically and physiologically justified pathways along which the drug (parent and metabolite, respectively) might be transported, as well as pathways between corresponding DAU-DOL compartments for possible metabolite formation.

A multicompartment representation of the rat is shown in Fig. 4.1. By partitioning in several ways, models of various size can be made.

**4.1.2.2 Small-scale Model.** Neglecting metabolism, the rat model should comprise at least seven compartments (Fig. 4.2a). Plasma, as a general transfer fluid; urine, as excretion compartment; liver and spleen, as tissues with a special status: unlike other organs the spleen passes drug on to the liver instead of returning it to the plasma, and the liver, next to exchanging drug with the plasma, also excretes into the bile; and two separate other\_tissues compartments containing all other organs of importance, divided into observed and unobserved ones. Another way of allocating organs to either other\_tissues compartment may be based on their well-perfused or peripheral (respectively, fast or slow drug exchange with plasma) nature. Neglecting a possible enterohepatic cycle, i.e., reuptake of drug after release from the gallbladder via the gut lumen into the liver, this model involves 10 transfer rate constants.

In a model used, not for any practical pharmacokinetics but for purpose of evaluating the computational algorithms, only five compartments and seven transfer rate constants were considered (Fig. 4.2b).

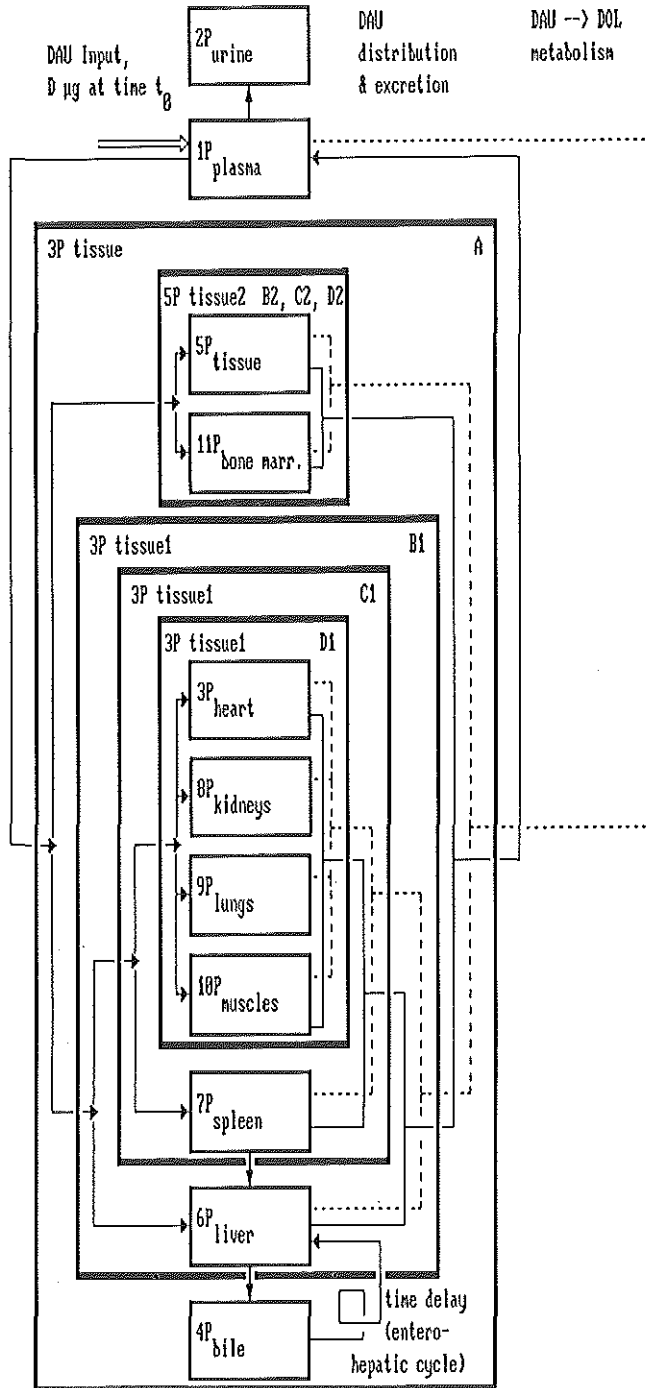
Small-scale models that do consider metabolism but are physiologically less realistic are a 2x3 compartment model (plasma, excretion, tissues; Fig 4.1b:

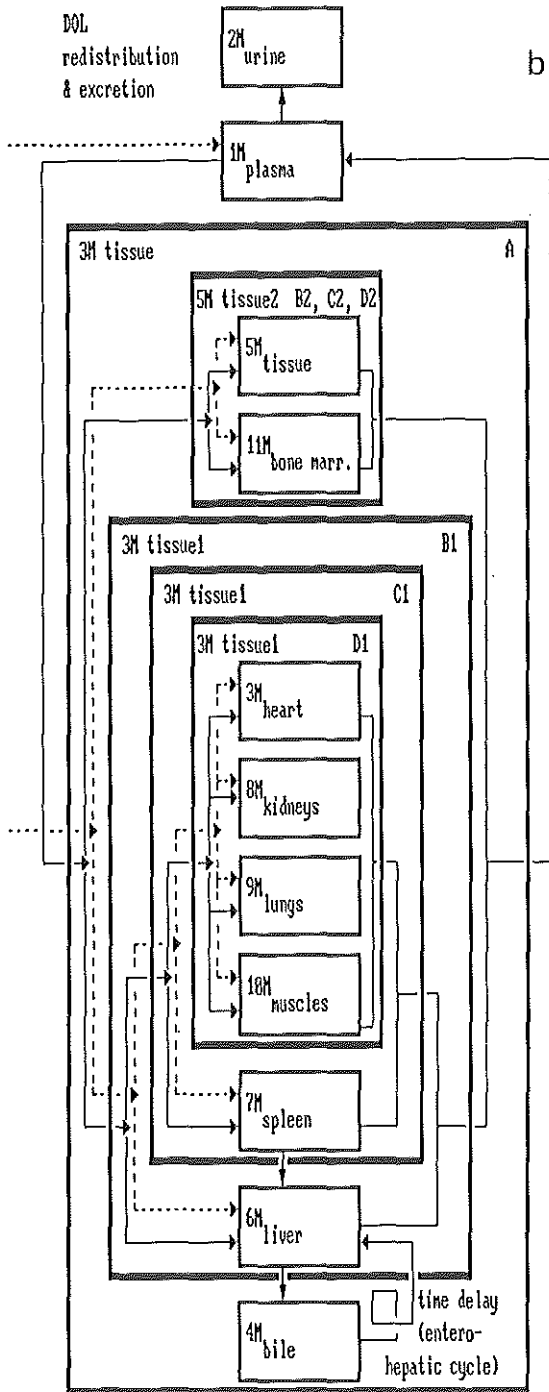
Fig. 4.1

Configurations of postulated models for DAU-DOL pharmacokinetics in the rat

The rat body (schematically in Fig. 4.1a) is lumped into 11 spaces representing tissue regions and body fluids that play a relevant role in the distribution, metabolism and elimination kinetics of daunomycin; each space has two compartments (P and M), respectively, for the parent drug and its major metabolite, daunomycinol; pathways are indicated along which DAU and DOL transport (solid lines) or DAU→DOL metabolism (broken lines) may take place.

A: tissues for 2x3 compartment model; B1,B2: tissues\_1, tissues\_2 for 2x5 compartment model; C1,C2: idem, 2x6 compartment model; D1,D2: idem, 2x7 compartment model





b

a

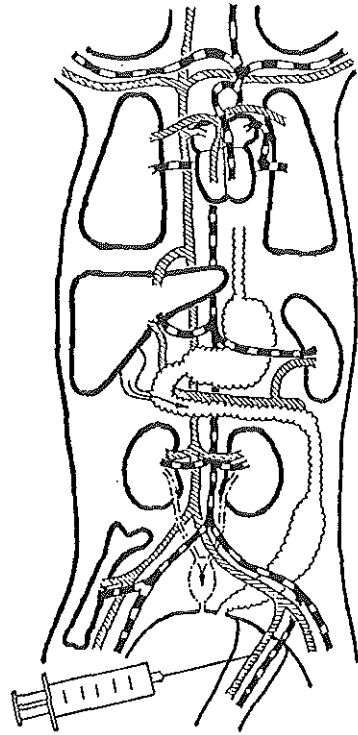
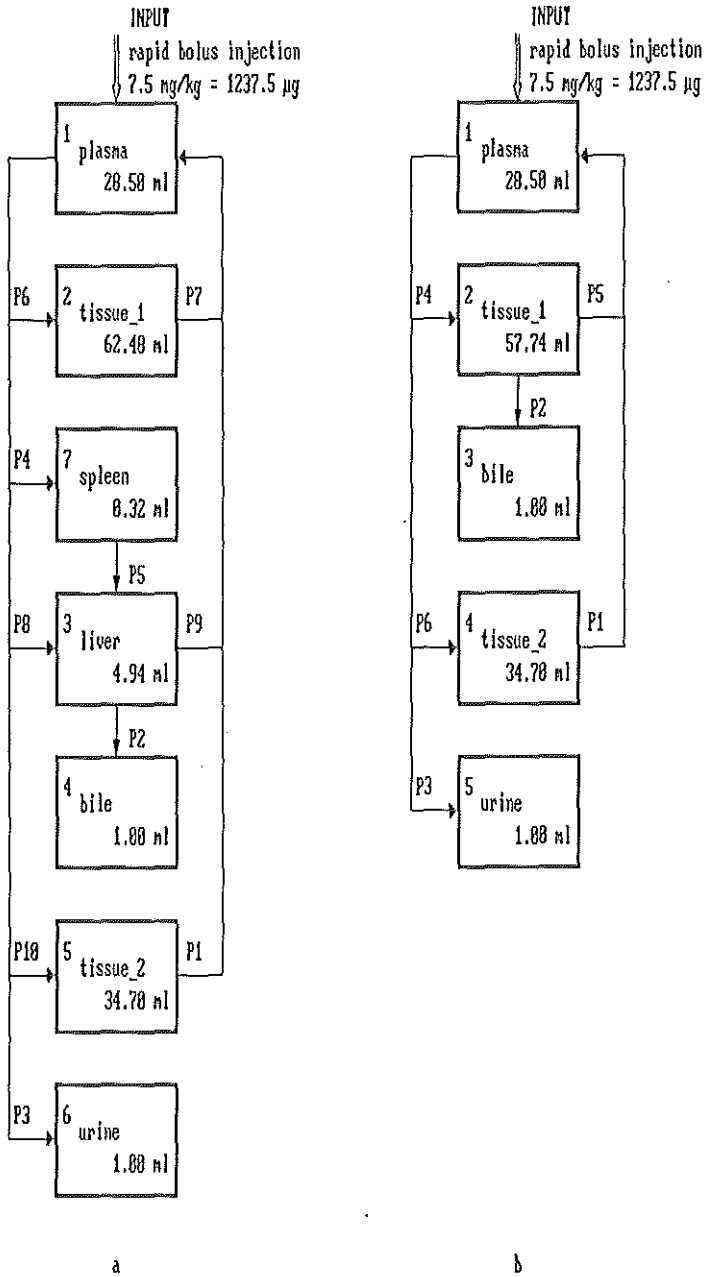


Fig. 4.2 Configurations of a seven (a) and a five (b) compartment model



A); a 2x5 compartment model (plasma, urine, observed and unobserved tissues, bile; Fig. 4.1b: B); and a 2x6 compartment model (as model B, but liver taken out of the observed tissues compartment; Fig. 4.1b: C).

**4.1.2.3 Medium-scale Model.** This model is the smallest model to be physiologically realistic. The rat body is now lumped into fourteen compartments, i.e., seven to model the DAU distribution and elimination, and again seven for DOL formation, distribution and excretion (Fig. 4.3). Data for observed organs that have been taken together can be obtained by pooling the separate measured drug concentrations, keeping in mind the relative volumes of the organs. Again neglecting the enterohepatic cycle, this model has 25 parameters.

**4.1.2.4 Large-scale Model.** Now, the most elaborate model (Fig. 4.4) comprises eleven double compartments: plasma plus extracellular body water; liver, spleen, heart, kidneys, lungs, bone marrow, muscles and other tissues; and urine and bile excretion. Various interlinking pathways can be assumed, and several transport/metabolism mechanisms. Six different cases will be considered (Table 4-1). First it is assumed that the formation of DOL is a first-order process that takes place in every organ at the same rate. After its formation DOL is excreted into urine through the plasma without redistribution. It is, however, more probable that DOL, transferred from an organ to the plasma, is exchanged again with other organs (redistribution) before it is excreted. A third possibility is that some organs are better able to form the metabolite than others; therefore, unequal transfer rates should be considered. On the other hand, metabolism might take place only in the liver—the chemical factory of the body. This means that presence of DOL in other organs is necessarily due to redistribution. Finally, metabolite formation might be a (saturable) zero-order, rather than a first-order linear process.

The most probable model must be identified by evaluating the models' adequacy in predicting the observed tissue drug concentrations.

**4.1.2.5 Assumptions on Drug Transport.** Within each compartment the drug is assumed to be instantaneously well-mixed throughout the compartment's volume. Thus, within a compartment itself there is no spatial concentration gradient. Between compartments such a gradient can exist and drug then is transported from one compartment to the other. In view of observed concentration—time histories, looking like exponential decay for most drugs in many organs (Fig. 4.5), for the transport processes first order kinetics are assumed, yielding first order ordinary linear differential equations. This corresponds with a state—space representation of passive diffusion processes. The amount of drug transported per unit time from a compartment  $i$  to an adjacent compartment  $j$ , is

Fig. 4.3 Configuration of a medium-scale model for DAU-DOL pharmacokinetics in the rat

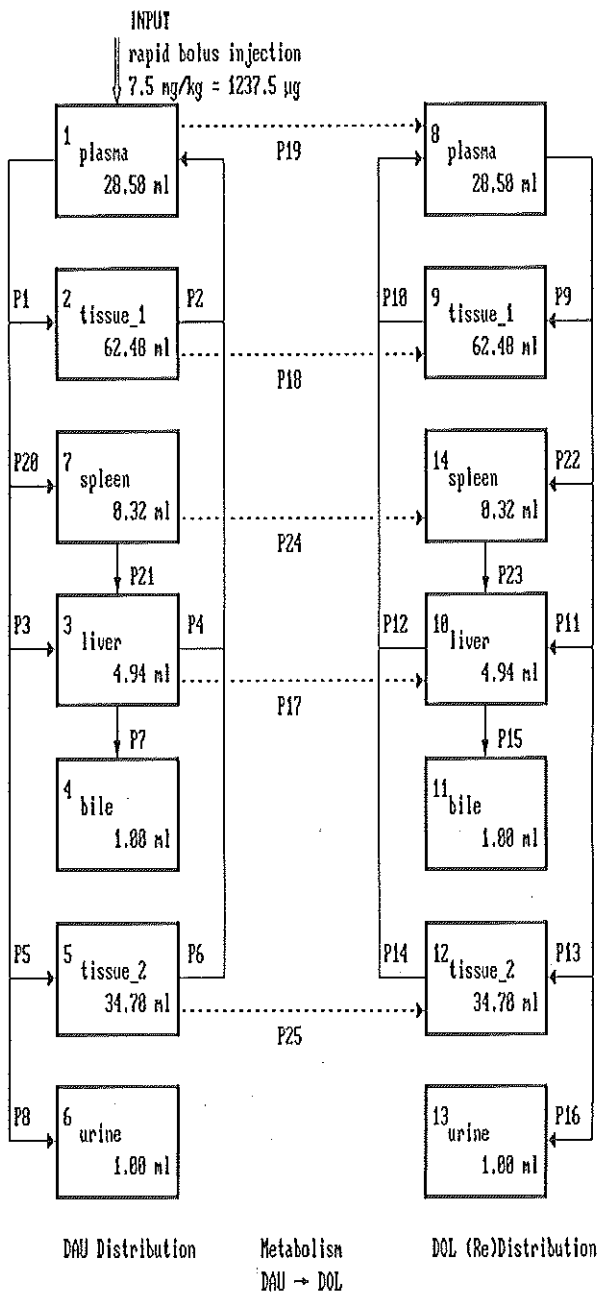


Fig. 4.4 Configuration of a large-scale model for DAU-DOL pharmacokinetics in the rat

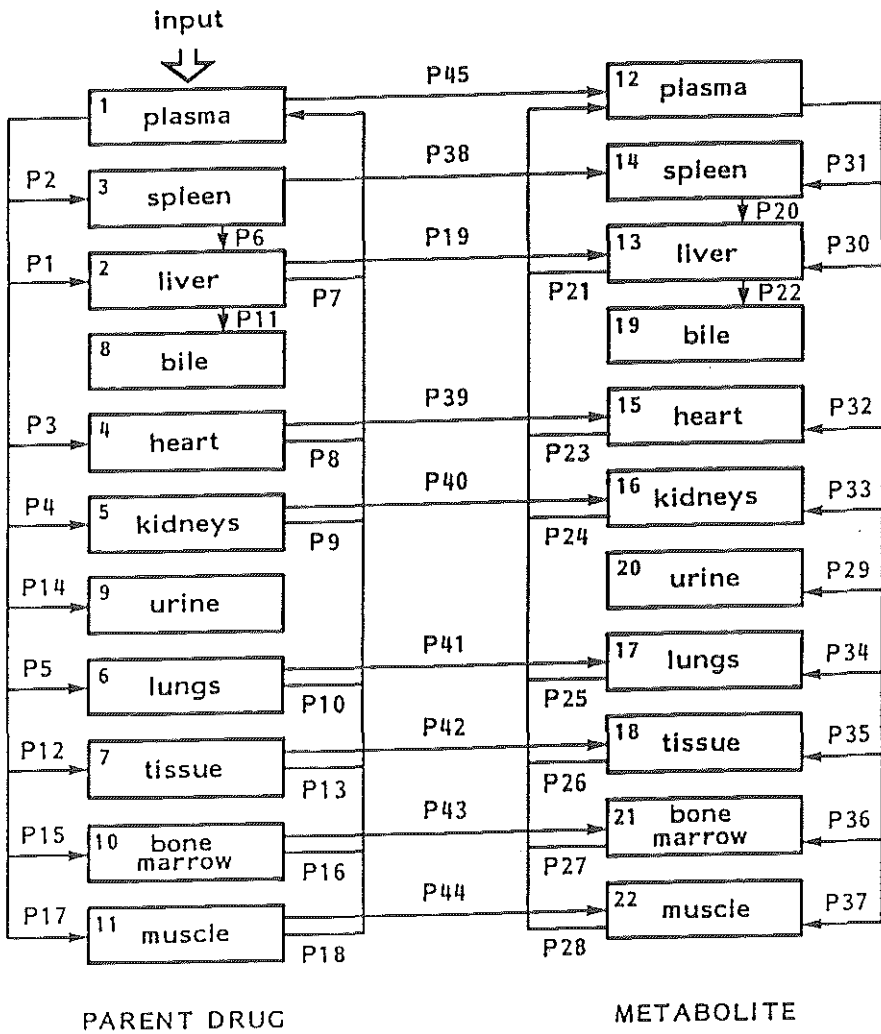


Fig. 4.5 Qualitative system response to a single rapid *i.v.* drug input  
 Three phases in the response can be distinguished: A = fast equilibrium between plasma and well-perfused organs and slow transfer to peripheral organs; B = excretion from well-perfused organs and plasma and still slow transfer to peripheral organs; C = return from peripheral organs to plasma, constant rate excretion, and a constant ratio of concentrations in plasma and well-perfused organs (after [Garrett, 1980])

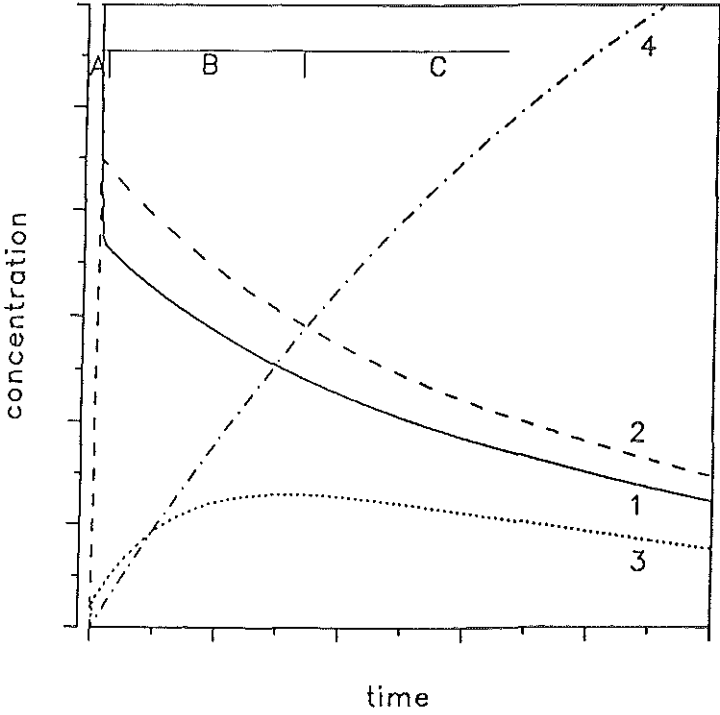
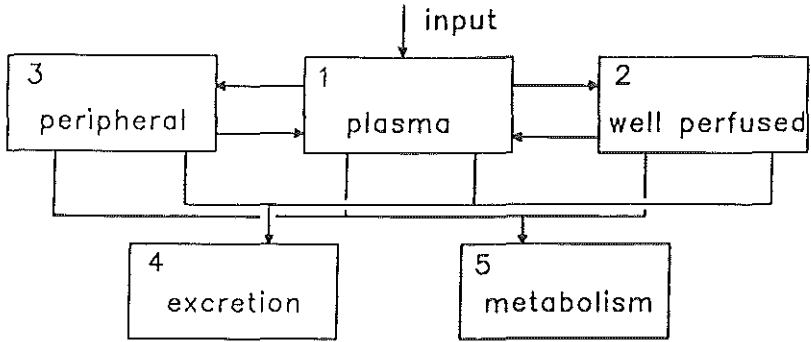




TABLE 4-1 LARGE-SCALE PHARMACOKINETIC MODELS FOR DAU-DOL (see Fig. 4.4)

| <u>model</u> | <u>remarks</u>  |
|--------------|---|
| 01           | <p><u>compartments:</u><br/>                     plasma, heart, kidneys, lungs, muscles, liver, spleen, bile, bone marrow, other_tissue, urine</p> <p><u>observations:</u><br/>                     no observations in other_tissue and bile</p> <p><u>metabolism:</u><br/> <u>first-order process</u> with equal rate constants in all organs (<math>P_{44}=P_{43}=P_{42}=P_{41}=P_{40}=P_{39}=P_{38}=P_{19}</math>), but not in plasma: <math>P_{45}=0</math>; no DOL redistribution, only excretion via plasma: <math>P_{30}=P_{31}=P_{32}=P_{33}=P_{34}=P_{35}=P_{36}=P_{37}=0</math></p> |
| 02           | as 01, but DOL redistribution pathways added ( $P_{30}$ through $P_{37}$ )  |
| 03           | as 02, but metabolic rate constants ( $P_{19}$ , $P_{38}$ through $P_{44}$ ) may differ among organs  |
| 04           | as 03, but metabolism in liver only<br>( $P_{45}=P_{44}=P_{43}=P_{42}=P_{41}=P_{40}=P_{39}=P_{38}=0$ )  |
| 05           | as 03, but also metabolism in plasma ( $P_{45} \neq 0$ )  |
| 06           | as 05, but <u>zero-order metabolism</u> process   |

N.B., other\_tissue: other tissues that can be reached by the drug but for which no observations are available. Tissues that cannot be reached by the drug, for example brain tissue and bones, are excluded from the models

proportional to the amount of drug present in the former compartment. Writing  $k_{j,i}$  for the time-invariant transfer rate constant (proportionality or diffusion constant), and  $V_i$  for the volume of compartment  $i$ , the rate of change of drug concentration  $x$  is given by:

$$\dot{x}_i = \frac{dx_i}{dt} = -k_{j,i} \cdot x_i \quad \text{and} \quad \dot{x}_j = k_{j,i} \cdot \frac{V_i}{V_j} \cdot x_i, \quad (4.1)$$

for the principle of mass conservation must hold.

If there is an external input of drug, e.g., at rate  $u_i(t)$ , Eq.(4.1) expands to:

$$\dot{x}_i = -k_{j,i} \cdot x_i + u_i \quad \text{and} \quad \dot{x}_j = k_{j,i} \cdot \frac{V_i}{V_j} \cdot x_i + u_j. \quad (4.2)$$

Similar equations can be drafted for other compartments. They can conveniently be combined in one equation in vector-matrix notation (see section 4.1.2.7; see

Appendix D that surveys a few elementary matrix manipulations):

$$\dot{\underline{x}} = A \cdot \underline{x} + B \cdot \underline{u}. \quad (4.3)$$

**4.1.2.6 Assumptions on Metabolism.** Again assuming first order kinetics, knowing that one mol of DAU yields one mol of DOL, equations analogous to Eq.(4.1) can be written. Because of a possibly limited enzyme capacity constant rate metabolism also must be considered. Then, per unit time a constant amount of drug is metabolized, thus transferred within a same tissue region from compartment i to its double, j:

$$\dot{x}_i = -K_{j,i} \quad \text{and} \quad \dot{x}_j = K_{j,i} \cdot \frac{V_i}{V_j}. \quad (4.4)$$

Although not considered in the present study, a combined constant/concentration—dependent metabolic process might be modeled with the Michaelis-Menten equation [Michaelis and Menten, 1913]:

$$\dot{x}_i = -\frac{K1_{j,i} \cdot x_i}{K2_{j,i} + x_i}, \quad (4.5)$$

which reduces to Eq.(4.4) for large  $x_i$ , and to Eq.(4.1) for small  $x_i$ .

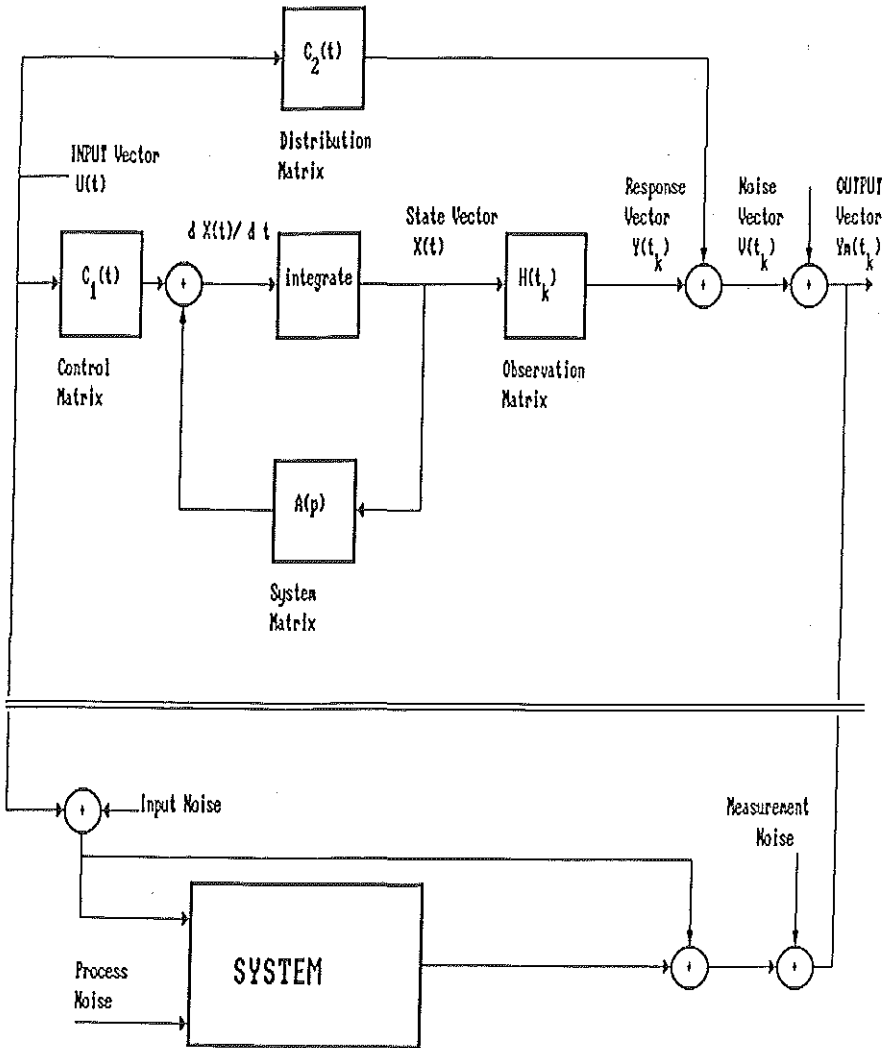
**4.1.2.7 Model Equations.** A block diagram representing the general case of a linear multivariable physical system is shown in Fig. 4.6. Variables that characterize the state of the system are elements of the state vector,  $\underline{x}(t)$ . Observed variables are elements of the response vector,  $\underline{y}(t)$ . Input variables are elements of the input vector  $\underline{u}(t)$ . The temporal change in the state of the system is related to the current state through the system matrix, A, which depends on (time-invariant) system parameters (elements of the parameter vector,  $\underline{p}$ ). It also depends on the input, through the control matrix,  $C_1(t)$ . This is expressed in the state equation:

$$\dot{\underline{x}}(t) = A(\underline{p}) \cdot \underline{x}(t) + C_1(t) \cdot \underline{u}(t). \quad (4.6)$$

The output equation relates the model response to the state through observation matrix  $H(t)$ —and to the input, through throughput or distribution matrix  $C_2(t)$ , although this latter matrix usually is zero for most physical systems:

$$\underline{y}(t) = H(t) \cdot \underline{x}(t) \quad [ + C_2(t) \cdot \underline{u}(t) ]. \quad (4.7)$$

Fig. 4.6 Block diagram for a general linear multivariable physical system  
 The system is characterized by the time-invariant system matrix  $A$  (function of parameters  $p$ ); the input  $u(t)$ , distributed by matrices  $C_1(t)$  and  $C_2(t)$ ; and state  $x(t)$  which is related by observation matrix  $H(t)$  to response  $y(t)$ . Noise from various sources, often modeled as a single noise vector  $v(t)$ , may pollute the observations  $y_m(t_k)$



Input Noise: fluctuations in control variables, e.g., dose or administration rate

Process Noise: fluctuations in uncontrolled process variables, e.g., temperature or pH

Various sources of noise may act on a system, causing a change in the response while the (nominal) input was kept constant (Fig. 4.6). Noise may corrupt the input (e.g., random fluctuations in rate of drug administration), the system itself (e.g., changes in ambient temperature, intermittent food uptake) and the output. In the present study, for simplification all noise will be considered to accumulate in random (measurement) errors in the output only.

The state vector in a model existing of  $2n$  compartments ( $n$  for DAU and  $n$  for DOL distribution, respectively) contains the  $2n$  drug concentrations. These are zero before time  $t_0$ . In the present case the input to the system can be modeled as an initial condition. The dose  $D$  is administered such, that instantaneous loading of the first compartment (plasma-DAU; volume  $V_1$ ) can be assumed at time  $t_0$ . Let the pathway structure be such, that  $m$  transfer rate constants (the model parameters) can be distinguished. Let there be  $N$  observation times, with observations available for  $w$  out of  $2n$  compartments (N.B., for any of these  $w$  observed compartments not necessarily all observation times must yield a measured value). If process noise is neglected, general model equations for the state—space representation of this first order linear time-invariant system reduce to:

$$\dot{\underline{x}}(t) = A(\underline{p}) \cdot \underline{x}(t) \quad \text{with initial condition } \underline{X}(t_0) = \text{col}[D/V_1, 0, \dots, 0], \quad (4.8)$$

$$\underline{y}(t_k) = H(t_k) \cdot \underline{x}(t_k), \quad (4.9)$$

$$\underline{y}_m(t_k) = \underline{y}(t_k) + \underline{v}(t_k), \quad (4.10)$$

where

- $\underline{x}(t)$  is a  $2n \times 1$  state vector, i.e., its elements are the concentrations in the  $2n$  defined compartments,
- $A(\underline{p})$  is a  $2n \times 2n$  system matrix whose elements depend on  $m$  parameters (transfer rate constants) that are gathered in the  $m \times 1$  parameter vector  $\underline{p}$ ,
- $\underline{y}(t_k)$  is a  $w \times 1$  ( $w \leq 2n$ ) model response vector at sample time  $t_k$ , whose elements thus are the concentrations in the  $w$  observed compartments in the system,
- $H(t_k)$  is a  $w \times 2n$  observation matrix relating the model response to the state vector at time  $t_k$ ,
- $\underline{y}_m(t_k)$  is a  $w \times 1$  vector with observations at time  $t_k$ ,
- $\underline{v}(t_k)$  is a  $w \times 1$  measurement noise vector at time  $t_k$ , which is assumed to originate from Gaussian random distributions characterized by a zero mean value and a certain variance.

TABLE 4-2 STATE EQUATIONS FOR THE MEDIUM-SCALE MODEL OF FIG. 4.3

The state equations are, based on mass balance:

$$\begin{aligned}
 \dot{x}_1(t) &= -(p(1) + p(3) + p(5) + p(8) + p(20) + p(19)) \cdot x_1(t) + p(2) \cdot x_2(t) \cdot V(2)/V(1) + \\
 &\quad p(4) \cdot x_3(t) \cdot V(3)/V(1) + p(6) \cdot x_5(t) \cdot V(5)/V(1), \\
 \dot{x}_2(t) &= p(1) \cdot x_1(t) \cdot V(1)/V(2) - (p(2) + p(18)) \cdot x_2(t), \\
 \dot{x}_3(t) &= p(3) \cdot x_1(t) \cdot V(1)/V(3) - (p(4) + p(7) + p(17)) \cdot x_3(t) + p(21) \cdot x_7(t) \cdot V(7)/V(3), \\
 \dot{x}_4(t) &= p(7) \cdot x_3(t) \cdot V(1)/V(4), \\
 \dot{x}_5(t) &= p(5) \cdot x_1(t) \cdot V(1)/V(5) - (p(6) + p(25)) \cdot x_5(t), \\
 \dot{x}_6(t) &= p(8) \cdot x_1(t) \cdot V(1)/V(6), \\
 \dot{x}_7(t) &= p(20) \cdot x_1(t) \cdot V(1)/V(7) - (p(21) + p(24)) \cdot x_7(t), \\
 \\
 \dot{x}_8(t) &= p(19) \cdot x_1(t) - (p(9) + p(11) + p(13) + p(16) + p(22)) \cdot x_8(t) + \\
 &\quad p(10) \cdot x_9(t) \cdot V(9)/V(8) + p(12) \cdot x_{10}(t) \cdot V(10)/V(8) + p(14) \cdot x_{12}(t) \cdot V(12)/V(8), \\
 \dot{x}_9(t) &= p(18) \cdot x_2(t) + p(9) \cdot x_8(t) \cdot V(8)/V(9) - p(10) \cdot x_9(t), \\
 \dot{x}_{10}(t) &= p(17) \cdot x_3(t) + p(11) \cdot x_8(t) \cdot V(8)/V(10) - (p(12) + p(15)) \cdot x_{10}(t) + \\
 &\quad p(23) \cdot x_{14}(t) \cdot V(14)/V(10), \\
 \dot{x}_{11}(t) &= p(15) \cdot x_{10}(t) \cdot V(10)/V(11), \\
 \dot{x}_{12}(t) &= p(25) \cdot x_5(t) + p(13) \cdot x_8(t) \cdot V(8)/V(12) - p(14) \cdot x_{12}(t), \\
 \dot{x}_{13}(t) &= p(16) \cdot x_8(t) \cdot V(8)/V(13), \\
 \dot{x}_{14}(t) &= p(24) \cdot x_7(t) + p(22) \cdot x_8(t) \cdot V(8)/V(14) - p(23) \cdot x_{14}(t),
 \end{aligned}$$

in which the transfer rate constants,  $k_{j,i}$ , are combined in a vector,  $\mathbf{p}$ , and compartment volumes in a vector,  $\mathbf{V}$ ; e.g.,  $p(21) = k_{3,7}$  (to liver from spleen),  $V(1)$  is the volume of compartment 1. These equations can be written as:

$$\begin{aligned}
 \dot{x}_1(t) &= A(1,1) \cdot x_1(t) + A(1,2) \cdot x_2(t) + \dots + A(14,1) \cdot x_{14}(t), \\
 \dot{x}_2(t) &= A(2,1) \cdot x_1(t) + A(2,2) \cdot x_2(t) + \dots + A(14,2) \cdot x_{14}(t), \\
 &\quad \dots \\
 \dot{x}_{14}(t) &= A(14,1) \cdot x_1(t) + A(14,2) \cdot x_2(t) + \dots + A(14,14) \cdot x_{14}(t),
 \end{aligned}$$

where the elements of system matrix  $A$  depend on the parameters  $p(1)$  through  $p(25)$  and on the compartment volumes  $V(1)$  through  $V(14)$ :

$$\begin{aligned}
 A(1,1) &= -(p(1) + p(3) + p(5) + p(8) + p(20) + p(19)), \\
 A(1,2) &= p(2) \cdot V(2)/V(1), \\
 A(1,3) &= p(4) \cdot V(3)/V(1), \\
 A(1,4) &= 0; \quad \text{etc.},
 \end{aligned}$$

or, in matrix-vector notation:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \vdots \\ \dot{x}_{14}(t) \end{bmatrix} = \begin{bmatrix} A(1,1) & A(1,2) & \dots & A(14,1) \\ A(2,1) & A(2,2) & \dots & A(14,2) \\ \vdots & \vdots & \ddots & \vdots \\ A(14,1) & A(14,2) & \dots & A(14,14) \end{bmatrix} \cdot \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{14}(t) \end{bmatrix}$$

which is equivalent with:  $\dot{\mathbf{x}}(t) = \mathbf{A} \cdot \mathbf{x}(t)$

TABLE 4-3 A FEW MATHEMATICAL MANIPULATIONS

Runge-Kutta integration

Given state vector  $\underline{x}$  at time  $t_0$ ,  $\underline{x}(t_0) = \underline{x}_0$ ; to find the state vector at time  $t_1 = t_0 + \Delta t$ ,  $\underline{x}(t_1) = \underline{x}_1$ , calculate subsequently:

$$\begin{array}{llll} \underline{x}_0 & f(\underline{x}_0) & = A \cdot \underline{x}_0 & \underline{k}_1 = \Delta t \cdot f(\underline{x}_0) \\ \underline{x}_0 + \frac{1}{2} \cdot \underline{k}_1 & f(\underline{x}_0 + \frac{1}{2} \cdot \underline{k}_1) & = A \cdot (\underline{x}_0 + \frac{1}{2} \cdot \underline{k}_1) & \underline{k}_2 = \Delta t \cdot f(\underline{x}_0 + \frac{1}{2} \cdot \underline{k}_1) \\ \underline{x}_0 + \frac{1}{2} \cdot \underline{k}_2 & f(\underline{x}_0 + \frac{1}{2} \cdot \underline{k}_2) & = A \cdot (\underline{x}_0 + \frac{1}{2} \cdot \underline{k}_2) & \underline{k}_3 = \Delta t \cdot f(\underline{x}_0 + \frac{1}{2} \cdot \underline{k}_2) \\ \underline{x}_0 + \underline{k}_3 & f(\underline{x}_0 + \underline{k}_3) & = A \cdot (\underline{x}_0 + \underline{k}_3) & \underline{k}_4 = \Delta t \cdot f(\underline{x}_0 + \underline{k}_3) \\ \underline{x}_1 = \underline{x}_0 + (\underline{k}_1 + 2 \cdot \underline{k}_2 + 2 \cdot \underline{k}_3 + \underline{k}_4) / 6 & & & \end{array}$$

Simpson's rule

The integral of a function  $f(x)$  on interval  $[a,b]$  can be approximated—if the value of  $f(x)$  is known in  $2 \cdot n + 1$  equidistant points  $c_i$  on this interval ( $c_0 = a$ ,  $c_{2n} = b$ )—by:

$$\int_a^b f(x) dx \approx \frac{b-a}{6 \cdot n} \left\{ f(a) + 4 \cdot \sum_{i=1}^n f(c_{2i-1}) + 2 \cdot \sum_{i=1}^{n-1} f(c_{2i}) + f(b) \right\}$$

Cramer's rule

To solve  $x_1$  and  $x_2$  from a set of 2 algebraic equations,

$$\begin{array}{l} a_{11} \cdot x_1 + a_{12} \cdot x_2 = b_1 \\ a_{12} \cdot x_1 + a_{22} \cdot x_2 = b_2 \end{array} \quad \text{or} \quad \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix},$$

according to Cramer's rule:

$$x_1 = \frac{\det(A_1)}{\det(A)}, \quad x_2 = \frac{\det(A_2)}{\det(A)},$$

where

$$\begin{array}{l} \det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}, \\ \det(A_1) = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix} = b_1 \cdot a_{22} - a_{12} \cdot b_2 \text{ and} \\ \det(A_2) = \begin{vmatrix} a_{11} & b_1 \\ a_{12} & b_2 \end{vmatrix} = a_{11} \cdot b_2 - b_1 \cdot a_{12}. \end{array}$$

E.g.,  $a_{11} = 2$ ,  $a_{12} = 1$ ,  $a_{21} = 1$ ,  $a_{22} = 2$ ,  $b_1 = 13$  and  $b_2 = 11$  yield:  $x_1 = 15/3 = 5$  and  $x_2 = 9/3 = 3$ . This can easily be extended to larger sets of algebraic equations in an analogous way.

As an example, the state equations describing the medium-scale model of Fig. 4.3 are shown in Table 4-2.

**4.1.2.8 Solving the Model Equations.** For small-scale models it is possible to solve the model equations analytically, e.g., by using Laplace transformation. This method is elucidated in Appendix E for the 2x3 compartment model shown in Fig. 4.7. Although in theory possible, this method quickly becomes unwieldy when larger models are concerned. Then it is easier to solve the 2n model equations numerically than analytically. The state vector as function of time can be found in different ways, e.g., through direct numerical integration (*Runge-Kutta algorithm*) or by the so-called *transition matrix method*.

Runge-Kutta Integration. For a given set of parameter values, i.e., the m transfer rate constants combined in a vector  $\underline{p}_0$ , the system matrix A is known. Starting from the known value at time  $t_0$  (initial condition) the state vector at time t, a time step  $\Delta t$  later ( $t = t_0 + \Delta t$ ), can be calculated from:

$$\underline{x}(t) = \underline{x}(t_0) + \int_{t_0}^t \dot{\underline{x}}(T) dT = \underline{x}(t_0) + \int_{t_0}^t A \cdot \underline{x}(T) dT. \quad (4.11)$$

The Runge-Kutta algorithm to perform the integration is illustrated in Table 4-3. To avoid inaccuracy due to accumulation of round-off errors it is necessary to take a small time step  $\Delta t$  (in practice,  $\frac{1}{2}$ -2 min). To calculate the state vector at large t, the interval  $[t, t_0]$  should be divided into many small time steps  $\Delta t$  and Eq.(4.11) should be used recursively. Observation times ranging up to 48 h, this may turn out to be a time-consuming procedure.

Transition Matrix Method. It can easily be proven that a general solution for the state vector can be written as:

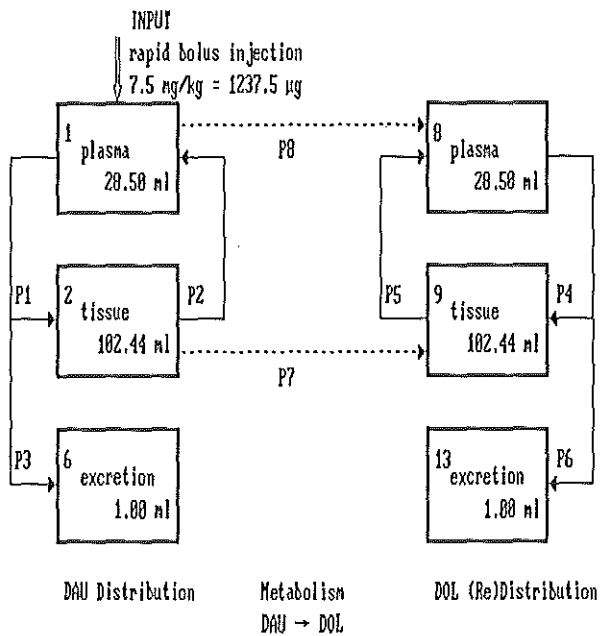
$$\underline{x}(t) = \Phi(t, t_0) \cdot \underline{x}(t_0), \quad (4.12)$$

provided that for the 2nx2n transition matrix  $\Phi$  the following formula holds:

$$\dot{\Phi}(t, t_0) = A(\underline{p}_0) \cdot \Phi(t, t_0); \quad \text{initial condition } \Phi(t_0, t_0) = I \quad (4.13)$$

(I denotes the 2nx2n identity matrix). From Eq.(4.13) the transition matrix can be solved, e.g., again by Runge-Kutta integration. Alternatively, the transition matrix may be obtained from a numerical evaluation of its analytical solution,

Fig. 4.7 Configuration of a 2x3 compartment model for DAU-DOL pharmacokinetics



$$\Phi(t, t_0) = \exp\{A \cdot (t - t_0)\} = \exp(A \cdot \Delta t) \tag{4.14}$$

To this purpose the exponential function is expanded into the series:

$$\exp(A \cdot \Delta t) = I + A \cdot \Delta t + \frac{A^2 \cdot \Delta t^2}{2!} + \frac{A^3 \cdot \Delta t^3}{3!} + \dots \tag{4.15}$$

(c!, c factorial, denotes 1x2x3...xc; so, e.g., 3! = 1x2x3 = 6).

As Eq.(4.15) converges for all  $\Delta t$ , only a limited number of terms needs to be evaluated to arrive at a sufficiently accurate approximation. At present the first 25 terms are computed and time step  $\Delta t = 0.5$  min is taken. The state vector subsequently follows from Eq.(4.12).

The advantage of the transition matrix method is that this matrix needs only be computed for a short time interval  $[t_0, t]$ , as for larger time intervals it can be found by simple matrix multiplications [Kwakernaak, 1972]. This means a considerable reduction in computation time. For example, it can be easily proven that, for integer a and b:



$$\Phi((a+b) \cdot t, t_0) = \Phi(a \cdot t, t_0) \cdot \Phi(b \cdot t, t_0). \quad (4.16)$$

Thus, if the state vector is to be known at regular time intervals  $\Delta t$  only, a recursive formula can be used:

$$\underline{x}(t+\Delta t) = \Phi(t+\Delta t, t) \cdot \underline{x}(t), \quad \text{or} \quad \underline{x}_{k+1} = \Phi(\Delta t) \cdot \underline{x}_k, \quad (4.17)$$

in which  $\Phi(\Delta t)$  is a constant and needs to be computed only once.

If effects of constant rate metabolism are to be included, the state Eq.(4.8) must be slightly modified. The system matrix  $A$  involves only the parameters associated with the drug (re)distribution and a  $2 \times 1$  vector  $\underline{z}$  is added, whose non-zero elements involve the metabolism parameters:

$$\dot{\underline{x}}(t) = A(\underline{p}_0) \cdot \underline{x}(t) + \underline{z}(\underline{p}_0). \quad (4.18)$$

These state eqs.(4.18) replacing eqs.(4.8) have the following solution:

$$\underline{x}(t) = \Phi(t, t_0) \cdot \underline{x}(t_0) + \int_{t_0}^t \Phi(t, T) \cdot \underline{z} dT, \quad (4.19)$$

where the transition matrix  $\Phi$  again can be obtained by numerical integration of Eq.(4.13), or through Eq.(4.15), followed by selfmultiplication(s). The integral in Eq.(4.19) can be evaluated by application of the well-known Simpson's rule (see Table 4-3) to each matrix element separately.

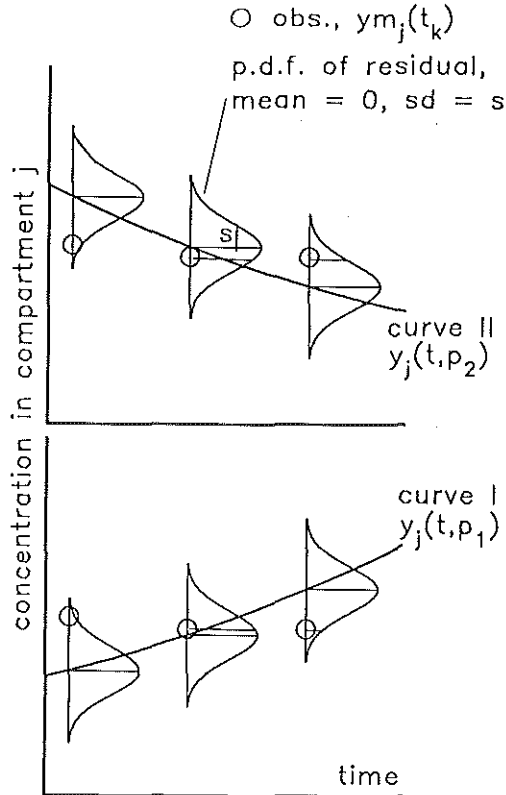
### 4.1.3 Parameter Estimation

If the chosen pharmacokinetic model is the right one, the model response for a given drug input should correspond to a high degree with concentrations observed as output. (Of course, as the observed concentrations are noise-corrupted, a perfect correspondence will be extremely rare.) Given the model structure, the model response varies with the choice of the parameter values (transfer rates). This choice should therefore be optimized before it can be decided whether the model is or is not a good model of the pharmacokinetic system.

Apart from the method used with the  $2 \times 3$  compartment model (Appendix E), for parameter estimation a few well-known techniques are available [e.g., Eijkhoff, 1974; Gill, Murray and Wright, 1981]. In principle, the procedure is as follows. For any parameter vector  $\underline{p}_0$ , whose elements can in

Fig. 4.8 Maximum Likelihood Optimization

Parameter values ( $p_2$ ) yielding response  $y_j(t, p_2)$  in compartment  $j$  (i.e., curve II) are more likely than those ( $p_1$ ) yielding response  $y_j(t, p_1)$  (i.e., curve I), because in the former case the likelihood of the residuals  $y_{mj}(t_k) - y_j(t_k)$ ,  $k = 1, 2, \dots$  is higher



principle be chosen 'arbitrarily' (i.e., in practice, based on available *à priori* information, the most appropriate guess is taken), the time course of the state vector can be calculated with Eq.(4.14) and Eq.(4.17) or Eq.(4.19). Then, for all  $N$  time points  $t_k$  the model response  $y(t_k)|_{p_0}$  is compared with the datapoints in all observed compartments simultaneously, from which a new set of parameter values—closer to the true ones—is deduced. In the present case a stochastic maximum likelihood (ML) estimation technique [e.g., Åström, 1979; Eijkhoff, 1974] was chosen because of its properties of asymptotic unbiasedness and efficiency (i.e., if the number of observations becomes large enough the true parameter values will be found eventually, respectively, the variance-covariance matrix of the parameters approaches a lower bound given by the inverted

Fisher's information matrix ( $E^{-1}$ , see below), thus yielding an estimate of the accuracy of the estimated parameter values). In this approach the parameters are considered random variables with a Gaussian probability density function ( $N\{\underline{p}_{ML}; E^{-1}\}$ ), rather than unknown constants.

**4.1.3.1 ML Technique.** Assuming that the chosen model equations describe the system correctly, in principle the ML method finds those parameter values and corresponding model response that yield the highest probability of occurrence of the residuals (i.e., the remaining deviations between model response,  $\underline{y}$ , and observations,  $\underline{y}_m$ ; see Fig. 4.8). Therefore, some probability density function (pdf) for this occurrence must be assumed. Usually a Gaussian function can be expected, with a zero mean value and some variance-covariance matrix. Writing  $\underline{v}(t_k)$  for the residuals at  $t_k$ , let this Gaussian pdf,  $N\{\underline{0}, Q(t_k)\}$ , be given by:

$$\begin{aligned} Ep\{\underline{v}(t_k)\} &= \underline{0} \quad \text{and} \\ Ep\{[\underline{v}(t_k) - Ep\{\underline{v}(t_k)\}] \cdot [\underline{v}(t_k) - Ep\{\underline{v}(t_k)\}]^T\} &= Q(t_k). \end{aligned} \quad (4.20)$$

( $Ep$  denotes the expected value of;  $Pr$  denotes the probability of;  $\underline{0}$  is a  $w \times 1$  vector and  $Q$  a  $w \times w$  matrix). The probability of the present residuals, for  $\underline{p}_0$ , will then be:

$$\begin{aligned} pdf\{\underline{v}(t_k) | \underline{p}_0\} &= Pr\{\underline{v}(t_k) | \underline{p}_0\} = \\ &= (2\pi)^{-1/2w} \cdot |Q(t_k)|^{-1/2} \cdot \exp\{-1/2 \cdot \underline{v}(t_k)^T \cdot Q(t_k)^{-1} \cdot \underline{v}(t_k)\}. \end{aligned} \quad (4.21)$$

If stationarity is assumed, then Eq.(4.21) is valid for every  $t_k$ . If it is further assumed that each observation is independent on any other one (both in the same compartment at different times and in different compartments at the same time, i.e., temporal and spatial independence), then  $Q(t_k) = Q = \text{diag}(\text{var}_i)$ , where  $i=1$  through  $w$ , and  $\text{var}_i$  is the variance of the residuals in compartment  $i$ . In that case, for the residuals the joint pdf, or likelihood function  $L$ , is given by:

$$\begin{aligned} L &= pdf\{\underline{v}(t_1), \underline{v}(t_2), \dots, \underline{v}(t_N) | \underline{p}_0\} = \\ &= (2\pi)^{-1/2wN} \cdot |Q|^{-1/2N} \cdot \exp\{-1/2 \cdot \sum_{k=1}^N \underline{v}(t_k)^T \cdot Q^{-1} \cdot \underline{v}(t_k)\} = \\ &= \prod_{k=1}^N Pr\{\underline{v}(t_k) | \underline{p}_0\} \quad \text{for } N \text{ observation times.} \end{aligned} \quad (4.22)$$

Substituting the ML values,  $\underline{p}_{ML}$ , for  $\underline{p}_0$  maximizes  $L$  and corresponds with the

optimum model response, under the given assumptions. If, in that situation, it does not satisfactorily explain the observations, then the model used is obviously not adequate and should be replaced.

To find the ML values, starting from arbitrary initial values  $\mathbf{p}_0$ , some numerical search algorithm must be employed that maximizes the function  $L$  with respect to the parameters. By optimizing this likelihood function, instead of any other function of  $\mathbf{p}$ , parameter values are obtained that have the highest probability of being true (provided that the used model is correct). Minimizing the function  $-\ln(L)$ :

$$-\ln(L) = \frac{1}{2} \cdot N \cdot w \cdot \ln(2\pi) + \frac{1}{2} \cdot N \cdot \ln(|Q|) + \frac{1}{2} \cdot PF(\mathbf{p}), \quad (4.23)$$

by minimizing the so-called performance index, PF:

$$PF = \sum_{k=1}^N \underline{y}(t_k)^T \cdot Q^{-1} \cdot \underline{y}(t_k), \quad (4.24)$$

where  $\underline{y}$  denotes measurement noise at sample time  $t_k$ :

$$\underline{y} = \underline{y}_m - \underline{y}(\mathbf{p}) = \underline{y}_m - H \cdot \underline{x}(\mathbf{p}), \quad (4.25)$$

yields the same result, but may be more convenient.

Next to its asymptotic unbiasedness the other advantageous property of the ML method is its asymptotic efficiency. In other words, the pdf of the parameters tends to a Gaussian distribution with the estimated values as mean and the inverted Fisher information matrix (or Cramer-Rao lower boundary) as variance-covariance matrix. The Fisher information matrix is defined by:

$$E = -E_p \left\{ \frac{\partial^2 \ln(L)}{\partial \mathbf{p}^2} \right\} = E_p \left\{ \left[ \frac{\partial \ln(L)}{\partial \mathbf{p}} \right]^T \left[ \frac{\partial \ln(L)}{\partial \mathbf{p}} \right] \right\}. \quad (4.26)$$

Minimization of  $-\ln(L)$  yields the wanted ML parameter values. This, in turn, means that the derivatives, with respect to the parameters of the function  $PF(\mathbf{p})$  (Eq.(4.24)) must all be set to zero. If  $Q$  is unknown—as it usually is—it can be estimated by maximization of  $\ln(L)$  with respect to  $Q$ . An approximation for this is:

$$Q = \frac{1}{N} \cdot \sum_{k=1}^N (\underline{y}_m - H \cdot \underline{x})_{t_k} \cdot (\underline{y}_m - H \cdot \underline{x})_{t_k}^T \quad (4.27)$$

**4.1.3.2 Minimization Routines.** Starting from some—in principle—arbitrarily chosen set of initial parameter values, an iterative computer program must take steps in the parameter space toward the ML values of the parameters that maximize the function L. A schematic flow diagram for the computations is given in Table 4-4. Three different numerical and iterative optimization routines were considered:

- 1) a *modified Gauss-Newton* (MGN) method, which requires solving of analytical sensitivity (ordinary first order differential) equations;
- 2) a *finite differences* (FD) approach to the computation;
- 3) a *direct determination of the Hessian matrix* of second order derivatives and the gradient vector of the log-likelihood function in the parameter space, again via finite differences (DDH method).

**MGN Optimization Algorithm.** This gradient search method is derived from the well-known powerful Newton-Raphson procedure [Allen, 1983]. Figure 4.9 shows the principle of the NR-procedure for a one-dimensional parameter space. According to a Taylor-series expansion the value of a function  $J(\underline{p})$  for parameter vector  $\underline{p}_1$  can be found from its value for parameter vector  $\underline{p}_0$ , if  $\underline{p}_1$  and  $\underline{p}_0$  differ by a small amount,  $\Delta \underline{p} = \underline{p}_1 - \underline{p}_0$ :

$$J(\underline{p}_1) = J(\underline{p}_0) + \frac{\partial J(\underline{p}_0)}{\partial \underline{p}_0} \cdot \Delta \underline{p} + \frac{1}{2!} \cdot \Delta \underline{p}^T \cdot \left[ \frac{\partial^2 J(\underline{p}_0)}{\partial \underline{p}_0^2} \right] \cdot \Delta \underline{p} + O(\Delta \underline{p}^3). \quad (4.28)$$

Substitution of PF for J and neglecting the higher order ( $O(\Delta \underline{p}^3)$ ) rest term, Eq.(4.28) converts to:

$$\frac{PF(\underline{p}_1) - PF(\underline{p}_0)}{\Delta \underline{p}} = \frac{\Delta PF}{\Delta \underline{p}} = \frac{\partial PF(\underline{p}_0)}{\partial \underline{p}_0} + \frac{1}{2} \cdot \Delta \underline{p}^T \cdot \left[ \frac{\partial^2 PF(\underline{p}_0)}{\partial \underline{p}_0^2} \right]. \quad (4.29)$$

This, in turn, by considering the limit  $\Delta \underline{p} \rightarrow \underline{0}$  and taking  $\Delta PF/\Delta \underline{p} = 0$  at the minimum of the function PF, converts to:

TABLE 4-4 COMPUTATIONAL FLOW CHART

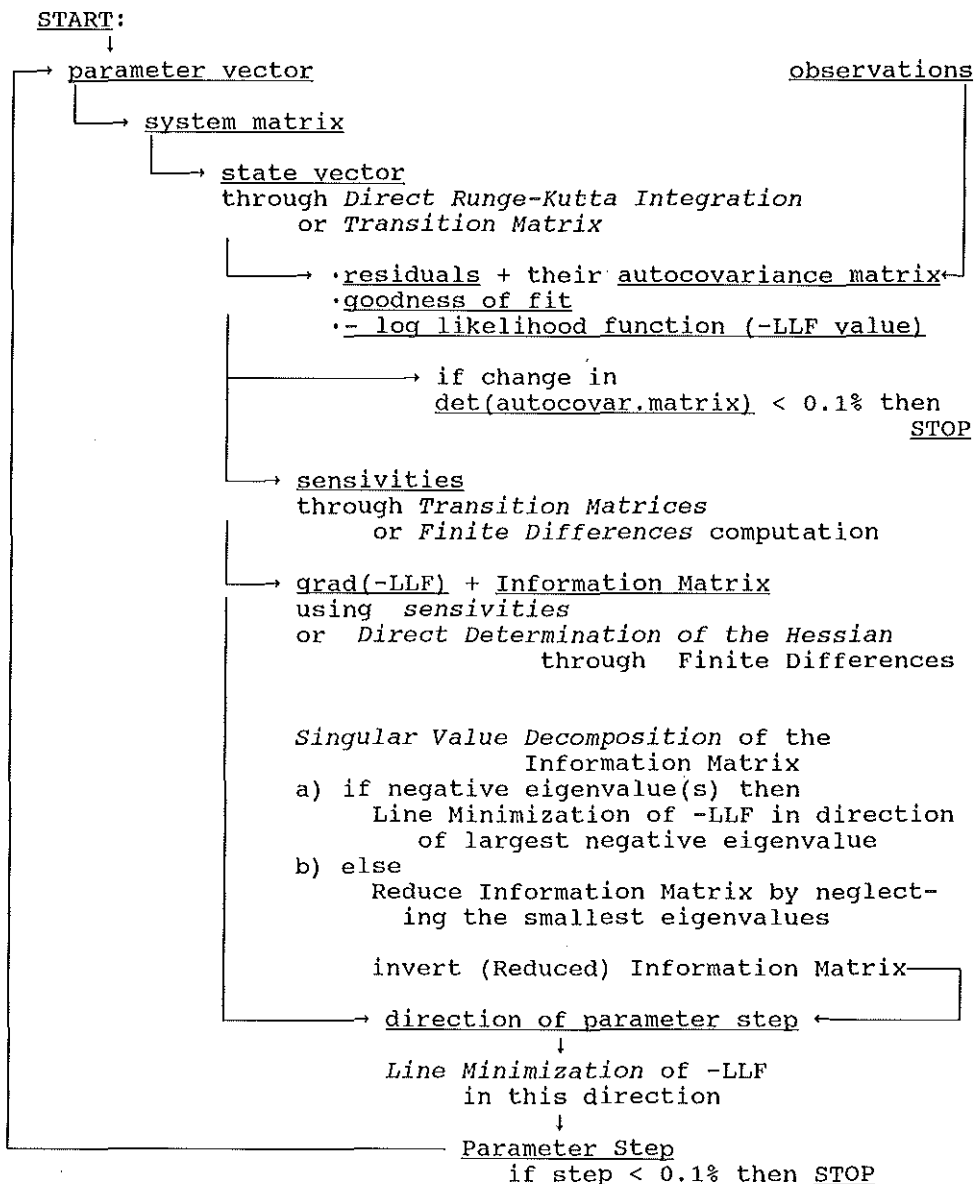
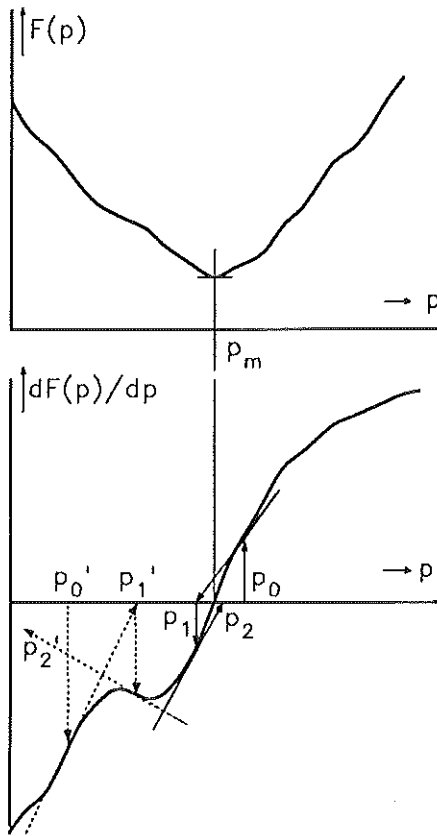


Fig. 4.9 Function  $F$  of parameter  $p$ ; the minimum is found at  $F(p_m)$  by the *Newton-Raphson algorithm*; starting at  $p_0$ :

$$p_1 = p_0 \cdot \left\{ \frac{d^2 F(p_0)}{dp^2} \right\}^{-1} \cdot \frac{dF(p_0)}{dp}, \quad p_2 = p_1 \cdot \left\{ \frac{d^2 F(p_1)}{dp^2} \right\}^{-1} \cdot \frac{dF(p_1)}{dp}, \quad \text{etc.}$$

N.B., starting from  $p_0'$  results in divergence, i.e., a parameter step away from the minimum



$$\underline{\Delta p} = -2 \cdot \left[ \frac{\partial^2 PF(\underline{p}_0)}{\partial \underline{p}_0^2} \right]^{-1} \cdot \left[ \frac{\partial PF(\underline{p}_0)}{\partial \underline{p}_0} \right]^T, \quad (4.30)$$

yielding an expression for the Newtonian step to be taken in parameter space toward minimum PF. Both second and first order derivatives occur.

The GN gradient method was chosen, because it has the advantage that it does not require the time-consuming computation of the second order derivatives. Instead, in an approximation of the Hessian matrix (of second order derivatives) it only involves the computation of first order derivatives of PF( $\underline{p}$ ) with respect to  $\underline{p}$ .

Differentiation of Eq.(4.25) yields:

$$grad(-\ln(L)) = \frac{\partial -\ln(L)}{\partial \underline{p}} = \sum_{k=1}^N \left\{ \left[ \frac{\partial \underline{y}}{\partial \underline{p}} \right]^T \cdot Q^{-1} \cdot (\underline{y}_m - \underline{y}) \right\}_{t_k}^T. \quad (4.31)$$

Differentiating once more yields the Hessian:

$$\frac{\partial^2 -\ln(L)}{\partial \underline{p}^2} = - \sum_{k=1}^N \left\{ \left[ \frac{\partial \underline{y}}{\partial \underline{p}} \right]^T \cdot Q^{-1} \cdot \left[ \frac{\partial \underline{y}}{\partial \underline{p}} \right] \right\}_{t_k} + \sum_{k=1}^N \left\{ \left[ \frac{\partial^2 \underline{y}}{\partial \underline{p}^2} \right]^T \cdot Q^{-1} \cdot (\underline{y}_m - \underline{y}) \right\}_{t_k} \quad (4.32)$$

$$\approx - \sum_{k=1}^N \left\{ \left[ \frac{\partial \underline{y}}{\partial \underline{p}} \right]^T \cdot Q^{-1} \cdot \left[ \frac{\partial \underline{y}}{\partial \underline{p}} \right] \right\}_{t_k}, \quad (4.33)$$

neglecting the second order term. Therefore, under the assumption of independent observations, E can be written as:

$$E = \left[ \sum_{k=1}^N \left\{ \left[ \frac{\partial \underline{y}(t_k)}{\partial \underline{p}} \right]^T \cdot Q^{-1} \cdot \left[ \frac{\partial \underline{y}(t_k)}{\partial \underline{p}} \right] \right\} \right]. \quad (4.34)$$

Thus, iteratively the GN parameter step vector  $\underline{\Delta p}$  can be calculated, until it becomes smaller than some preset value,  $\underline{e}$ , according to (for the  $i$ th iteration):



$$\underline{\Delta p}_i = p_{i+1} - p_i = E^{-1} \cdot \{grad(-\ln(L))\}^T = \quad (4.35)$$

$$= \left[ \sum_{k=1}^N \left\{ \left[ \frac{\partial y}{\partial p} \right]^T \cdot Q^{-1} \cdot \left[ \frac{\partial y}{\partial p} \right]_{t_k} \right\} \right]^{-1} \cdot \sum_{k=1}^N \left\{ \left[ \frac{\partial y}{\partial p} \right]^T \cdot Q^{-1} \cdot (y_m - y)_{t_k} \right\}$$

However, at present the iterations will be continued as long as the change in the determinant of the variance-covariance matrix of the residuals,  $|Q|$ , remains larger than 0.1%. If the iterative procedure has converged, the parameter values after the last iteration are considered the true ones (ML estimates), and their accuracies are estimated from the final  $E^{-1}$ .

Sensitivity Equations through Transition Matrices. Sensitivity equations show the sensitivity of the state vector with respect to the parameters. As the GN routine requires evaluation of the derivatives of the model response with respect to the parameters (Eq.(4.35)), and

$$\frac{\partial y}{\partial p} = \frac{\partial H \cdot x}{\partial p} = H \cdot \frac{\partial x}{\partial p}, \quad (4.36)$$

a set of sensitivity equations is drafted and solved simultaneously with the state equations. For each of the  $j=1, \dots, m$  parameters, define the sensitivity vector:

$$\underline{s}_j(t) = \frac{\partial x}{\partial p_j}. \quad (4.37)$$

The sensitivity equations then are found by differentiation of Eq.(4.8):

$$\dot{\underline{s}}_j(t) = \frac{\partial \underline{s}_j(t)}{\partial t} = A \cdot \underline{s}_j(t) + B_j \cdot \underline{x}(t), \quad \text{where } B_j = \frac{\partial A}{\partial p_j}. \quad (4.38)$$

Each element of the matrix  $B_j$  can be found by actually differentiating the corresponding element of the system matrix  $A$  with respect to the parameter  $p_j$ . The initial condition of the sensitivity equations is  $\underline{s}_j(t_0) = \underline{0}$ . It is not difficult to prove that the solutions are:

$$\underline{s}_j(t) = \Phi(t, t_0) \cdot \underline{s}_j(t_0) + \Psi_j(t, t_0) \cdot \underline{x}(t_0), \quad (4.39)$$

where transition matrices  $\Psi_j$  can be found by numerical integration (presently performed for  $t-t_0=0.5$  min by Runge-Kutta method) of:

TABLE 4-5 SENSIVITIES BY FINITE DIFFERENCES (FD)

- subsequently, vary each of m parameters by 1% ⇒  
m + 1 parameter vectors (including the original one);
- calculate corresponding state vectors;  
at time point k the state vector for the j<sup>th</sup> parameter vector is x(k,j);
- write the i<sup>th</sup> element, x<sub>i</sub>(k,j) as a first-order polynomial in p; e.g.,

$$\begin{aligned} x(i,k,1) &= s_0(i,k) + s_1(i,k) \cdot p_1(1) + s_2(i,k) \cdot p_2(1) + \dots + s_m(i,k) \cdot p_m(1) \\ x(i,k,2) &= s_0(i,k) + s_1(i,k) \cdot p_1(2) + s_2(i,k) \cdot p_2(2) + \dots + s_m(i,k) \cdot p_m(2) \\ &\vdots \\ x(i,k,m+1) &= s_0(i,k) + s_1(i,k) \cdot p_1(m+1) + s_2(i,k) \cdot p_2(m+1) + \dots + s_m(i,k) \cdot p_m(m+1) \end{aligned}$$

- subtract the first m equations from the last one:

$$\begin{aligned} \Delta x(i,k,1) &= s_1(i,k) \cdot \Delta p_1(1) + s_2(i,k) \cdot \Delta p_2(1) + \dots + s_m(i,k) \cdot \Delta p_m(1) \\ \Delta x(i,k,2) &= s_1(i,k) \cdot \Delta p_1(2) + s_2(i,k) \cdot \Delta p_2(2) + \dots + s_m(i,k) \cdot \Delta p_m(2) \\ &\vdots \\ \Delta x(i,k,m) &= s_1(i,k) \cdot \Delta p_1(m) + s_2(i,k) \cdot \Delta p_2(m) + \dots + s_m(i,k) \cdot \Delta p_m(m) \end{aligned}$$

as the  $\Delta x$ 's and  $\Delta p$ 's are known, from this set of m equations the m unknown sensitivities  $s_1(i,k) = \partial x_i / \partial p_1$ ,  $s_2(i,k) = \partial x_i / \partial p_2$ , .. ,  $s_m(i,k) = \partial x_i / \partial p_m$  at time point k can be solved algebraically;

similar sets can be drafted for the other time points,  $k = 1, \dots, N$ , as well as for the other state vector elements,  $i = 1, \dots, 2n$

$$\dot{\Psi}_j(t, t_0) = A \cdot \Psi_j(t, t_0) + B_j \cdot \Phi(t, t_0), \quad \Psi_j(t_0, t_0) = 0, \quad (4.40)$$

or by evaluation of the analytical solution:

$$\Psi_j(t, t_0) = B_j \cdot (t - t_0) \cdot \exp(A \cdot (t - t_0)). \quad (4.41)$$

For larger time intervals, by matrix multiplications:

$$\Psi_j((a+b) \cdot t, t_0) = \Psi_j(a \cdot t, t_0) \cdot \Phi(b \cdot t, t_0) + \Phi(a \cdot t, t_0) \cdot \Psi_j(b \cdot t, t_0). \quad (4.42)$$

FD Approach to the Solution of the Sensitivity Equations. The use of transition matrices demands the solving of an extended number of simultaneous differential equations. The finite difference method deals with ordinary algebraic equations and obtains approximate values for the sensitivity vectors in the following way. From the initial m-dimensional parameter vector m new and slightly different

parameter vectors are derived by giving, in sequence, each parameter value a 1% deviation. For each of these  $m+1$  parameter vectors the model response is calculated (Eq.(4.11) or Eq.(4.17)). The state vector  $\underline{x}(\underline{p})$  now is written as a first order polynomial in  $\underline{p}$ , e.g., for the  $i^{\text{th}}$  element of the state vector at the  $k^{\text{th}}$  sample time and the  $l^{\text{th}}$  parameter set:

$$X(i,k,l) = s_0(i,k) + s_1(i,k) \cdot p_1(l) + s_2(i,k) \cdot p_2(l) + \dots + s_m(i,k) \cdot p_m(l), \quad (4.43)$$

in which  $x$  and  $p$  are known and  $s_1$  through  $s_m$  can be regarded as the sensitivities of the state variable in the  $i^{\text{th}}$  compartment at time  $t_k$ :

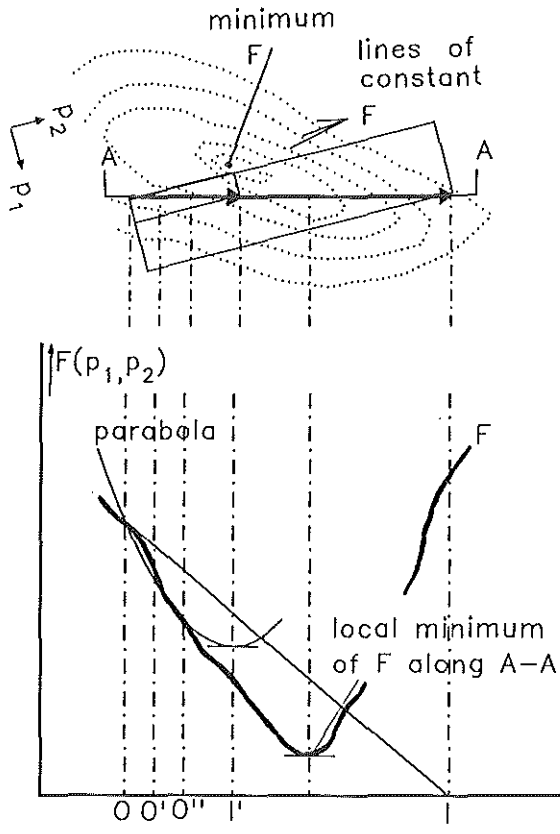
$$s_j(i,k) = \frac{\partial x_i(t_k)}{\partial p_j}. \quad (4.44)$$

Drafting Eq.(4.43) for all  $m+1$  parameter vectors yields a set of equations from which the  $i$ th elements of the  $m$  sensitivity vectors at time  $k$  can be easily calculated after some algebraic manipulations rather than after solving differential equations (Table 4-5). This is repeated for each of the  $i = 1..2n$  elements. Next, the parameter step vector can be computed as in the MGN procedure. For the second iteration only one new model response needs to be calculated, i.e., using the new found parameters. They substitute for the set with the worst model response (largest contribution to  $-\ln(L)$ ) in the previous iteration. Then, again the sensitivities can be computed, etc.

Convergence. An advantage of gradient methods like the GN method is that they can be fast; a disadvantage is that convergence problems may arise, i.e., that successive parameter steps may drift away from the optimum, instead of getting closer to it. This may especially occur when initial estimates are poor, i.e., far from the true parameter values. To diminish the risk of divergence in executing the calculated parameter step the following modifications have been applied.

1.) After the calculation of a parameter step vector (direction and magnitude), instead of taking this step and proceeding with the next iteration a line minimization (parabolic expansion) is performed first. Two small steps of equal size are taken in the original step direction and the function to be minimized,  $-\ln(L)$ , is evaluated. The function value in the starting point also is known. A parabola is drawn through the three points. The location of its minimum determines the magnitude of the parameter step to be actually executed. From this point in parameter space the next iteration is started, etc., until the parameter values no longer change more than 0.1 percent. Figure 4.10 illustrates a two-dimensional case. If necessary, a built-in step limitation procedure reduces again the magnitude of the step, to prevent any of the parameters from becoming negative (the transfer rate constants cannot be less than zero).

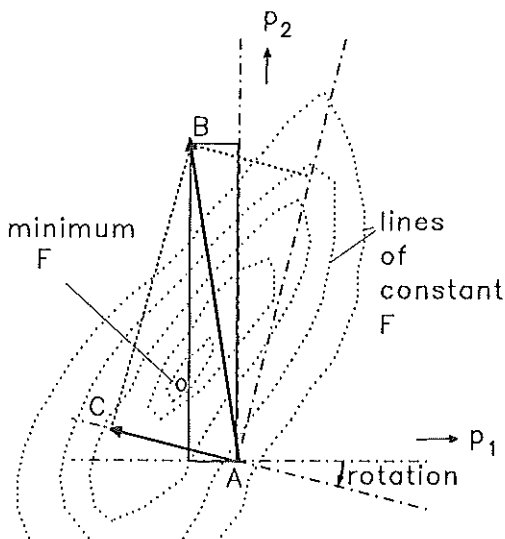
Fig. 4.10 Function  $F$  of two parameters; *Line Minimization* along A-A  
top shows lines of constant  $F$  in the  $p_1$ - $p_2$  plane; bottom shows cross-section A-A. The conventional step O-I (Gauss-Newton gradient) results in divergence; take equal steps  $OO'$  and  $O'O''$  and evaluate the function in  $O$ ,  $O'$  and  $O''$ ; find the minimum of the parabola through these points; take the corresponding step  $O'I'$  which results in convergence



2.) A singular value decomposition (SVD; [Golub and Reinsch, 1970]) of the information matrix  $E$  rotates the  $m$  parameter axes into  $m$  mutually independent eigendirections. Divergence is caused by large steps in the least significant eigendirections. Negligence of these unimportant eigendirections that hardly contribute to the function minimization, thus reducing the dimension of the parameter space from  $m$  to  $m' < m$ , will in general improve convergence. Figure 4.11 shows a two-dimensional example. Computationally, the SVD involves resolving  $E$  as a product of its eigenvalue and eigenvector matrices:

Fig. 4.11 Function  $F$  of two parameters; Minimization after *Singular Value Decomposition* (SVD)

conventional step AB (Gauss-Newton gradient) results in divergence; SVD a) rotates the parameter axes into eigendirections; b) neglects the step into the least important eigendirection (corresponding to the smallest eigenvalue of the information matrix); so, c) takes step AC which results in convergence



$$E = T \cdot Y \cdot T^{-1} = T \cdot Y \cdot T^T, \quad (4.45)$$

where a) eigenvalue matrix  $Y = \text{diag}(e_1, e_2, \dots, e_m)$  with eigenvalues  $e_i$  following from  $|E - e_i \cdot I| = 0$ ; and b)  $\underline{T}_i$  is an eigenvector, from  $E \cdot \underline{T}_i = e_i \cdot \underline{T}_i$ , forming the  $i^{\text{th}}$  column of the eigenvector matrix  $T$ .

Rearranging  $Y$  and  $T$  in order of decreasing  $e_i$ , and subsequently ignoring the  $m-m'$  smallest eigenvalues and their corresponding eigenvectors, yields a reduced and sorted information matrix:

$$E \approx T' \cdot Y' \cdot T'^T \quad (4.46)$$

to be used in Eq.(4.35) for the determination of the direction of the parameter step.

Should the information matrix yield negative eigenvalues, which can be

TABLE 4-6 DIRECT DETERMINATION OF THE GRADIENT, GRAD(F) AND HESSIAN, H(F) OF A FUNCTION F(p) IN p<sub>0</sub>

- Calculate with parameter vector
 

|   |   |                  |
|---|---|------------------|
| p <sub>0</sub> , elements p <sub>i</sub> = p <sub>0i</sub> ,                  | p <sub>i</sub> = p <sub>0j</sub> ⇒ function value | F <sub>0</sub> , |
| p <sub>A</sub> , elements p <sub>i</sub> = p <sub>0i</sub> + d <sub>i</sub> , | p <sub>i</sub> = p <sub>0j</sub> ⇒                | F <sub>A</sub> , |
| p <sub>B</sub> , elements p <sub>i</sub> = p <sub>0i</sub> - d <sub>i</sub> , | p <sub>i</sub> = p <sub>0j</sub> ⇒                | F <sub>B</sub> , |

 (i = 1 or 2 or ... or m; j = 1, ..., m; j ≠ i; small deviations d<sub>i</sub>).
  
- Then, calculate
  - grad(F) = ∂F/∂p, with i<sup>th</sup> element: ∂F(p<sub>0</sub>)/∂p<sub>i</sub> = {(F<sub>A</sub>-F<sub>0</sub>)-(F<sub>B</sub>-F<sub>0</sub>)} / 2d<sub>i</sub>,
  - H(F) = ∂<sup>2</sup>F/∂p<sup>2</sup>, with i<sup>th</sup> diagonal element: H<sub>ii</sub> = ∂<sup>2</sup>F/∂p<sub>i</sub><sup>2</sup> = {(F<sub>A</sub>-F<sub>0</sub>) + (F<sub>B</sub>-F<sub>0</sub>)} / d<sub>i</sub><sup>2</sup>.
  
- Calculate with parameter vector
 

|   |  |                  |
|---|--|------------------|
| p <sub>0</sub> , elements p <sub>i</sub> = p <sub>0i</sub> ,                  | p <sub>i</sub> = p <sub>0j</sub> ⇒ function value                                      | F <sub>0</sub> , |
| p <sub>C</sub> , elements p <sub>i</sub> = p <sub>0i</sub> + d <sub>i</sub> , | p <sub>i</sub> = p <sub>0j</sub> + d <sub>j</sub> , p <sub>k</sub> = p <sub>0k</sub> ⇒ | F <sub>C</sub> , |

 (i = 1 or 2 or ... or m; j = 1 or 2 or ... or m; k = 1, ..., m; j ≠ i; k ≠ i, k ≠ j).
  
- Calculate
  - H(F) = ∂<sup>2</sup>F/∂p<sup>2</sup>, with element on row i and column j (j ≠ i):  

$$H_{ij} = \partial^2 F / \partial p_i \partial p_j = \{2 \cdot (F_C - F_0) \cdot d_i^2 \cdot H_{ii} \cdot d_j^2 \cdot H_{jj}\} / (2 \cdot d_i \cdot d_j)$$
, while H<sub>ii</sub> = H<sub>ij</sub>.

This requires 1 + 2·m + ½·(m<sup>2</sup> - m) evaluations of the function F

shown to correspond to a "saddle-point" situation, then the subsequent line minimization of -ln(L) is performed into the direction of the largest negative eigenvalue, after which the next iteration starts.

DDH Method. Instead of first having to solve a large set of differential equations to find sensitivities ∂x/∂p, the gradient of the log likelihood function (LLF) and its Hessian matrix are calculated for p<sub>0</sub> in a direct way. Again, several parameter vectors are chosen in the neighborhood of p<sub>0</sub>. Starting from the m-dimensional vector p<sub>0</sub>, first, each parameter separately is augmented, and also decreased, by factors 0.01 and 0.02. Then, again starting from p<sub>0</sub>, parameters are varied simultaneously by a factor ±0.01. This yields 1 + 4m + 2<sup>m</sup> different parameter vectors, for which the corresponding values of the LLF are calculated (Eq.(4.11) or Eq.(4.17)). Through these values a second degree surface can be fitted in the least squares sense. According to Newton,

$$LLF(\underline{p}_1) = LLF(\underline{p}_0) + \frac{\partial LLF}{\partial \underline{p}_0^T} \cdot (\underline{p}_1 - \underline{p}_0) + \frac{1}{2} \cdot (\underline{p}_1 - \underline{p}_0)^T \cdot \frac{\partial^2 LLF}{\partial \underline{p}_0^2} \cdot (\underline{p}_1 - \underline{p}_0), \quad (4.47)$$

from which the gradient  $\partial LLF/\partial \underline{p}_0$  and the Hessian  $\partial^2 LLF/\partial \underline{p}_0^2$  can be obtained as LF,  $\underline{p}_0$  and  $\underline{p}_1$  ( $i=1, \dots, 1+4m+2^m$ ) are known (Table 4-6). Next, a parameter step vector can be calculated, etc., as in the MGN method.

Approximation by a first degree surface, hopefully without great loss of information, reduces the number of evaluations of LLF, which otherwise will quickly grow very large with increasing number of parameters. Using the fact that the Hessian is a symmetrical matrix, i.e.,

$$\frac{\partial^2 LLF}{\partial p_a \partial p_b} = \frac{\partial^2 LLF}{\partial p_b \partial p_a}, \quad (4.48)$$

further reduces the number of evaluations to  $1 + 2m + \frac{1}{2} \cdot (m^2 - m)$ .

#### 4.1.4 Error Estimates

**4.1.4.1 Estimation of the Variance-Covariance Matrix Q.** Before every new iteration the variance-covariance matrix of the residuals, Q, is updated, using the sensitivity vectors and parameter step vector from the preceding iteration. Assuming that residuals are mutually independent, the covariances are set to zero ( $\text{cov}\{y_{mi} - y_i, y_{mj} - y_j\} = 0; j \neq i$ ). Then

$$Q = \text{diag}[var_1, var_2, \dots, var_w], \quad \text{where for the } (k+1)^{\text{st}} \text{ iteration:} \quad (4.49)$$

$$var_i|_{k+1} = \frac{1}{N-1} \cdot \sum_{l=1}^N \left\{ (y_{mi} - y_{lk}) - \sum_{j=1}^m \left[ \frac{\partial y_i}{\partial p_j} \cdot \Delta p_j \right]_k \right\}_{l_i}^2, \quad i = 1, \dots, w.$$

It is possible to put extra weight to the observations in some compartment  $i$ , for instance weighing by a factor  $f$  means that  $var_i$  must be set to  $var_i \cdot f^2$ . This has been done sometimes to emphasize the DAU-plasma observations, whose levels are low compared to the tissue concentrations and therefore might be neglected by the optimization algorithm.

**4.1.4.2 Estimation of the Variance-Covariance Matrix  $V_p$ .** The inverted Fisher's information matrix approximates the variance-covariance matrix,  $V_p$ , of the parameters. Thus, after the final iteration step the diagonal elements of the

matrix  $E^{-1}$  (Eq.(4.34)) yield the standard deviations of the parameter values.

**4.1.4.3 Estimation of the Variance-Covariance Matrix  $V_{\underline{x}}$ .** The estimated error in the final model response can be computed from the variance-covariance matrix,  $V_{\underline{x}}$ , of the state vector:

$$V_{\underline{x}}(t) = \left[ \frac{\partial \underline{x}(t)}{\partial \underline{p}} \right] \cdot E^{-1} \cdot \left[ \frac{\partial \underline{x}(t)}{\partial \underline{p}} \right]^T \quad (4.50)$$

Thus, at any time  $t$  the standard deviation of the  $i^{\text{th}}$  element of the state vector results from the  $i^{\text{th}}$  diagonal element of the matrix  $V_{\underline{x}}(t)$ .

#### 4.1.5 Performance Criteria

To compare the performance of the different optimization routines, and to enable the comparison of the results for different model structures, the following criterion variables were selected:

- 1) CPU time needed to go through the iterations;
- 2) the value of the log likelihood function, LLF, after each iteration;
- 3) the distance between the final and initial estimates of the parameter values;
- 4) the goodness of fit of the model response to the observations, expressed in the total correlation coefficient, TCC, defined by:

$$TCC = \left\{ 1 - \frac{\sum_k [y_{m_i}(t_k) - y_i(t_k)]^2}{\sum_k [y_{m_i}(t_k)]^2} \right\}^{1/2}, \quad (4.51)$$

where  $y_m$  and  $y$  denote observed and calculated values, respectively, and the summation is performed over all sample times  $k = 1, \dots, N$  in a compartment. The overall TCC is computed by:

$$TCC_t = \left\{ 1 - \frac{\sum_i \sum_k [y_{m_i}(t_k) - y_i(t_k)]^2}{\sum_i \sum_k [y_{m_i}(t_k)]^2} \right\}^{1/2}, \quad (4.52)$$

adding a summation over  $i = 1, \dots, w$  observed compartments. The TCC is a good measure for comparing the goodness of fit of various curves to a same set of data points. In a case of perfect fit TCC becomes 1, in general TCC will be less than 1.

5) SSR; the sum of the sums—per compartment—of squared residuals divided by their variances. This variable allows comparison of the goodness of fit of



various curves to several sets of data points; the lower SSR, the better the fit.

$$SSR = \sum_i SSR_i; \quad SSR_i = \frac{\sum_k [y_{m_i}(t_k) - y_i(t_k)]^2}{var\{[y_m(t_k) - y(t_k)]_i\}}. \quad (4.53)$$

6) PF; the performance index (Eq.(4.24)) is basically the same as the SSR, and should equal it exactly if the variance-covariance matrix of the residuals would really be a diagonal matrix.

7) AIC; the Akaike Information Criterion [Akaike, 1974] compares various models by considering the goodness of fit (SSR), the number of observations that are used (NR) and the number of parameters (m) to be estimated (model complexity). The lower the AIC value, the better the accuracy—corrected for the degrees of freedom in fitting—of the data representation by the model.

$$AIC = NR \cdot \ln(SSR) + 2 \cdot m. \quad (4.54)$$

8) The magnitude of the variable EC, an error criterion showing to what extent a final parameter value approaches its true value. EC is given by

$$EC = \exp(x), \quad \text{where } x = |\ln(p_{true}/p_{estimated})|. \quad (4.55)$$

This error criterion is meaningful only if the true parameter value is known. As Eq.(4.55) shows, EC will have the value 1 if the estimated parameter value equals the true value. For both positive and negative deviations EC will increase, in both cases to a same extent, i.e., estimating p either too large or too small by a factor of k will have the same effect on the magnitude of EC:  $|\ln(1/k)| = |-\ln(k)| = |\ln(k)|$ .

#### 4.1.6 Computation

With the above considerations a computer program was designed, written in ALGOL60 and implemented to run on an IBM 370/158 main frame computer. This program version was used to identify the large-scale models. With a model comprising 22 compartments, 132 observations at 16 observation times and 45 parameters to be estimated the program requires 3600K of memory and some 150 s computing time per iteration step. The 2-minute transition matrices are obtained by integration. By repetitive multiplications by the current state vector the model response is calculated for every 2 minutes and can be printed and plotted subsequently.

A FORTRAN77 version of the program, suitable for running on a mini-computer system, has also been developed. The system used for testing the



TABLE 4-8 INITIAL PARAMETER SETS FOR THE DATA-SENSITIVITY EVALUATION

| parameters             | initial, set 1    | %true | EC    | initial, set 2    | %true | EC    |
|------------------------|-------------------|-------|-------|-------------------|-------|-------|
| true min <sup>-1</sup> | min <sup>-1</sup> |       |       | min <sup>-1</sup> |       |       |
| 1 .1429E+0             | .1842E+0          | 128.9 | 1.289 | .1242E+0          | 86.9  | 1.151 |
| 2 .7500E-2             | .7020E-2          | 93.6  | 1.068 | .5243E-2          | 69.9  | 1.430 |
| 3 .6814E-2             | .8013E-2          | 117.6 | 1.176 | .6535E-2          | 95.9  | 1.043 |
| 4 .5854E-1             | .7950E-1          | 135.8 | 1.358 | .6433E-1          | 109.9 | 1.099 |
| 5 .1343E+0             | .1229E+0          | 91.5  | 1.093 | .1359E+0          | 101.2 | 1.012 |
| 6 .8180E+1             | .8843E+1          | 108.1 | 1.081 | .7010E+1          | 85.7  | 1.167 |
| 7 .4054E+0             | .3503E+0          | 86.4  | 1.157 | .4038E+0          | 99.6  | 1.004 |
| 8 .1726E+1             | .1498E+1          | 86.8  | 1.152 | .1850E+1          | 107.2 | 1.072 |
| 9 .6000E+0             | .4158E+0          | 69.3  | 1.559 | .6256E+0          | 104.3 | 1.043 |
| 10 .9040E+1            | .9962E+1          | 110.2 | 1.102 | .9299E+1          | 102.9 | 1.029 |

E±a denotes  $\times 10^{\pm a}$ .

The parameters were chosen at random from a Gaussian distribution with mean value  $\mu = 100\%$  times and standard deviation  $\sigma = 15\%$  times true value, using the subroutine GGNML from the IMSL standard FORTRAN F77 library. Both sets are realizations from the same distribution. As random number generator seeds were chosen 15 for set 1, and for set 2: 450. For the meaning of criterion EC, see Eq.(4.55)

#### 4.1.7 Set-up of Test and Evaluation Runs

**4.1.7.1 Data-sensitivity of the Identification Procedure.** It was studied to what extent the results of the identification procedure are sensitive to the number and the location (temporal and spatial) of the measured drug concentrations. The ten transfer rate constants ( $m = 10$ ) in a seven-compartment model (small-scale,  $2n = 7$ , Fig. 4.2a) were estimated using the MGN optimization method on the DG MV/10000. Simulated drug concentration measurements at  $N = 25$  time points during a 48 h period after a pulse drug input into plasma were generated by calculating the linear first order model response for a certain set of transfer rate constants (true parameter values). To account for measurement errors Gaussian noise was added to the model response in these points (Table 4-7).

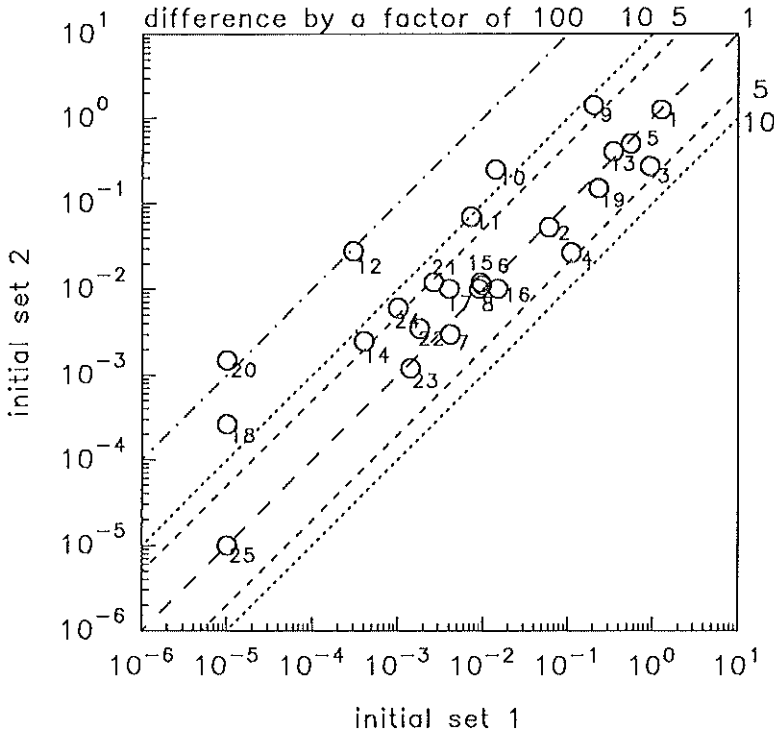
Two different sets of initial parameter values were used to assess how the choice of the starting point influences the results of the optimization routine. These sets were derived by drawing random values from a normal distribution that was characterized by a mean value amounting to 100%, and a standard deviation of 15% of the true parameter value. The random Gaussian deviations were added to the true parameter values. The thus obtained noise corrupted parameter values (Table 4-8) were used as initial estimates in several series of optimization experiments (Table 4-9) that were conducted with varying numbers

TABLE 4-9 SCHEME OF THE SIMULATION EXPERIMENTS (SEE TABLES 4-7 & 4-8)

- T: test, 25 noiseless observations in each of the 7 compartments;  
T1: initial parameters deviate  $\pm 10\%$  from true values;  
T2: initial parameters deviate  $\pm 30\%$  from true values;
- A: 25 observations in each of the 7 compartments;  
A1: initial parameter set 1;  
A2: initial parameter set 2.
- B: 12 observations in each of the 7 compartments  
(points 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24);  
B1: initial parameter set 1;  
B2: initial parameter set 2.
- C: 6 observations in each of the 7 compartments  
(points 1, 4, 10, 13, 22, 25);  
C1: initial parameter set 1;  
C2: initial parameter set 2.
- D: 25 observations in each of the 7 compartments, except tissues\_2;  
D1: initial parameter set 1;  
D2: initial parameter set 2.
- E: initial parameter set 1; 25 observations in plasma and urine.
- F: initial parameter set 1; 25 observations in plasma, urine and bile.
- G: initial parameter set 1; 25 observations in plasma, urine and bile;  
also 3 observations in spleen (points 1, 12, 25).
- H: initial parameter set 1; 9 observations in plasma, urine and bile  
(points 1, 4, 7, 10, 13, 16, 19, 22, 25);  
also 3 observations in spleen (points 1, 12, 25).
- I: initial parameter set 1; 25 observations in all compartments;  
deliberately introduced error in model structure:  
I1: pathway spleen-liver replaced by pathway spleen-plasma;  
I2: pathway plasma-urine replaced by tissues\_1-urine.

of observation times (N) and observed compartments (w). First, the software was tested by using undisturbed observations and initial parameters that deviated only a little from their true values (experiment T). Next, decreasing numbers of observations were allowed in all compartments (experiments A through C). In other cases (experiments D through H) observations were allowed in a few compartments only. In an ultimate case availability of observations was restrict-

Fig. 4.12 Evaluation of minimization routines; comparison of the initial values of the 25 parameters (see Fig. 4.3) in set 1 and set 2



ed to plasma and urine, simulating actual clinical circumstances where patients can be sampled in these compartments only. Knowing the true parameters, the accuracy with which the program determined the optimal parameter values in each case could be compared. This enables establishing of a minimum required amount of observations.

The sensitivity of the estimation procedure with respect to model structure errors (deliberately introduced errors in the distribution pathways) was also investigated (experiment I).

**4.1.7.2 Performance of the Different Minimization Routines.** The medium-scale model ( $2n = 14$  compartments, Fig. 4.3) and a small-scale model ( $2n = 5$  compartments, Fig. 4.2b) were used to evaluate the performance of the different minimization routines. The solutions of the state and sensitivity equations were computed by either direct Runge-Kutta integration or through transition matrices, whose analytical solutions for a 2-minutes interval were cal-

TABLE 4-10

OBSERVED CONCENTRATIONS AND CUMULATIVE AMOUNTS OF DAU AND DOL AFTER I.V. BOLUS INJECTION OF 7.5 mg/kg DAU (MEAN,SD; 4-5 RATS PER DATAPOINT); COMPARTMENT VOLUMES (MEAN OF 40 RATS)

| time<br>h | plasma<br>$\mu\text{g/ml}$ | urine<br>$\mu\text{g}$ | liver<br>$\mu\text{g/g}$ | spleen<br>$\mu\text{g/g}$ | heart<br>$\mu\text{g/g}$ | kidneys<br>$\mu\text{g/g}$ | lungs<br>$\mu\text{g/g}$ | muscles<br>$\mu\text{g/g}$ | bone marrow<br>$\mu\text{g/g}$ |
|-----------|----------------------------|------------------------|--------------------------|---------------------------|--------------------------|----------------------------|--------------------------|----------------------------|--------------------------------|
| 0.167     | 0.78,0.6                   |                        |                          |                           |                          |                            |                          |                            |                                |
| 0.333     | 0.20,0.5                   |                        |                          |                           |                          |                            |                          |                            |                                |
| 0.5       | 0.21,0.5                   |                        | 24.0,0.9                 | 16.8,1.0                  | 18.5,0.7                 | 34.0,0.1                   | 46.5,1.4                 | 3.8, -                     |                                |
| 0.75      | 0.24, -                    |                        |                          |                           |                          |                            |                          |                            |                                |
| 1         | 0.11,0.2                   |                        | 22.0,0.9                 | 17.4,1.0                  | 17.0,0.7                 | 24.0,0.3                   | 46.0,1.4                 | 2.3, -                     | 10.0,2.9                       |
| 1.5       |                            |                        | 11.0,1.0                 | 18.0,1.4                  | 14.0,0.7                 | 23.0,0.2                   | 37.5,2.1                 | 2.7, -                     |                                |
| 2         | 0.05,0.03                  |                        |                          | 16.0,1.0                  |                          |                            |                          | 2.6, -                     | 4.0, -                         |
| 3         |                            |                        | 8.0,0.8                  |                           | 7.1,0.2                  | 11.0,0.8                   | 16.9,0.9                 |                            | 5.0, -                         |
| 4         |                            |                        | 6.5,0.7                  | 15.5,1.0                  | 6.0,0.2                  | 10.0,0.7                   | 18.0,0.9                 |                            |                                |
| 6         | 0.01,0.01                  |                        | 7.7,0.7                  | 14.5,0.7                  | 6.0,0.2                  | 7.2,0.3                    | 12.0,0.7                 | 2.0, -                     |                                |
| 7         |                            | 18.0,5.3               |                          |                           |                          |                            |                          |                            |                                |
| 8         |                            |                        | 7.0,0.6                  | 13.0,0.7                  | 4.5,0.2                  | 6.5,0.2                    | 8.0,0.7                  | 1.0, -                     |                                |
| 9         |                            |                        | 5.0,0.3                  | 12.5,0.7                  | 4.8,0.2                  | 6.0,0.5                    | 9.0,0.7                  | 0.5, -                     |                                |
| 10        |                            | 20.0, -                |                          |                           |                          |                            |                          |                            |                                |
| 15        |                            | 22.0, -                |                          |                           |                          |                            |                          |                            |                                |
| 21        |                            | 23.0,5.1               |                          |                           |                          |                            |                          |                            |                                |
| 24        |                            |                        | 1.5,0.3                  | 7.5,0.5                   | 0.8,0.1                  |                            | 2.8,0.3                  |                            | 4.5,2.9                        |
| 48        | 0.00,0.01                  |                        | 0.7,0.1                  | 2.3,0.3                   | 0.0, -                   | 0.3,0.1                    | 0.4,0.1                  | 0.0, -                     |                                |

TABLE 4-10

CONTINUED

| time<br>h | plasma<br>$\mu\text{g/ml}$ | urine<br>$\mu\text{g}$ | liver<br>$\mu\text{g/g}$ | spleen<br>$\mu\text{g/g}$ | heart<br>$\mu\text{g/g}$ | kidneys<br>$\mu\text{g/g}$ | lungs<br>$\mu\text{g/g}$ | muscles<br>$\mu\text{g/g}$ | bone marrow<br>$\mu\text{g/g}$ |
|-----------|----------------------------|------------------------|--------------------------|---------------------------|--------------------------|----------------------------|--------------------------|----------------------------|--------------------------------|
| 0.167     | 0.54,0.2                   |                        |                          |                           | <u>DOL values</u>        |                            |                          |                            |                                |
| 0.333     | 0.14,0.2                   |                        |                          |                           |                          |                            |                          |                            |                                |
| 0.5       | 0.11,0.1                   |                        | 3.0,0.4                  | 0.9,0.7                   | 4.0,0.3                  | 4.0,0.7                    | 3.5,0.3                  | 0.3, -                     |                                |
| 0.75      | 0.13, -                    |                        |                          |                           |                          |                            |                          |                            |                                |
| 1         | 0.07,0.06                  |                        | 2.2,0.3                  | 3.0,0.7                   | 5.0,0.7                  | 10.0,0.6                   | 4.5,0.7                  | 0.5, -                     | 1.0,0.2                        |
| 1.5       |                            |                        | 2.0,0.1                  | 2.7,0.7                   | 7.5,0.7                  | 9.0,0.1                    | 3.0,0.3                  | 0.3, -                     |                                |
| 2         | 0.04,0.05                  |                        |                          | 3.3,0.5                   |                          |                            |                          | 0.8, -                     | 0.4, -                         |
| 3         |                            |                        | 3.5,0.3                  |                           | 7.6,0.7                  | 8.0,0.5                    | 2.0,0.1                  |                            | 0.6, -                         |
| 4         |                            |                        | 2.3,0.5                  | 2.5,0.3                   | 5.8,0.7                  | 6.0,0.6                    | 4.5,0.3                  | 0.9, -                     |                                |
| 6         | 0.01,0.03                  |                        | 3.0,0.2                  | 5.0,0.5                   | 4.0,0.3                  | 4.8,0.6                    | 2.3,0.1                  | 1.0, -                     | 1.0,0.2                        |
| 7         |                            | 9.5,6.1                |                          |                           |                          |                            |                          |                            |                                |
| 8         |                            |                        | 6.0,0.2                  | 4.0,0.3                   | 4.3,0.3                  | 7.0,0.1                    | 6.0,0.7                  | 1.2, -                     |                                |
| 9         |                            |                        | 5.5,0.2                  | 6.0,0.5                   | 3.0,0.3                  | 6.0,0.4                    | 5.0,0.7                  | 0.6, -                     |                                |
| 10        |                            | 10.5, -                |                          |                           |                          |                            |                          |                            |                                |
| 15        |                            | 15.0, -                |                          |                           |                          |                            |                          |                            |                                |
| 21        |                            | 18.0,8.6               |                          |                           |                          |                            |                          |                            |                                |
| 24        |                            |                        | 3.0,0.6                  | 7.0,0.9                   | 2.2,0.1                  | 4.0,0.2                    | 3.3,0.3                  |                            | 1.7,2.5                        |
| 48        | 0.04,0.01                  |                        | 4.5,0.4                  | 3.3,0.3                   | 0.4,0.1                  | 1.6,0.1                    | 1.5,0.1                  | 0.0, -                     |                                |

compartment volumes (ml):

28.50

4.94

0.32

0.56

1.04

1.38

49.50

1.08

other\_tissues: 33.70; nb. plasma volume = plasma (5.70) + extracellular body water (22.80)

TABLE 4-11 MORE TESTED MODELS FOR DAU-DOL PHARMACOKINETICS

| <u>model</u> | <u>n</u> | <u>w</u> | <u>m</u> | <u>NR</u> | <u>method</u> | <u>remarks</u>  |
|--------------|----------|----------|----------|-----------|---------------|---|
| 3a1.1        | 2x3      | 2x2      | 9        | 16        | seq.          | excretion = urine;                                      |
| 3a1.2        | 2x3      | 2x2      | 9        | 16        | seq.          | as 3a1.1, other initial parameter values;               |
| 3b1          | 2x3      | 2x2      | 9        | 22        | sim.          | as 3a1.1, plasma DAU weighted 100x, plasma volume free; |
| 3a2          | 2x3      | 2x2      | 9        | 20        | seq.          | excretion = urine + bile;                               |
| 3b2          | 2x3      | 2x2      | 8        | 22        | sim.          | as 3a2;   |
| 3b3          | 2x3      | 2x2      | 9        | 22        | sim.          | as 3a2, plasma DAU weighted 100x, plasma volume free;   |
| 5b1          | 2x5      | 2x3      | 15       | 32        | sim.          | metabolism in plasma and tissues                        |
| 5b2          | 2x5      | 2x3      | 14       | 32        | sim.          | as 5b1, metabolism in tissues only;                     |
| 5b3          | 2x5      | 2x3      | 15       | 32        | sim.          | as 5b1, modified algorithm;                             |
| 6b1          | 2x6      | 2x5      | 19       | 76        | sim.          | metabolism in plasma, liver, tissues_1;                 |
| 7b1          | 2x7      | 2x6      | 24       | 94        | sim.          | metabolism in plasma, liver, spleen, tissues_1;         |

n = number of compartments; w = number of observed compartments; m = number of parameters; NR = number of observations; seq. = sequential fitting (analytical method); sim. = simultaneous solution (numerical method);

tissues\_1 = well perfused tissues: liver, spleen, heart, kidneys, lungs and skeletal muscles; tissues\_2 = poorly perfused tissues: bone marrow and other tissues that can be reached by the drug, but for which no observations are available. Tissues that cannot be reached by the drug, for example brain tissue and bones, are excluded from the models

culated\* (MGN method). In parallel computation runs, instead of solving the sensitivities, the gradient and Hessian matrix of the log likelihood function—necessary to establish the direction of the step in parameter space toward the minimum—were approximated in a direct way by means of finite differences (DDH method). A third approach involves the solving of the sensitivities by finite differences instead of transition matrices (FD method). The various performance criteria resulting from the different runs were compared.

The numerical minimization procedures are always started with the same two sets (set 1 and 2) of initial parameter values (Fig. 4.12). Most of the corresponding values in these sets differ less than a factor of 5 (68%), some differ between a factor of five and a factor of 10 (12%), the remaining ones (20%) differ more than a factor of 10 (up to a factor of over 100).

Observations used are derived from actually observed DAU and DOL concentrations. Considering the relative organ volumes, measurements in several organs were pooled into observations for the composite tissues\_1 compartment (Table 4-10).



The computations were performed on the GOULD minicomputer system.

#### **4.1.7.3 Identification of the System of DAU-DOL Pharmacokinetics.**

Various structures of the large-scale model ( $2n = 22$ , Fig 4.4) were tested for the best fit of the observed DAU-DOL concentration—time data (Table 4-10; at that time no bile observations were available yet). These six structures are listed in Table 4-1. The MGN minimization method was used, finding transition matrices for a 2 minutes time interval by numerical integration. For the large-scale model computations were performed on the IBM main frame, using the ALGOL source code.

In addition, a few smaller-scale models were evaluated as well (Table 4-11). The same program was used, this time in FORTRAN F77 on the DG MV/10000. The 2x3 compartment model was evaluated both numerically and analytically (see Appendix E).

## **4.2 RESULTS AND DISCUSSION**

### **4.2.1 Experimental Data**

The observed quantities of DAU and DOL in the organs and body fluids of the rat after the *i.v.* administration of a dose of DAU, as well as the measured organ volumes, are listed in Table 4-12. The observed standard deviations are rather high, probably due to the biological variation in the test animals. In general however, similar patterns are seen in the concentration—time histories, both with respect to the order of magnitude of the data and the qualitative behavior.

**Plasma, DAU and DOL.** A very rapid decrease to low concentration levels is seen after almost instantaneous heavy loading of the compartments (which is remarkable for DOL especially as it seems to suggest a very fast metabolism process). After two hours the low levels have been reached and a very slow further decrease follows.

**Urine/Bile, DAU and DOL.** In these compartments the compounds are accumulated. In urine similar amounts of DAU and DOL are excreted eventually, while DAU is being accumulated somewhat faster in an early stage. For bile quantitative data became available at a later stage. Large interindividual differences resulted in large standard deviations for the observed concentrations. Just like with urine, it was seen that almost equal amounts of DAU and DOL were excreted into the bile, with a little faster rate of accumulation of DAU during the first few hours.

TABLE 4-12 OBSERVED CONCENTRATIONS AND CUMULATIVE AMOUNTS OF DAU AND DOL AFTER I.V. BOLUS INJECTION OF 7.5 mg/kg DAU (MEAN,SD; 4-5 RATS PER DATAPOINT); COMPARTMENT VOLUMES (MEAN OF 40 RATS)

| time<br>h | plasma<br>μg/ml | urine<br>μg | liver<br>μg/g     | spleen<br>μg/g | tissues_1<br>μg/g | bile<br>μg |
|-----------|-----------------|-------------|-------------------|----------------|-------------------|------------|
|           |                 |             | <u>DAU values</u> |                |                   |            |
| 0.167     | 1.07,0.6        | - , -       | - , -             | - , -          | - , -             | - , -      |
| 0.333     | 1.18,0.5        | - , -       | - , -             | - , -          | - , -             | - , -      |
| 0.5       | 1.25,0.5        | - , -       | 24.0,0.9          | 16.8,1.0       | 5.7,0.8           | - , -      |
| 0.75      | - , -           | - , -       | - , -             | - , -          | - , -             | - , -      |
| 1         | 0.54,0.2        | - , -       | 22.0,0.9          | 17.4,1.0       | 2.9,0.8           | 55.0,49.5  |
| 1.5       | - , -           | - , -       | 11.0,1.0          | 18.0,1.4       | 4.1,0.8           | - , -      |
| 2         | 0.14,0.03       | - , -       | - , -             | 16.0,1.0       | - , -             | 56.4,56.1  |
| 2.5       | - , -           | - , -       | - , -             | - , -          | - , -             | - , -      |
| 3         | 0.09,0.03       | - , -       | 8.0,0.8           | - , -          | - , -             | - , -      |
| 3.5       | - , -           | - , -       | - , -             | - , -          | - , -             | - , -      |
| 4         | - , -           | - , -       | 6.5,0.7           | 15.5,1.0       | - , -             | 90.7,59.8  |
| 5         | - , -           | - , -       | - , -             | - , -          | - , -             | - , -      |
| 6         | 0.06,0.01       | 22.1,5.3    | 7.7,0.7           | 14.5,0.7       | 2.4,0.8           | 97.2,67.7  |
| 7         | - , -           | - , -       | - , -             | - , -          | - , -             | - , -      |
| 8         | - , -           | - , -       | 7.0,0.6           | 13.0,0.7       | 1.3,0.8           | 105.7,69.0 |
| 9         | - , -           | - , -       | 5.0,0.3           | 12.5,0.7       | 0.9,0.8           | - , -      |
| 10        | - , -           | - , -       | - , -             | - , -          | - , -             | 109.7,67.8 |
| 12        | - , -           | - , -       | - , -             | - , -          | - , -             | 111.8,67.8 |
| 14        | - , -           | - , -       | - , -             | - , -          | - , -             | 113.4,67.3 |
| 15        | - , -           | - , -       | - , -             | - , -          | - , -             | - , -      |
| 16        | - , -           | - , -       | - , -             | - , -          | - , -             | 115.1,66.6 |
| 18        | - , -           | - , -       | - , -             | - , -          | - , -             | 117.0,65.8 |
| 20        | - , -           | - , -       | - , -             | - , -          | - , -             | 119.6,64.4 |
| 21        | - , -           | - , -       | - , -             | - , -          | - , -             | - , -      |
| 22        | - , -           | - , -       | - , -             | - , -          | - , -             | 133.2,72.6 |
| 24        | - , -           | 35.5,5.1    | 1.5,0.3           | 7.5,0.5        | - , -             | 134.1,73.0 |
| 28        | - , -           | - , -       | - , -             | - , -          | - , -             | - , -      |
| 32        | - , -           | - , -       | - , -             | - , -          | - , -             | - , -      |
| 48        | 0.00,0.01       | - , -       | 0.7,0.1           | 2.3,0.3        | 0.02,0.8          | - , -      |

TABLE 4-12

CONTINUED

| time<br>h                 | plasma<br>$\mu\text{g/ml}$ | urine<br>$\mu\text{g}$ | liver<br>$\mu\text{g/g}$ | spleen<br>$\mu\text{g/g}$ | tissues_1<br>$\mu\text{g/g}$ | bile<br>$\mu\text{g}$ |                      |
|---------------------------|----------------------------|------------------------|--------------------------|---------------------------|------------------------------|-----------------------|----------------------|
| 0.167                     | 1.10,0.2                   | - , -                  | - , -                    | - , -                     | - , -                        | - , -                 | n.b. plasma volume = |
| 0.333                     | 0.81,0.2                   | - , -                  | - , -                    | - , -                     | - , -                        | - , -                 | plasma (5.70) +      |
| 0.5                       | 0.71,0.1                   | - , -                  | 3.0,0.4                  | 0.9,0.7                   | 0.49,0.08                    | - , -                 | extracellular body   |
| 0.75                      | - , -                      | - , -                  | - , -                    | - , -                     | - , -                        | - , -                 | water (22.80)        |
| 1                         | 0.39,0.06                  | - , -                  | 2.2,0.3                  | 3.0,0.7                   | 0.84,0.08                    | 22.0,29.4             |                      |
| 1.5                       | - , -                      | - , -                  | 2.0,0.1                  | 2.7,0.7                   | 0.62,0.08                    | - , -                 | tissues_1 data by    |
| 2                         | 0.11,0.05                  | - , -                  | - , -                    | 3.3,0.5                   | - , -                        | 49.4,45.0             |                      |
| 2.5                       | - , -                      | - , -                  | - , -                    | - , -                     | - , -                        | - , -                 | pooling heart,lungs, |
| 3                         | 0.11,0.05                  | - , -                  | 3.5,0.3                  | - , -                     | - , -                        | - , -                 | kidneys and muscles  |
| 3.5                       | - , -                      | - , -                  | - , -                    | - , -                     | - , -                        | - , -                 | of Table 4-10        |
| 4                         | - , -                      | - , -                  | 2.3,0.5                  | 2.5,0.3                   | 1.15,0.08                    | 57.4,39.5             |                      |
| 5                         | - , -                      | - , -                  | - , -                    | - , -                     | - , -                        | - , -                 |                      |
| 6                         | 0.09,0.03                  | 25.6,6.1               | 3.0,0.2                  | 5.0,0.5                   | 1.14,0.08                    | 88.3,50.4             |                      |
| 7                         | - , -                      | - , -                  | - , -                    | - , -                     | - , -                        | - , -                 |                      |
| 8                         | - , -                      | - , -                  | 6.0,0.2                  | 4.0,0.3                   | 1.47,0.08                    | 103.3,54.9            |                      |
| 9                         | - , -                      | - , -                  | 5.5,0.2                  | 6.0,0.5                   | 0.85,0.08                    | - , -                 |                      |
| 10                        | - , -                      | - , -                  | - , -                    | - , -                     | - , -                        | 113.4,58.2            |                      |
| 12                        | - , -                      | - , -                  | - , -                    | - , -                     | - , -                        | 121.6,61.2            |                      |
| 14                        | - , -                      | - , -                  | - , -                    | - , -                     | - , -                        | 127.6,62.3            |                      |
| 15                        | - , -                      | - , -                  | - , -                    | - , -                     | - , -                        | - , -                 |                      |
| 16                        | - , -                      | - , -                  | - , -                    | - , -                     | - , -                        | 133.7,62.1            |                      |
| 18                        | - , -                      | - , -                  | - , -                    | - , -                     | - , -                        | 139.4,63.0            |                      |
| 20                        | - , -                      | - , -                  | - , -                    | - , -                     | - , -                        | 146.4,63.0            |                      |
| 21                        | - , -                      | - , -                  | - , -                    | - , -                     | - , -                        | - , -                 |                      |
| 22                        | - , -                      | - , -                  | - , -                    | - , -                     | - , -                        | 146.4,70.0            |                      |
| 24                        | - , -                      | 49.6,8.6               | 3.0,0.6                  | 7.0,0.9                   | - , -                        | 143.8,59.0            |                      |
| 28                        | - , -                      | - , -                  | - , -                    | - , -                     | - , -                        | - , -                 |                      |
| 32                        | - , -                      | - , -                  | - , -                    | - , -                     | - , -                        | - , -                 |                      |
| 48                        | 0.04,0.01                  | - , -                  | 4.5,0.4                  | 3.3,0.3                   | 0.08,0.08                    | - , -                 |                      |
| compartment volumes (ml): |                            |                        |                          |                           |                              |                       |                      |
|                           | 28.50                      |                        | 4.94                     | 0.32                      | 52.48                        |                       | (tissues_2: 34.78);  |

**Organs, DAU.** Almost all organs take up DAU very fast; the peak concentration is reached within one hour. Then, elimination occurs, rapidly at first and slowly later when low concentration levels have been attained. In spleen and bone marrow both DAU uptake and elimination appear to take place at a lower rate, and the 24 h concentrations in these organs still are relatively high. The muscles compartment also is an exception as the DAU concentrations remain comparatively low overall.

**Organs, DOL.** For the metabolite the maximum attained concentration levels are much lower than for the parent drug. As can be expected considering that time is needed for metabolite formation and redistribution processes, the DOL concentrations are built up at a slower pace (peak concentrations well past one hour). The rate of elimination appears to be lower as well, for considerable amounts of DOL are still present after 24 h (compared to parent drug). In spleen and bone marrow the accumulation of DOL exceeds the elimination and concentrations are still rising after 24 h.

From the experimental observations already two conclusions can be drawn with clinical relevance. The plasma concentrations of DAU and DOL do not reflect, at least not in a simple way, the compounds' levels in the organs; while plasma concentrations are already very low the levels in other organs may be fast decaying or even be rising still. The second conclusion is that large interindividual differences may be encountered at a same sample time.

#### 4.2.2 Data-Sensitivity

**4.2.2.1 Test of the Optimization Routine.** A test was conducted to check whether the MGN optimization routine performed well. Two initial parameter vectors were chosen to start with, one with near-true values (10% deviation; case 1), the other one a poorer estimate ( $\pm 30\%$  deviation; case 2). In this test all observations used were noiseless. The routine returned final parameter values that yielded very good fits of the model response to the observations, TCC being equal to one in all compartments but plasma—for case 2—for which TCC = 0.92 was found. The log likelihood function was reduced by a factor of  $10^4$  and  $10^3$  for case 1 and 2, respectively, showing that the optimization routine performs better when the initial parameter vector is nearer to the true one. Yet, even with the very good fits resulting, the final parameter values still showed inaccuracy. For case 1 only 20% of the final parameter were more accurate than 5%, 70% had an accuracy better than 10%, and all were more accurate than 12%. For case 2 most final parameter values showed an accuracy of not more than 20-30%. *This demonstrates that even under optimum conditions, with a near perfect fit achieved, most transfer rate constants cannot be estimated but*

Fig. 4.13 Data-sensitivity experiments; The percentage of parameters for which the error criterion EC (Eq.(4.55)) has a value smaller than indicated; mean initial (○) and final (□) parameter vectors compared (all compartments: 100% observations)

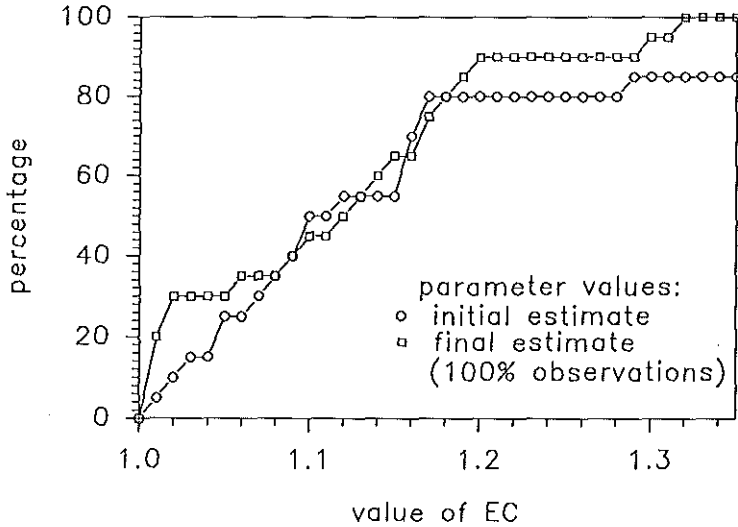


Fig. 4.14 Data-sensitivity experiments; The percentage of parameters for which the error criterion EC (Eq.(4.55)) has a value smaller than indicated; comparison of final parameter vectors (all compartments: ○ 100%, □ 50% or ◇ 25% observations)

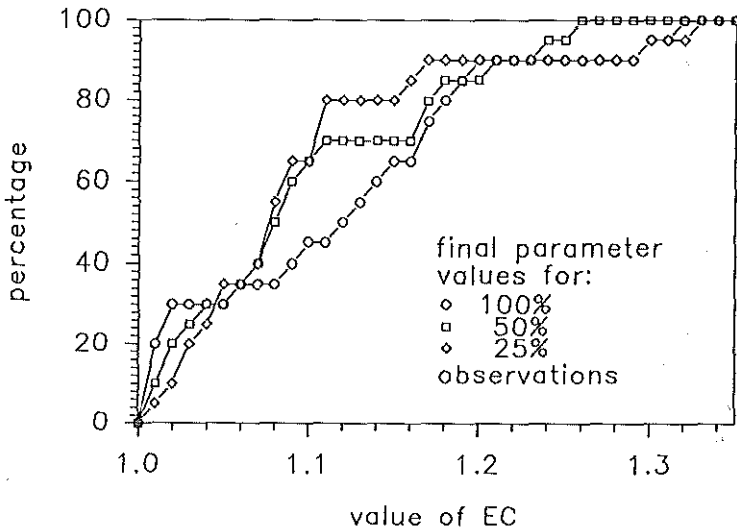
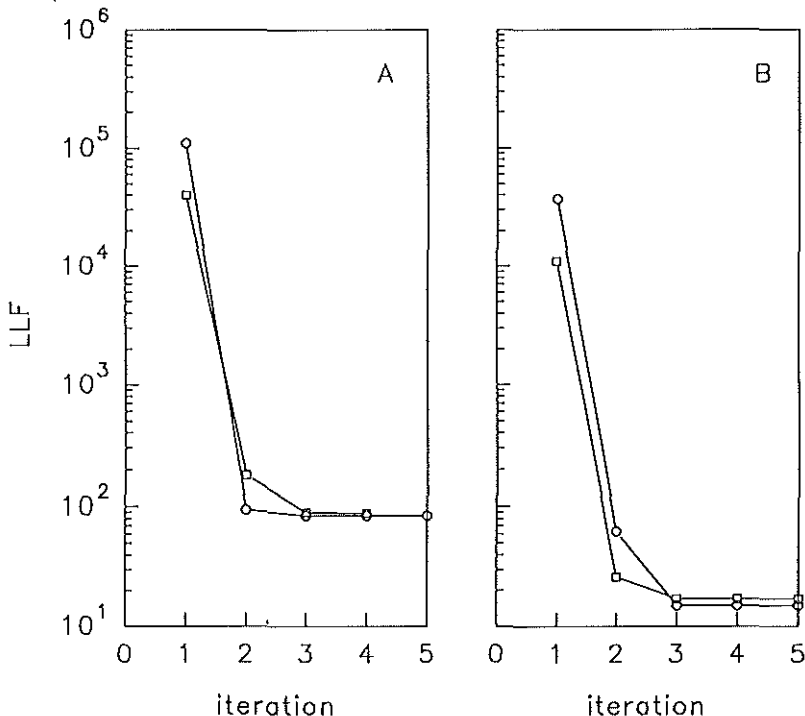


Fig. 4.15 Data-sensitivity experiments; The decrease in log likelihood function with the number of iteration steps, starting with parameter sets 1 ( $\square$ ) or 2 ( $\circ$ ) (see Table 4-8); all compartments observed, a: 100%, b: 25% observations



with at least 5-10% inaccuracy.

**4.2.2.2 Experiments A through C; reduction of the number of data-points while all compartments are observed.** Figures 4.13 through 4.15 show for the various experiments how the estimated parameters are distributed with respect to corresponding values of the error criterion EC (Eq.(4.55)). Figure 4.13 compares the mean initial parameter vector to the final one for 100% observations. In the latter case all parameter values have an EC-value smaller than 1.32 (against 85% in the former case), and a third has an EC-value smaller than 1.02 (against 10% in the former case). With the final parameter values the fit of the model response to the datapoints is much improved (overall TCC going from 0.96052 to 0.99997) and—within a few iteration steps—the log likelihood function is reduced by a factor of  $10^3$  (Fig. 4.15a). Also, the remaining variance of the residuals arrives at values that are near the ones expected on the bases of the Gaussian distributions of the added noise. *This shows that also in the case of noise polluted observations the optimization routine returns*

TABLE 4-13 LOG LIKELIHOOD FUNCTION, GOODNESS OF FIT AND RESIDUAL VARIANCE FOR THE DATA-SENSITIVITY EXPERIMENTS

For each of the experiments (Table 4-9) the model response associated with the final parameter values was calculated and compared with *all* observations in *all* compartments. Thus, not necessarily all observations were always used to *estimate* these final parameters!

The expected residual variances, based on the Gaussian distributions of the generated random noise, are .25E-2, .40E-1, .25E+0, .10E+1, .25E+0, .64E+0 and .36E+0 for, respectively, plasma, tissues\_1, liver, bile, tissues\_2, urine and spleen.

| EXPERIMENT     | A1      | A2      | B1      | B2      | C1      | C2      |
|----------------|---------|---------|---------|---------|---------|---------|
| LLF            | 89.29   | 85.11   | 85.91   | 80.30   | 87.31   | 82.75   |
| TCC            |         |         |         |         |         |         |
| <i>overall</i> | .999969 | .999971 | .999970 | .999972 | .999970 | .999972 |
| plasma         | .990703 | .990632 | .990869 | .990219 | .989327 | .989430 |
| tissues_1      | .998959 | .998938 | .998951 | .998880 | .998940 | .998931 |
| liver          | .997261 | .996562 | .997233 | .997071 | .997047 | .997029 |
| bile           | .999987 | .999987 | .999986 | .999988 | .999987 | .999988 |
| tissues_2      | .999634 | .999635 | .999645 | .999670 | .999604 | .999626 |
| urine          | .999921 | .999924 | .999920 | .999923 | .999923 | .999907 |
| spleen         | .998606 | .998862 | .998952 | .998919 | .999122 | .999120 |

Residual Variance

|           |          |          |          |          |          |          |
|-----------|----------|----------|----------|----------|----------|----------|
| plasma    | .3006E-2 | .3029E-2 | .2952E-2 | .3161E-2 | .3448E-2 | .3415E-2 |
| tissues_1 | .4435E-1 | .4525E-1 | .4470E-1 | .4772E-1 | .4515E-1 | .4551E-1 |
| liver     | .2834E+0 | .3555E+0 | .2863E+0 | .3030E+0 | .3055E+0 | .3073E+0 |
| bile      | .1382E+1 | .1309E+1 | .1426E+1 | .1216E+1 | .1409E+1 | .1219E+1 |
| tissues_2 | .3416E+0 | .3402E+0 | .3311E+0 | .3082E+0 | .3700E+0 | .3489E+0 |
| urine     | .7935E+0 | .7608E+0 | .8016E+0 | .7728E+0 | .9036E+0 | .9287E+0 |
| spleen    | .7235E+0 | .5907E+0 | .5438E+0 | .5611E+0 | .4558E+0 | .4565E+0 |

| EXPERIMENT     | D1      | D2      | E        | F         | G       | H       |
|----------------|---------|---------|----------|-----------|---------|---------|
| LLF            | 84.41   | 79.68   | 27894.-- | 113150.-- | 451.55  | 616.28  |
| TCC            |         |         |          |           |         |         |
| <i>overall</i> | .999970 | .999973 | .990363  | .960309   | .999845 | .999972 |
| plasma         | .990359 | .989974 | .990906  | .990921   | .991029 | .989430 |
| tissues_1      | .998890 | .998827 | .959249  | .938765   | .907298 | .998931 |
| liver          | .997072 | .996989 | .985678  | .943733   | .968042 | .997029 |
| bile           | .999987 | .999988 | .999575  | .999990   | .999959 | .999988 |
| tissues_2      | .999669 | .999679 | .998359  | .989559   | .990625 | .999626 |
| urine          | .999921 | .999924 | .999932  | .999926   | .999927 | .999907 |
| spleen         | .998957 | .999042 | .000000  | .000000   | .999064 | .999120 |

Residual Variance

|           |          |          |          |          |          |          |
|-----------|----------|----------|----------|----------|----------|----------|
| plasma    | .3117E-2 | .3240E-2 | .2940E-2 | .2936E-2 | .2901E-2 | .3232E-2 |
| tissues_1 | .4728E-1 | .4995E-1 | .1701E+1 | .2530E+1 | .3767E+1 | .4768E+1 |
| liver     | .3029E+0 | .3114E+0 | .1473E+1 | .5665E+1 | .3258E+1 | .5789E+1 |
| bile      | .1384E+1 | .1259E+1 | .4452E+2 | .1032E+1 | .1302E+1 | .1204E+1 |
| tissues_2 | .3094E+0 | .2995E+0 | .1531E+1 | .9696E+1 | .8711E+1 | .1152E+2 |
| urine     | .7884E+0 | .7671E+0 | .6851E+0 | .7450E+0 | .7351E+0 | .8814E+0 |
| spleen    | .5411E+0 | .4972E+0 | .1066E+4 | .4506E+4 | .4859E+0 | .4861E+0 |

Fig. 4.16 Data-sensitivity experiments;

Model response with estimated  $\pm 1$  standard deviation and observations with their accuracies ( $\pm 1$  standard deviation) in plasma and liver. All compartments: 100% observations. Performance criteria in boxes: left value: with respect to observations used for parameter estimation; right value: with respect to all observations (true values between brackets)

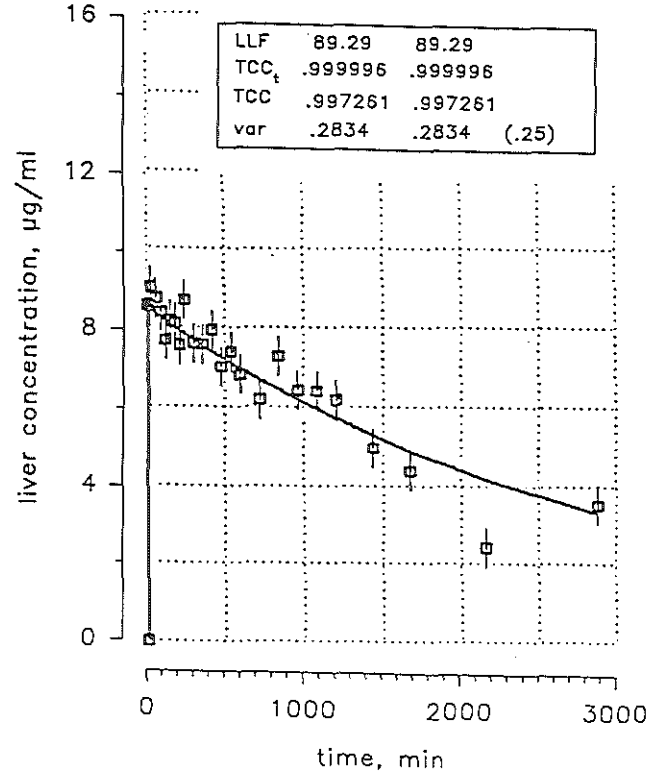
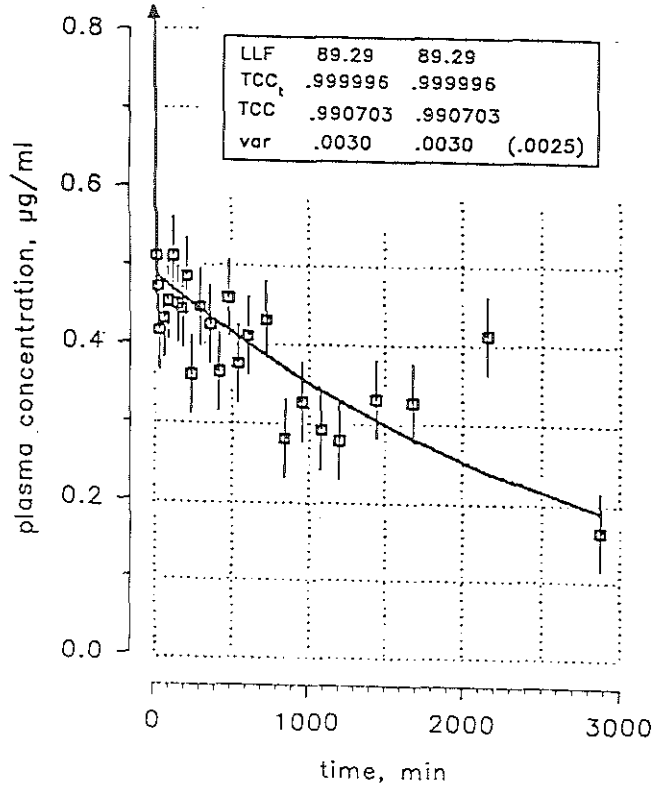
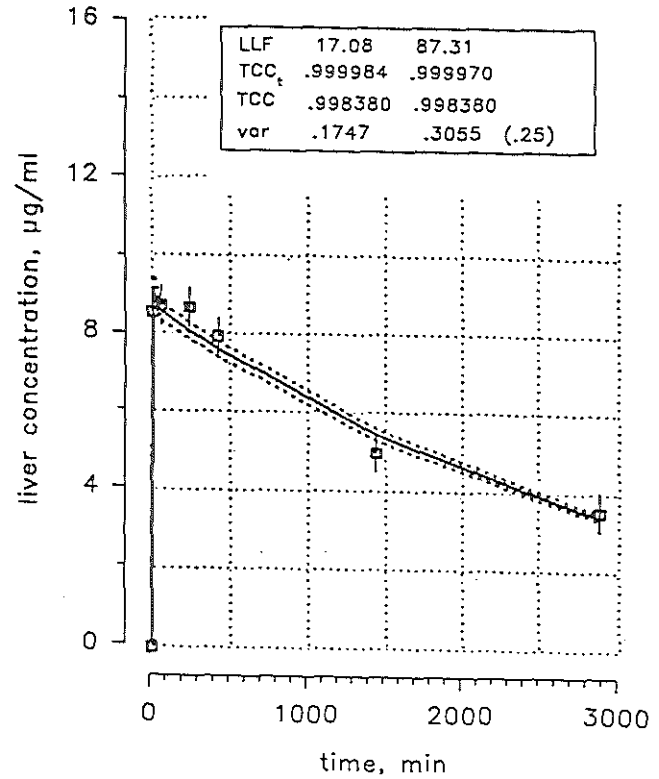
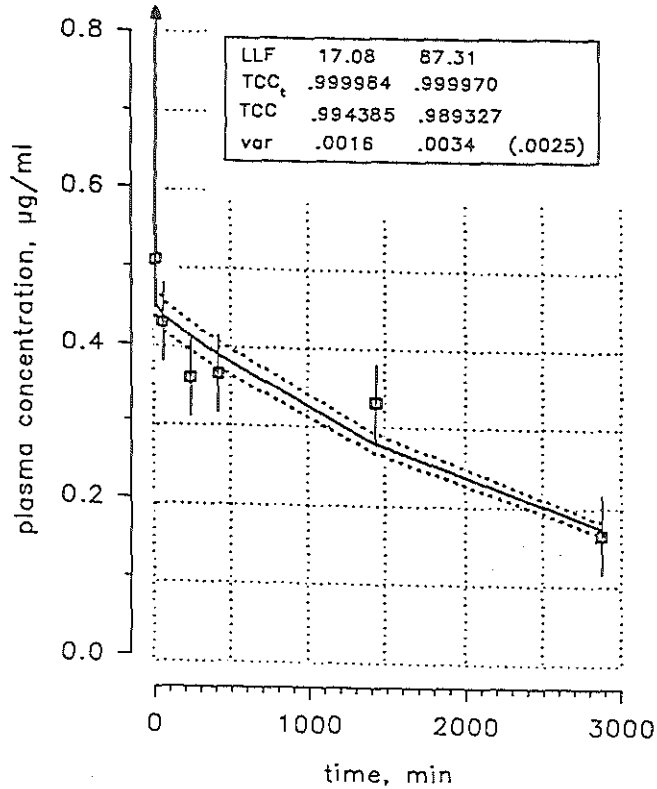




Fig. 4.17 Data-sensitivity experiments;

Model response with estimated  $\pm 1$  standard deviation and observations with their accuracies ( $\pm 1$  standard deviation) in plasma and liver. All compartments: 25% observations. Performance criteria in boxes: left value: with respect to observations used for parameter estimation; right value: with respect to all observations (true values between brackets)



*improved values with respect to initial estimates.* Figure 4.14 shows that reducing the number of observations to one half and even one fourth does not dramatically change the accuracy of the results. Always 100% of the final parameters have an EC-value of less than 1.32 and, in fact, with the reduced numbers of observations the number of parameters with EC-values less than 1.10 amounts to 65%, whereas in the case of all observations 45% is found. However, the number of very accurately estimated parameters decreases with decreasing observations. The number of parameters with EC-value less than 1.01 drops from 20% to 10% to %5 for the cases of 100%, 50% and 25% observations, respectively. A 1000-fold reduction in the log likelihood function still is attained when the number of observations decreases to a quart (Fig. 4.15).

Comparing the model responses from all experiments A through C to all datapoints—so, for B and C, also to datapoints not used for estimating the transfer rates—it can be seen (Table 4-13) that similar values are always obtained for the log likelihood function and the goodness of fit. The ultimate transfer rate constants returned depend to some extent on what initial parameter vector the optimization routine is started with, but the accuracy of the results is such that differences are not significant. *Therefore, it can be concluded that, as long as all compartments are observed, the number of observations per compartment necessary to estimate the transfer rate constants can safely be lowered from 25 to 6, without serious loss of accuracy. Other experiments, not reported here, seem to suggest that especially early data-points (0-10 h) can be left out without great effect.*

Final responses—after starting the optimization from parameter set 1—in, for example, plasma and liver are shown in Fig. 4.16 (100% observations) and Fig. 4.17 (25% observations). They are only a little bit poorer in the latter case, and so are the estimated response accuracies.

**4.2.2.3 Experiments D through H; datapoints in a few compartments only.** Leaving out the observations in one single compartment (tissues\_2) results in the unexplained phenomenon that the accuracy of the final parameter estimates improves (case D, Table 4-13; Fig. 4.18).

Allowing observations in plasma and urine only—all 25 in each—it appears that 40% of the parameters cannot be estimated with good accuracy (EC-value larger than 1.35). Only 10% has an EC-value better than 1.01 (Fig. 4.19). Adding all bile observations results in little improvement (30% of the parameters with EC-values larger than 1.35, zero with EC-value better than 1.01). Some improvement is seen if yet another compartment is at least partially observed, e.g., adding as few as 3 observations in spleen leaves only 10% of the parameters with an EC-value larger than 1.35. Then, in fact the number of observations in bile can be reduced again, from 25 to 9, at only small costs with

Fig. 4.18 Data-sensitivity experiments; The percentage of parameters for which the error criterion EC (Eq.(4.55)) has a value smaller than indicated; comparison of final parameter vectors ( $\square$ ) all compartments: 100% observations, versus ( $\circ$ ) all compartments: 100%, except tissues\_2: 0%

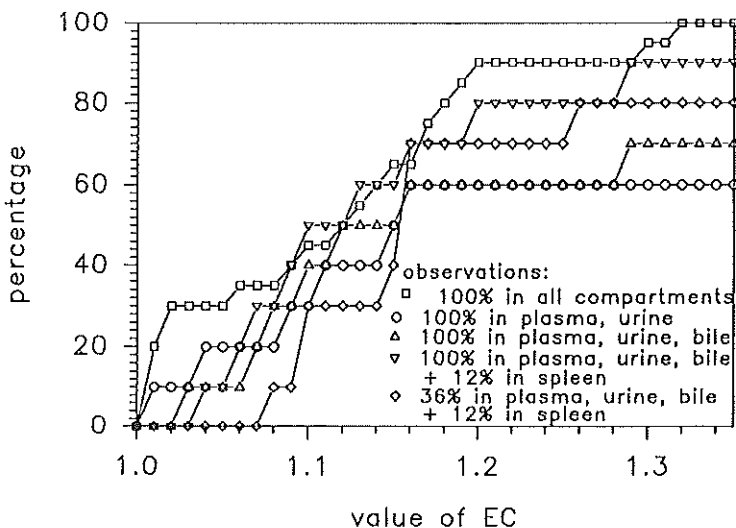
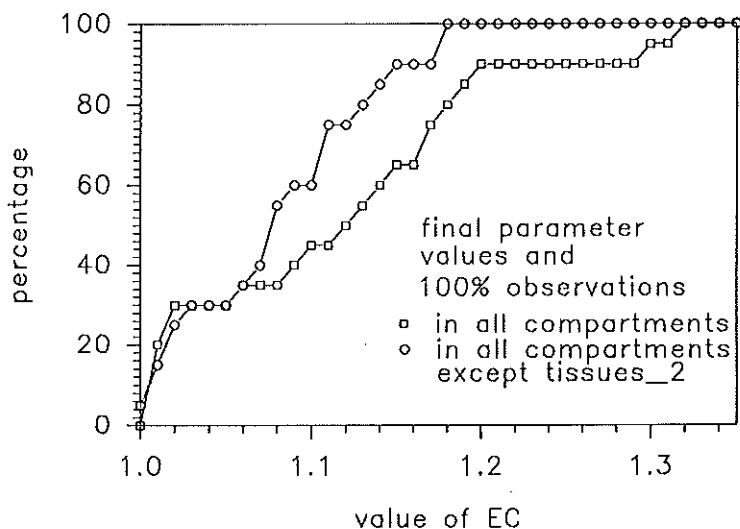


Fig. 4.19 Data-sensitivity experiments; The percentage of parameters for which the error criterion EC (Eq.(4.55)) has a value smaller than indicated; comparison of final parameter vectors ( $\square$ ) all compartments: 100% observations, versus, respectively, ( $\circ$ ) plasma and urine: 100%, other: 0% ( $\triangle$ ) plasma, urine and bile: 100%, other: 0% ( $\nabla$ ) plasma, urine and bile: 100%, spleen: 12%, other: 0% ( $\diamond$ ) plasma, urine and bile: 36%, spleen: 12%, other: 0%

respect to accuracy (20% of the parameters then have an EC-value larger than 1.35).

Comparing the model responses in Figs. 4.16, 4.20 and 4.22, respectively, all compartments 100% observed, only plasma and urine 100% observed, and plasma, urine, bile and spleen partially observed, it is seen that the plasma goodness of fit is not much influenced, while the liver response fluctuates considerably. Without at least a few organ observations the spleen response is very wrongly estimated (Figs. 4.21 and 4.23).

*Thus, these experiments show that it will be necessary to observe more compartments than only those that in practice can be easily accessed without sacrificing the laboratory animal. However, the number of samples in such inaccessible organs needs not be very high.*

**4.2.2.4 Experiments I and A; the influence of model errors.** As can be seen from Table 4-14, when three models with slightly different pathway structures are matched to the same datapoints the best results are obtained with the wrong model (Ia; lowest value of the log likelihood function, highest overall TCC). Only if it is *a priori* known that for instance the magnitude of  $p_8$  should be between 1.6 and 1.9 and  $p_3$  should be larger than 0.006, it can be deduced that Ia and Ib are based on wrong models. If not, such model errors may go unnoticed.

#### **4.2.2.5 Conclusions,**

- 1) Employing a modified Gauss-Newton optimization routine to minimize the log likelihood function, ML estimates of transfer rates were derived using various sets of simulated observations. These observations were obtained by sampling the model response to a pulse input, calculated for a chosen "true" parameter vector, and by adding random Gaussian noise. It was examined to what extent the true parameter values could be estimated with sufficient accuracy if the number of available observations was reduced.
- 2) Using the maximum number of observations and having all compartments observed, one fifth of the parameters is estimated with an accuracy of 1% or better; one twentieth with an accuracy not better than 30%. The final estimates are to some extent dependent on the initially chosen parameter values. A reduction in the number of observations to one fourth of the maximum number, equally distributed over all compartments, makes the accuracy of the former "1%" parameters shift into the 5%-10% range; the largest deviations from true parameter values remain approximately 30%. In particular, the omission of early observations (0-10 h) does not result in an unacceptable loss of accuracy. Thus, a considerable reduction in the number of laboratory animals to be used in similar future *in vivo* experiments will be possible.
- 3) If unobserved compartments are allowed a substantial loss of accuracy may

TABLE 4-14 PARAMETER VALUES, LOG(LIKELIHOOD FUNCTION) AND GOODNESS OF FIT FOR TRUE AND FALSE MODEL STRUCTURES; SEE FIG. 4.2A

The simulations below were started with initial parameter values according to set 1 (Table 4-8) and all observations in all compartments (Table 4-7). With respect to the true model structure (experiment A) the following false structures were introduced. In experiment I1 pathway spleen-to-liver was replaced by spleen-to-plasma. In experiment I2 pathway plasma-to-urine was replaced by tissues\_1-to-urine. The parameters involved in these changes are written in italics.

| Parameter values ( $\text{min}^{-1}$ ) and SD ( $\text{min}^{-1}$ ) |       |         |        |         |        |         |        |
|---|-------|---------|--------|---------|--------|---------|--------|
|   | true  | A1      |        | I1      |        | I2      |        |
| $p_1$   | .1429 | .1556,  | .0002  | .1549,  | .0005  | .1362,  | .0005  |
| $p_2$   | .0075 | .0074,  | <.0001 | .0074,  | .0001  | .0073,  | <.0001 |
| $p_3$   | .0068 | .0069,  | <.0001 | .0069,  | <.0001 | .0004,  | <.0001 |
| $p_4$   | .0585 | .0505,  | .0015  | .0626,  | .0033  | .0507,  | .0005  |
| $p_5$   | .1343 | .1164,  | .0034  | .1413,  | .0073  | .1177,  | .0012  |
| $p_6$   | 8.18  | 8.841,  | .0001  | 8.798,  | .0001  | 8.768,  | <.0001 |
| $p_7$   | .4054 | .4283,  | .0008  | .4279,  | .0018  | .4234,  | .0007  |
| $p_8$   | 1.726 | 1.474,  | .0007  | 2.105,  | .0015  | 2.833,  | .0012  |
| $p_9$   | .600  | .500,   | .0009  | .6885,  | .0046  | .8977,  | .0039  |
| $p_{10}$  | 9.040 | 9.962,  | .0001  | 9.938,  | .0001  | 9.260,  | .0003  |
| LLF   |       | 88.50   |        | 77.96   |        | 113.40  |        |
| TCC   |       |         |        |         |        |         |        |
| overall   |       | .999970 |        | .999973 |        | .999961 |        |
| plasma  |       | .990860 |        | .990879 |        | .990803 |        |
| tissues_1   |       | .998942 |        | .998903 |        | .997807 |        |
| liver   |       | .997204 |        | .997274 |        | .997297 |        |
| bile  |       | .999987 |        | .999988 |        | .999990 |        |
| tissues_2   |       | .999653 |        | .999661 |        | .999688 |        |
| urine   |       | .999921 |        | .999921 |        | .999913 |        |
| spleen  |       | .998636 |        | .999117 |        | .996344 |        |

result (e.g., two fifths of the parameters showing accuracies poorer than 30%). It cannot be avoided that in practice organs must be sampled that are inaccessible without sacrificing the laboratory animal. However, the number of sample points, thus the total number of animals consumed in an experiment, can be reduced without impairing the achievable accuracy of the obtained results. But, as confirmed by others [Erickson and Ackerman, 1986], *it is better to have fewer observations in many compartments than many observations in only two or three compartments.*

4) The sensitivity of the estimation procedure for deliberately introduced errors in the model structure (distribution pathways) was also investigated. Small model errors appeared to go unnoticed, unless the order of magnitude of the

Fig. 4.20 Data-sensitivity experiments;

Model response with estimated  $\pm 1$  standard deviation and observations with their accuracies ( $\pm 1$  standard deviation) in plasma and liver. Observations, plasma and urine: 100%; other: 0%. Performance criteria in boxes; left values: with respect to observations used for parameter estimation; right values: with respect to all observations (true values between brackets)

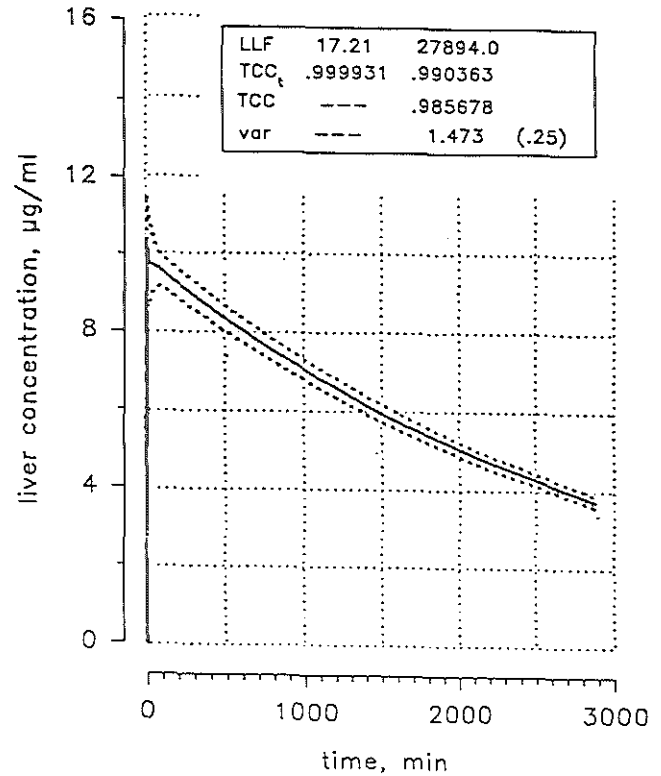
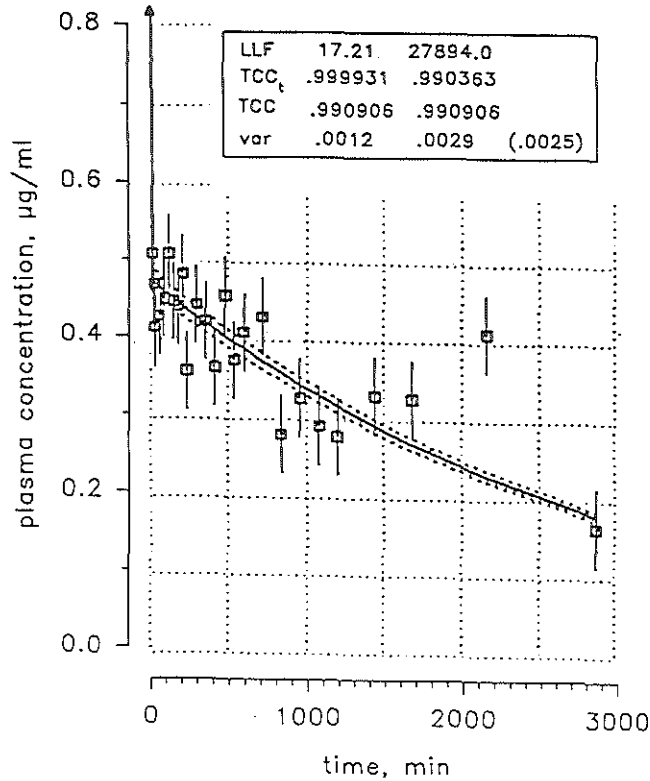


Fig. 4.21 Data-sensitivity experiments;

Model response with estimated  $\pm 1$  standard deviation and observations with their accuracies ( $\pm 1$  standard deviation) in spleen. Comparison of experiments A1 and E (see Table 4-9). Performance criteria in boxes; left values: with respect to observations used for parameter estimation; right values: = with respect to all observation (true values between brackets)

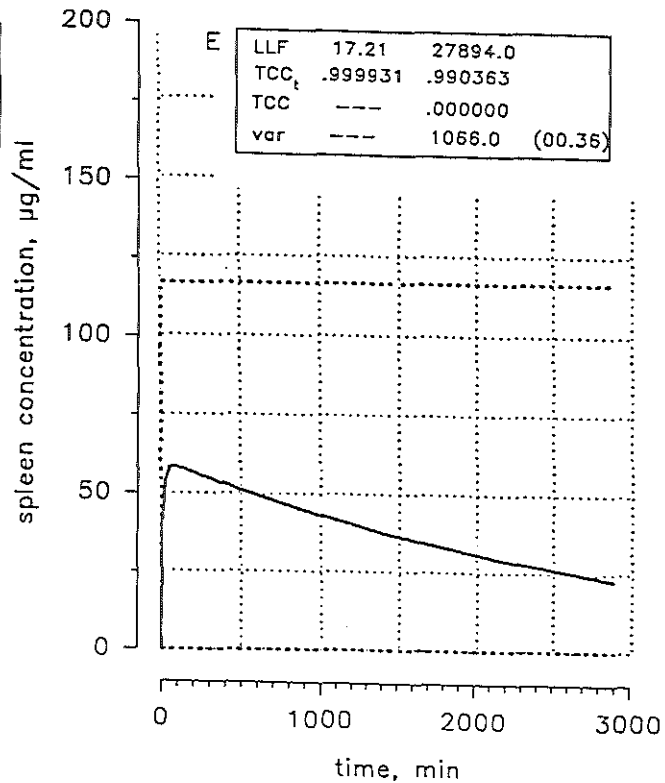
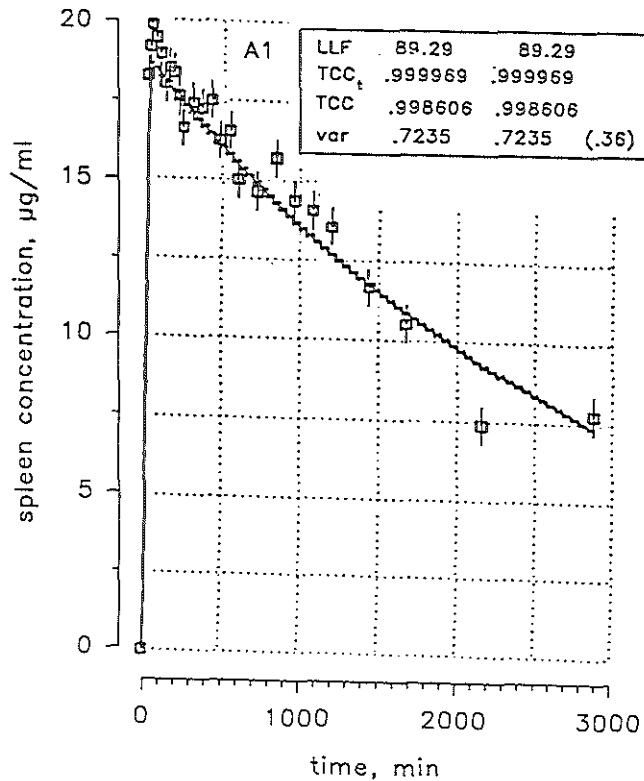


Fig. 4.22 Data-sensitivity experiments;

Model response with estimated  $\pm 1$  standard deviation and observations with their accuracies ( $\pm 1$  standard deviation) in plasma and liver. Observations, plasma, urine and bile: 36%; spleen: 12%; other: 0%. Performance criteria in boxes; left values: with respect to observations used for parameter estimation; right values: with respect to all observations (true values between brackets)

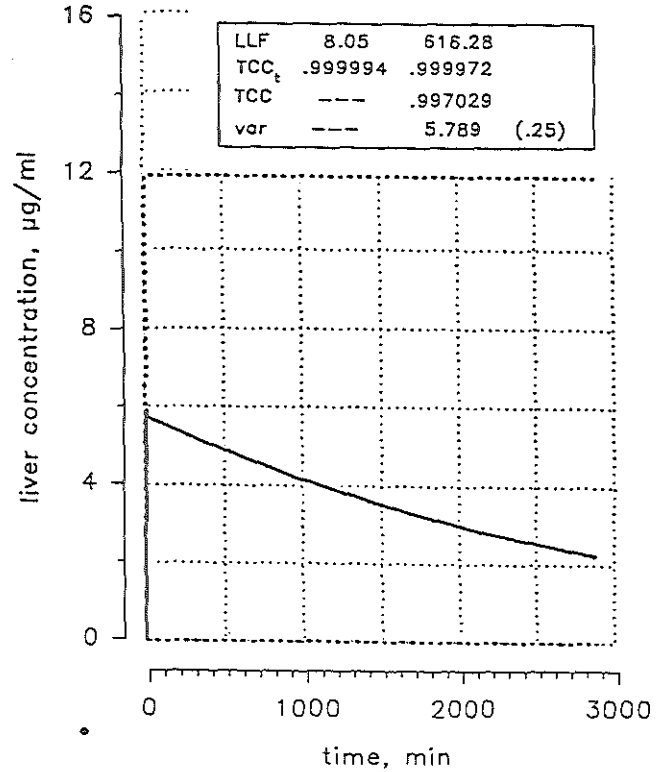
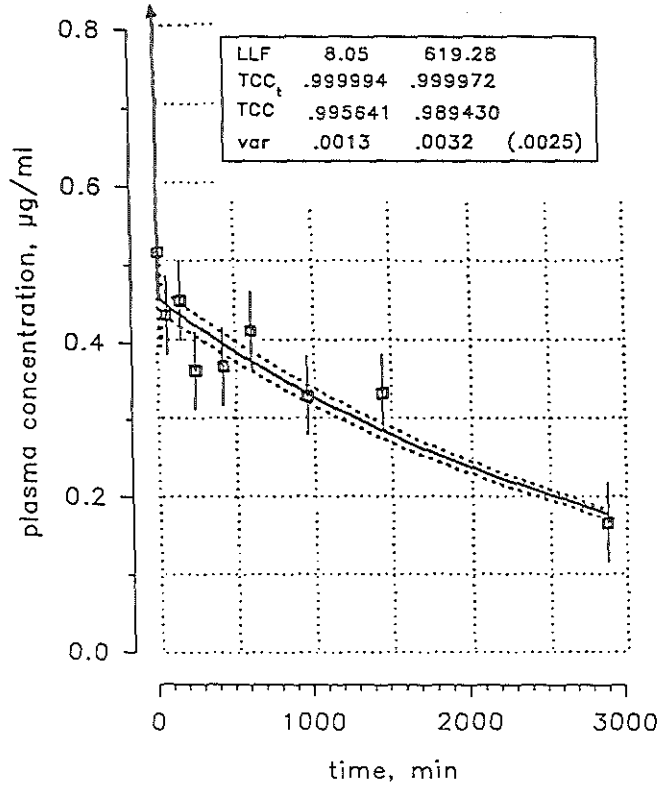




Fig. 4.23 Data-sensitivity experiments;

Model response with estimated  $\pm 1$  standard deviation and observations with their accuracies ( $\pm 1$  standard deviation) in spleen. Comparison of experiments A1 and H (see Table 4-9). Performance criteria in boxes; left values: with respect to observations used for parameter estimation; right values: = with respect to all observations (true values between brackets)

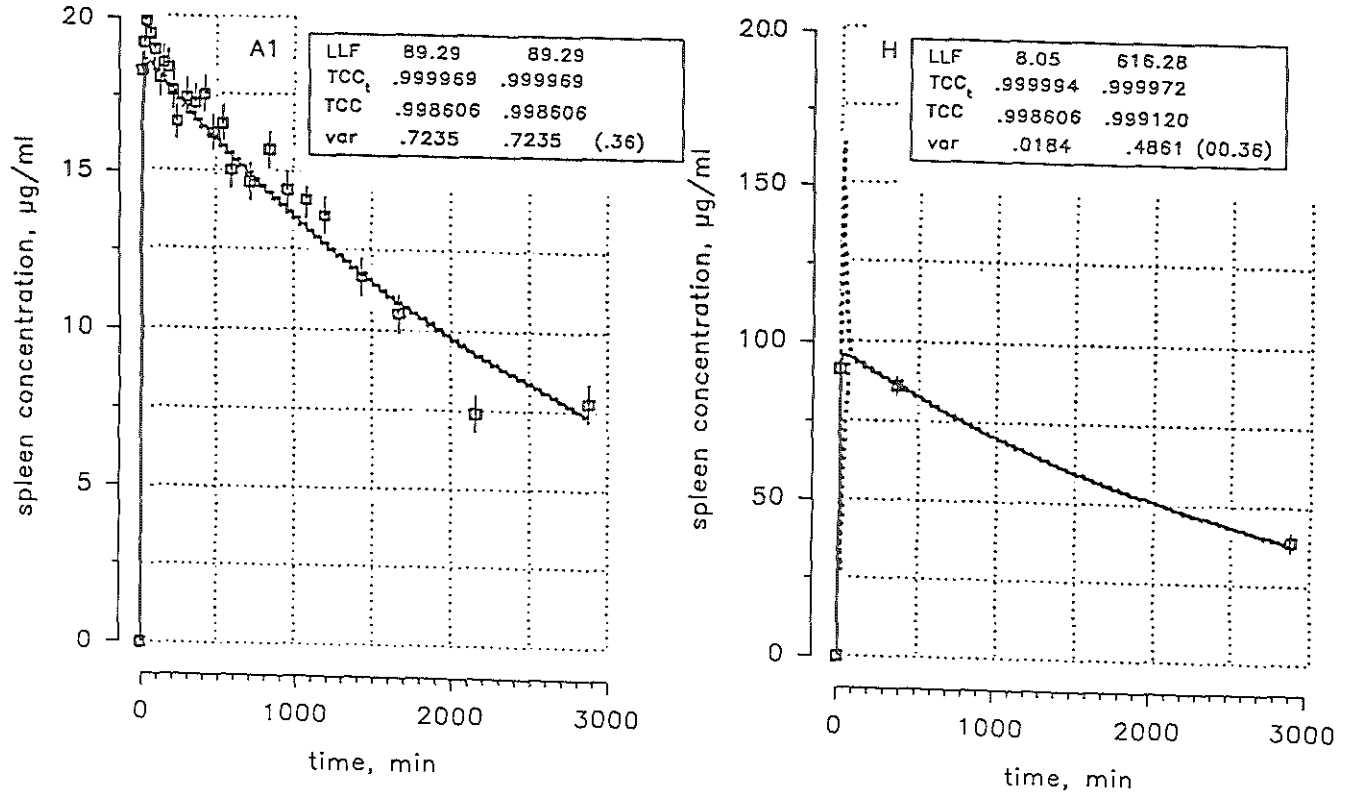
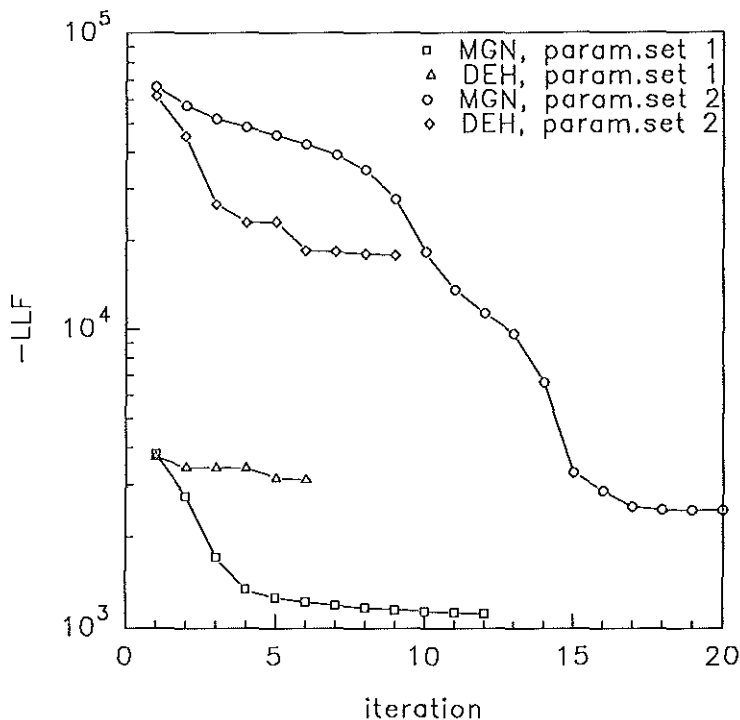


Fig. 4.24 The negative log likelihood function, LLF decreases with increasing number of iterations, after starting with parameter set 1 ( $\square$  and  $\Delta$ ) or 2 ( $\circ$  and  $\nabla$ ) for, respectively, the MGN and DDH routines



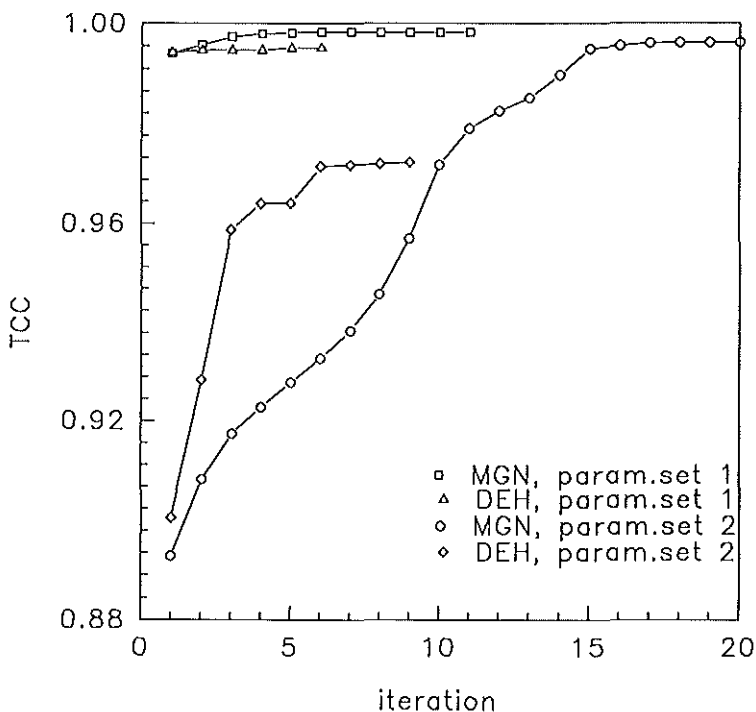
corresponding parameter values are *à priori* known.

5) *Making use of mathematical modeling and computer simulation techniques can yield improved designs of in vivo pharmacokinetic experiments, whenever minimizing the number of laboratory animals required to achieve reliable results is concerned.*

### 4.2.3 Comparison of the Minimization Routines

**4.2.3.1 MGN method.** For both parameter set 1 and 2 the log likelihood function decreases monotonously to steady levels; in 11 iterations from 3822 to 1126, respectively, in 20 iterations from 66565 to 2462 (Fig. 4.24). An increase of the overall goodness of fit is seen (Fig. 4.25). For the first parameter set the differences between initial and final parameter values are shown in Fig. 4.26. For the majority (64%) the change is less than a factor of 5; for 80% of the parameters the change is less than a factor of 10. The finally estimated values

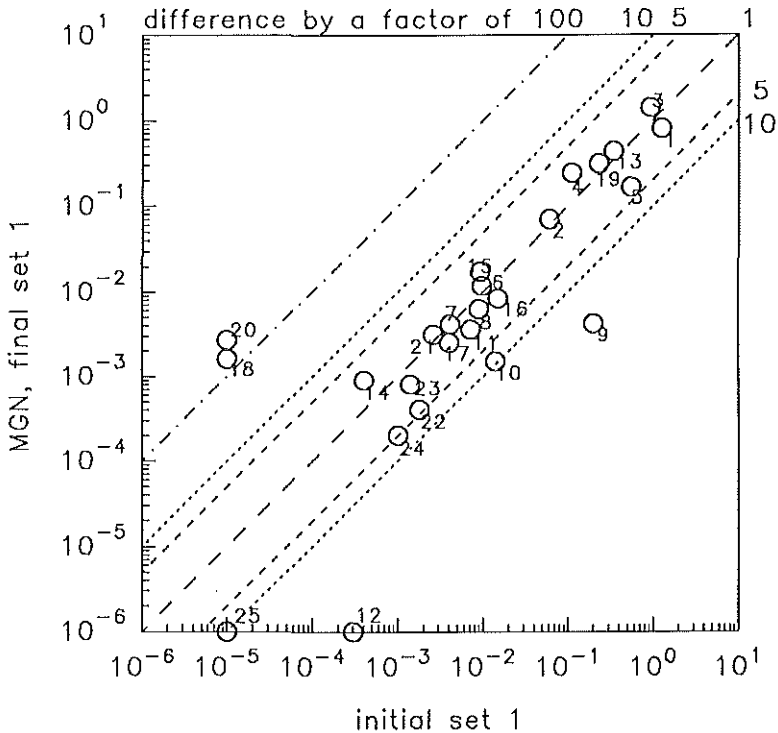
Fig. 4.25 The change in overall goodness of fit (TCC, Eq.(4.51)) with increasing number of iterations, after starting with parameter set 1 ( $\square$  and  $\triangle$ ) or 2 ( $\circ$  and  $\nabla$ ) for, respectively, the MGN and DDH routines



depend on the initial choice; Figure 4.27 shows over 10-fold disagreement for 40% of the parameters. The first set appears to be a 'luckier' initial choice, resulting in a lower log likelihood function value. Thus, the MGN routine converges both times, but not to the same point in the parameter space. The estimated SD's of the parameters (Fig. 4.28) reveal that for some parameters (about 60% of the cases) the differences are not significant while for other ones they are. For a few parameters, always associated with metabolite production or distribution, negative values are found. It is difficult to decide whether this result is caused by inaccuracy of the numerical method in finding the minimum LLF, or that any model assumption is wrong. However, the negative parameter values do not differ significantly from zero, except for one case:  $p_{18} = k_{9,2}$ , which would mean an inversed metabolism (DOL-DAU) in the tissues<sub>1</sub> compartment.

Figure 4.29 shows the consumed cpu-time for both minimization runs. The time required for one iteration step amounts to some 14 minutes. Because of exceeding an upper limit for batch job execution time, in one case a restart

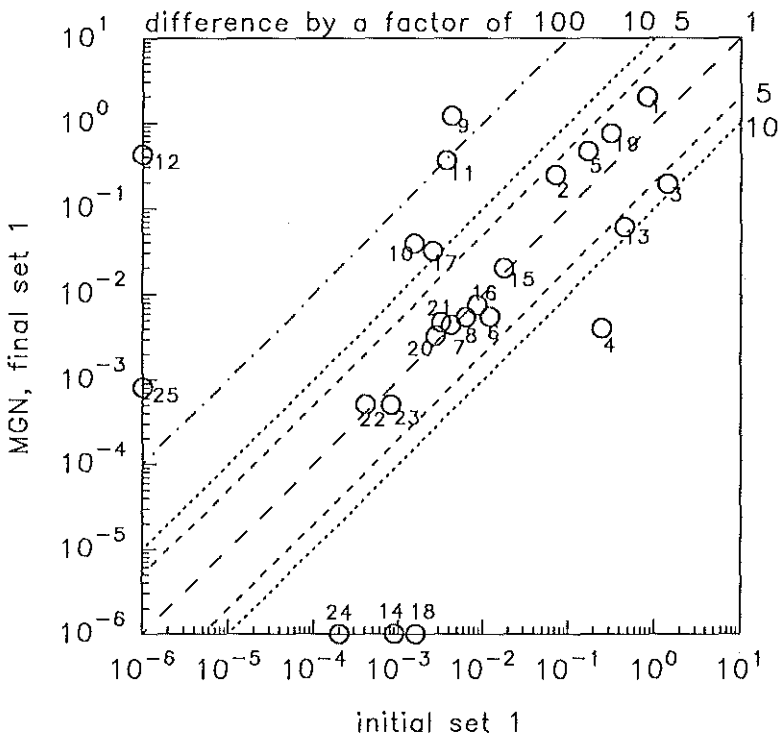
Fig. 4.26 Comparison of initial and final (set 1-)parameter values after MGN optimization. Bands of factor of 5, and factor of 10 differences are indicated



was necessary in the 14th iteration.

**4.2.3.2 FD method.** In the MGN method the calculation of the transition matrices is rather time consuming (about 400 s each iteration). An equally long time is needed for the computation of the model and sensitivity responses, the residuals and their autocovariance matrix, the information matrix, the SVD and the determination of the parameter step direction. In particular, the computation of the transition matrices for the sensitivities requires much cpu-time. In the FD method these sensitivities are computed in a different way, by evaluating the change in the model response due to a slight change in each of the  $m$  elements of the parameter vector. In the first iteration step it requires  $m$  ( $= 25$ ) extra calculations of the model response. Taking some 6.5 s each (see DDH method below), this adds up to 165 s. Subsequently, the computation of the elements of the sensitivity vectors involve the solution of  $2n$  sets of  $m$  simultaneous algebraic equations, e.g. by Cramer's rule, which must be repeated for every observation time  $N$ . For the present values of  $m$ ,  $2n$  and  $N$  (25, 14 and 26) this

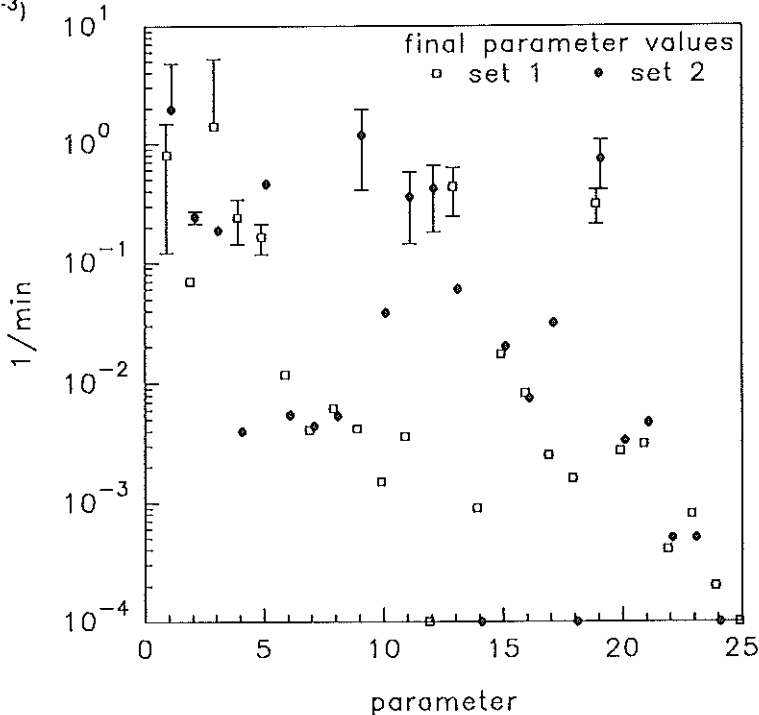
Fig. 4.27 Comparison of final parameter values obtained by MGN optimization after starting with initial sets 1 or 2. Bands of factor of 5, and factor of 10 differences are indicated



will require an estimated cpu-time of another 615 s, so, the total cpu-time for the model and sensitivity responses will amount to some 780 s. Time up to determination of parameter step direction will then be about 785 s, which compared to the 800 s needed by the MGN method means only a marginal improvement. During further iteration steps, however, always only one extra model response calculation (6.5 s instead of 165 s) is needed. Thus, the advantage of the FD approach increases with the number of iterations (about 30 min profit per 10 iterations). The corresponding fortran routine for the FD method is currently being implemented for testing.

**4.2.3.3 DDH method.** Like in the MGN method the negative log likelihood function decreases with increasing iteration number (Fig. 4.24). The rate of the decrease, however, is much lower; after 6 iterations LLF is still high. In the test runs LLF goes from 375 to 312, respectively, from 62000 to 18500 for parameter sets 1 and 2. The goodness of fit improves, but also only very slowly (Fig.

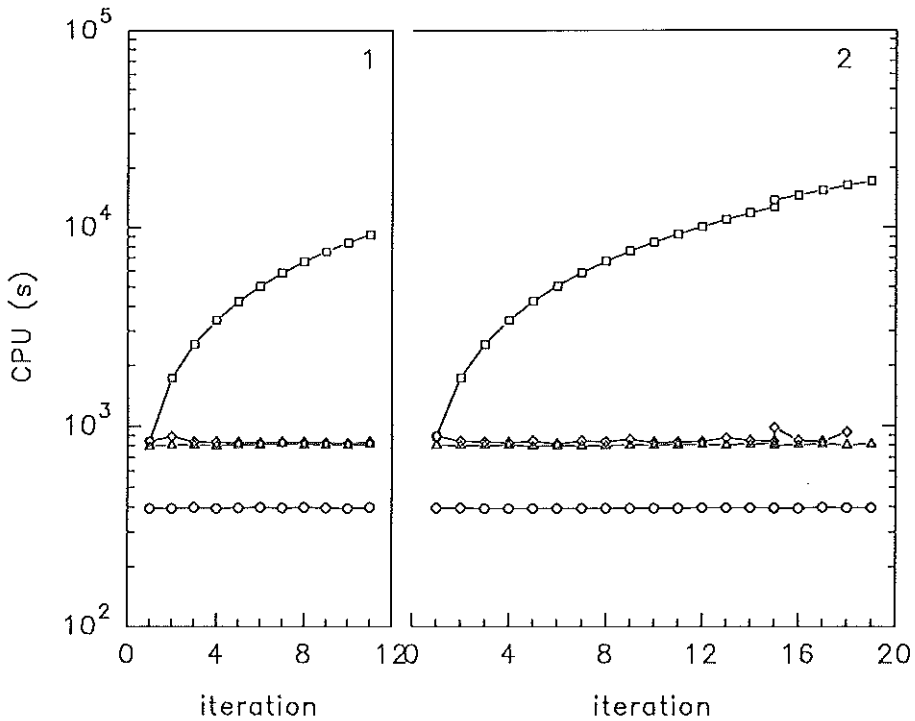
Fig. 4.28 MGN method: final parameter values and estimated SD after starting with parameter sets 1 ( $\square$ ) or 2 ( $\bullet$ ). Zero and negative values are drawn below the  $10^{-4}$  line (Set 1:  $p_{12} = -0.57 \times 10^{-2} \pm 0.66 \times 10^{-2}$ ,  $p_{25} = -0.56 \times 10^{-2} \pm 0.206 \times 10^{-1}$ ; Set 2:  $p_{14} = -0.1 \times 10^{-4} \pm 0.4 \times 10^{-4}$ ,  $p_{18} = -0.698 \times 10^{-1} \pm 0.344 \times 10^{-1}$ ,  $p_{24} = -0.5 \times 10^{-4} \pm 0.3 \times 10^{-3}$ )



4.25). Needing much cpu-time, the method is very slow. To estimate the elements of the Hessian matrix and the gradient vector of the likelihood function, 350 evaluations of the function LLF, requiring some 6.5 s each, must be carried out. So, each iteration lasts about 38 min (Fig. 4.30). For every 4 iterations by the DDH program the MGN program completes almost 14.

Although the DDH runs were not finished properly (they ended due to exceeding a maximum cpu-time limit) and the final estimates of the parameter values therefore were not computed, two of the intermediate values, corresponding with the tissues metabolic rates, persistently remained negative. This suggests that the model may not be quite correct as far as metabolism is concerned. Perhaps metabolic changes in reality show a nonlinear behavior, e.g., due to a limited capacity of the enzymes involved. This would imply incorporation of a nonlinear saturation process in the model. In that case the state equations (Eq.(4.8)) can no longer be conveniently solved by means of transition matrices (TM). Instead, some numerical integration method, e.g., according to Runge-Kutta (RK), must be used. This will cause an even higher

Fig. 4.29 MGN method: consumed cpu-time after starting with parameter sets 1 (left) or 2 (right);  $\square$  total accumulated time; time needed for computation of transition matrices  $\circ$ , determination of direction  $\triangle$  and magnitude  $\diamond$  of parameter step



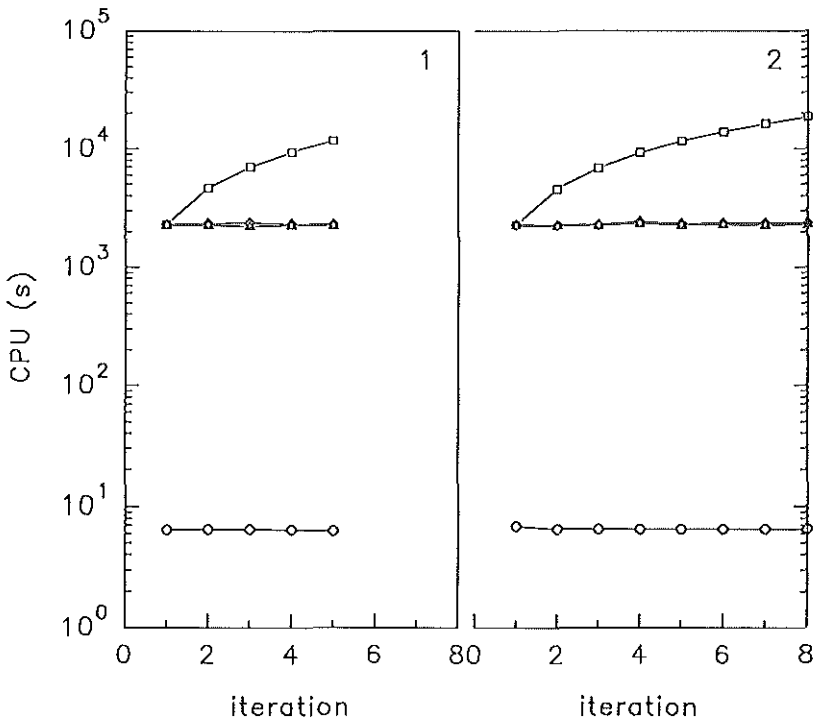
cpu-time consumption, as was demonstrated with a five compartment/seven parameter model (see below). Identification of large multicompartment pharmacokinetic models then becomes a rather tedious affair from a practical point of view, unless perhaps a mainframe computer is available to keep response times acceptable.

#### 4.2.3.4 Comparison of model responses, MGN and DDH methods.

Figure 4.31 shows, as an example, the model responses in spleen (DAU) and plasma (DOL) for the initial parameter values (set 1) and the final estimates according to the MGN and the DDH optimization. Both methods result in an improved goodness of fit, but the MGN procedure turns out to be the most successful. Goodness of fit for the other compartments is shown in Fig. 4.32.

A five compartment/seven parameter model (Fig. 4.2b) can be handled well on a minicomputer. Starting with a same set of initial parameter values three different optimization runs were performed (Table 4-15): A) using the TM method to compute model response and solve sensitivity equations according to

Fig. 4.30 DDH method: consumed cpu-time after starting with parameter sets 1 (left) or 2 (right);  $\square$  total accumulated time; time needed for computation of state vector  $\circ$ , determination of direction  $\triangle$  and magnitude  $\diamond$  of parameter step



the MGN routine; B) as A, but computation of the sensitivities according to the DDH routines; C) using a RK integration method (with a small time step for accuracy reasons) to compute model response and the DDH routines for the sensitivities. The average times (s) required for computation of the transition matrices/model response; parameter step; and a complete iteration amount to A: 5.6, 29.9, 35.5; B: 0.6, 20.7, 21.3; and C: 21, 901, 922, respectively. It shows that the RK integration method is very time consuming. The normal DDH iteration here is faster than the MGN method, instead of slower as happens when more parameters are involved. However, the method needs more iterations and does not arrive at the same minimum as the MGN method. Case A: LLF decreases from 2685 to 44, TCC from .656 to .995 in 5 iterations (3 min cpu); case B: LLF from 2685 to 282, TCC from .656 to .970 in 18 iterations (6.4 min cpu-time); and case C: LLF from 2685 to 345, TCC from .656 to .963 in 11 iterations (2.8 h cpu-time).



TABLE 4-15 PERFORMANCE OF DIFFERENT OPTIMIZATION ALGORITHMS

Model: 5 compartments; 7 parameters; 26 observation times

| Algorithm   | A            | B            | C            |
|---|--------------|--------------|--------------|
| <i>state vector</i> calculated with   | TM           | RK           | TM           |
| <i>information matrix</i> calculated with   | TM           | DDH          | DDH          |
| <u>Mean CPU Time per iteration (s)</u>  |              |              |              |
| • TM (2')   | 5.6          |              |              |
| • model response  |              | 21           | 0.6          |
| • model response + residuals +<br>information matrix + SVD +<br>direction of parameter step | 14.9         |              |              |
| • line minimization   | 15.0         |              |              |
| • Hessian (=35 model responses)   |              | 726          | 17.0         |
| • SVD + line minimization +<br>parameter step   |              | 175          | 3.7          |
| • <u>iteration</u>  | <u>35.5</u>  | <u>922</u>   | <u>21.3</u>  |
| Number of Iterations Required   | 5            | 11           | 18           |
| <u>Total CPU Time of run (s)</u>  | <u>177.5</u> | <u>10142</u> | <u>383.4</u> |
| -Log(Likelihood Function) value   |              |              |              |
| <i>start</i>  | 2685         | 2685         | 2685         |
| <i>final</i>  | 44           | 345          | 282          |
| Total Correlation Coefficient   |              |              |              |
| <i>start</i>  | .656         | .656         | .656         |
| <i>final</i>  | .995         | .963         | .970         |

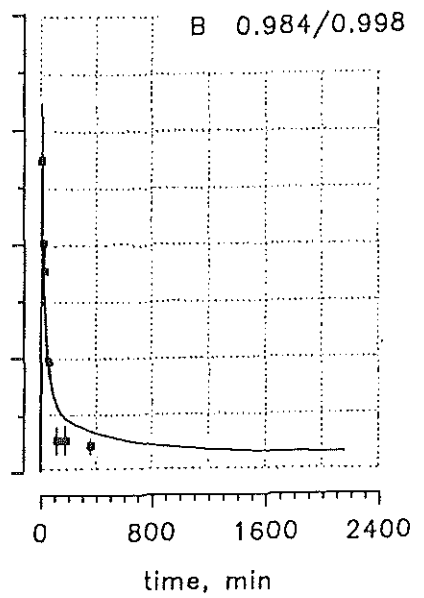
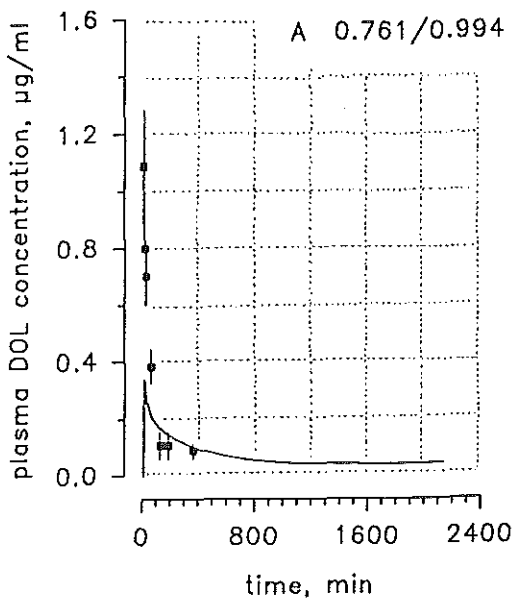
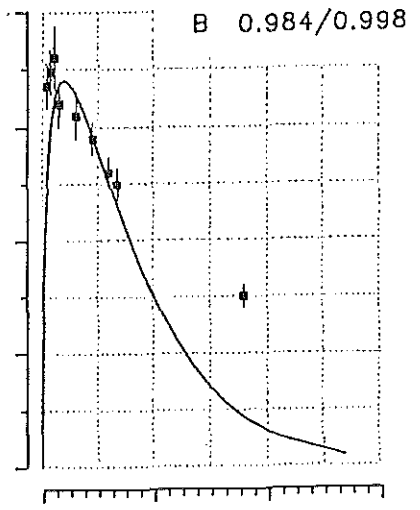
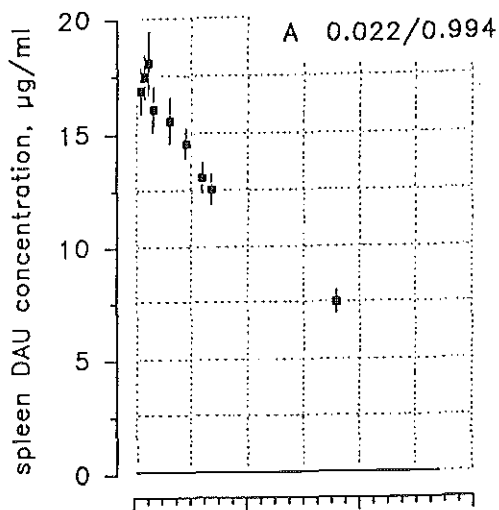
TM: transition matrix (matrices)

RK: direct integration by Runge-Kutta method

DDH: direct estimation of Hessian through finite differences

SVD: singular value decomposition

**4.2.3.5 Conclusion.** Clearly the application of the MGN method deserves preference to the DDH method. Both methods converge, but the MGN one converges faster and reaches a lower minimum. The FD approach to the computations will make the MGN routine faster, but the difference can only be noticed if the number of iterations is large. Whichever method is chosen, by recommendation several runs should be made, starting with different sets of parameter values, to make sure that the method has not converged to a local but to the global minimum.



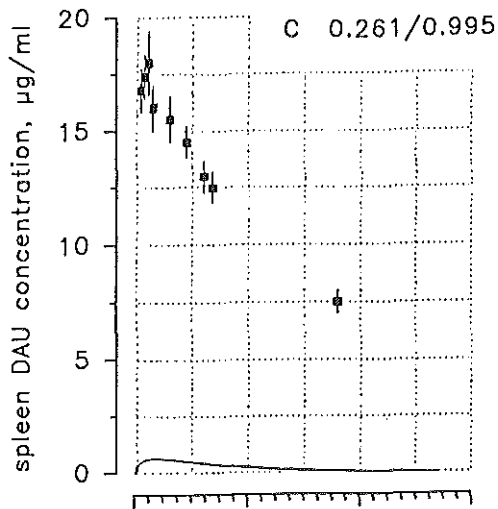


Fig. 4.31

2x7 compartment model

response, *upper row*: spleen DAU and *lower row*: plasma DOL concentrations as function of time

- A: initial parameter set 1
- B: after MGN optimization
- C: after DDH optimization

Goodness of fit: indicated are TCC in compartment/TCC overall

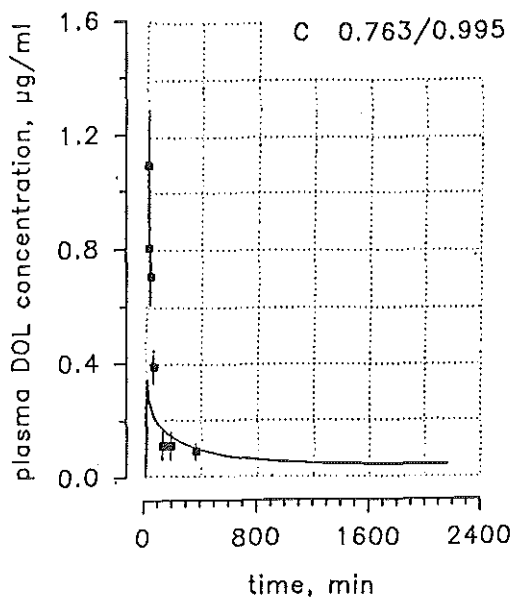
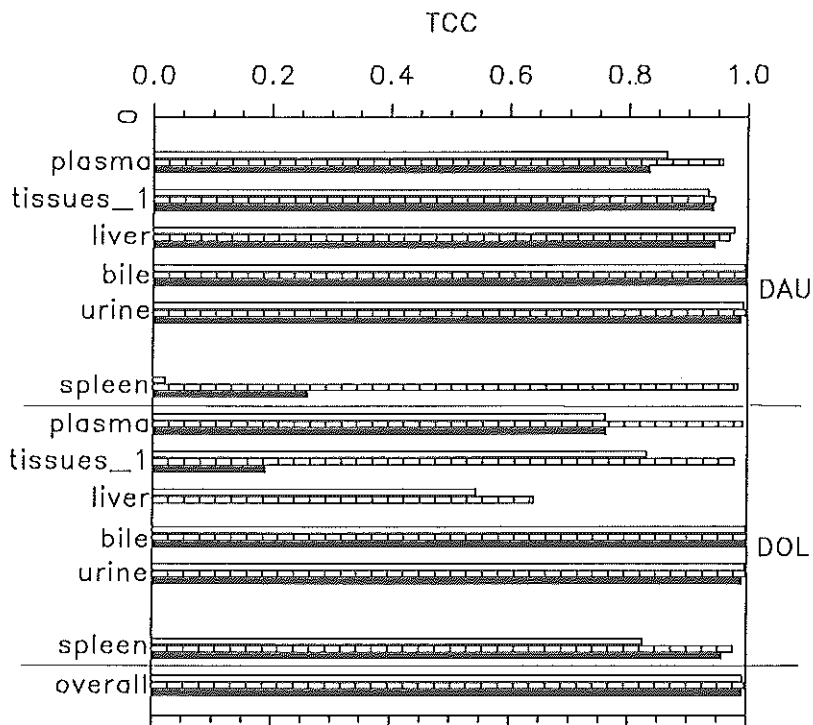


Fig. 4.32 2x7 compartment model; goodness of fit (TCC, Eq.(4.51,52)), in the compartments and overall, for □ initial parameter set 1; ● after MGN optimization; ▲ after DDH optimization



#### 4.2.4 Identification of the Daunomycin Pharmacokinetics

To study the pharmacokinetics of DAU and DOL in the rat after a single *i.v.* bolus injection of the parent drug various multicompartement models (2x3, -5, -6, -7 and -11) have been tested (Tables 4-1 and 4-11; Fig. 4.1). Increasing the complexity of the model structure, by considering more and more separate organs, allows incorporation and evaluation of enhanced pharmacokinetic detail. Identification of the increasing number of transfer rate constants for distribution and metabolism remains possible as the number of available observations increases proportionally; the observations/parameters ratio is always 2-4. The described numerical MGN optimization method has been used. Beside this simultaneous parameter optimization method, for the 2x3 compartment model the analytical solution and sequential fitting method have been tried as well.

The resulting values of the various performance criteria are listed in

TABLE 4-16 COMPARISON OF THE LARGE-SCALE MODELS

| model | m  | NR  | TCC  | SSR                 | PF                  | AIC    |
|-------|----|-----|------|---------------------|---------------------|--------|
| 01    | 29 | 132 | .926 | 3.6x10 <sup>4</sup> | 1690.7              | 1442.9 |
| 02    | 37 | 132 | .947 | 480.4               | 263.3               | 889.0  |
| 03    | 44 | 132 | .948 | 374.8               | 137.2               | 870.3  |
| 04    | 37 | 132 | .895 | 3.3x10 <sup>6</sup> | 2.8x10 <sup>6</sup> | 2056.6 |
| 05    | 45 | 132 | .948 | 207.1               | 133.9               | 794.0  |
| 06    | 45 | 132 | .893 | 945.0               | 4.1x10 <sup>5</sup> | 994.4  |

m = number of parameters; NR = number of observations;  
 for the definitions of the criterion variables TCC, SSR, PF and AIC, see Eq.(4.51) through Eq.(4.54), and Eq(4.24)

TABLE 4-17 ESTIMATED ML TRANSFER RATE CONSTANTS ( $k_{to,from}$ ), [min<sup>-1</sup>], AND SD, [%]; MODEL 05 (SEE FIG. 4.4; P: PARENT DRUG, M: METABOLITE)

| k      |              | k       |              | k       |              |
|--------|--------------|---------|--------------|---------|--------------|
| 2P,1P  | .3010 (0.02) | 1P,10P  | .0085 (3.5)  | 14M,12M | .0045 (13.3) |
| 3P,1P  | .0064 (7.8)  | 11P,1P  | .6008 (.002) | 15M,12M | .1E-5 ( * )  |
| 4P,1P  | .0196 (3.6)  | 1P,11P  | .0302 (0.7)  | 16M,12M | .7E-4 (714.) |
| 5P,1P  | .0716 (0.4)  | 13M,2P  | .0013 (30.8) | 17M,12M | .0106 (3.8)  |
| 6P,6P  | .0497 (0.4)  | 13M,14M | .0184 (0.5)  | 18M,12M | .1E-5 ( * )  |
| 2P,3P  | .0033 (12.1) | 12M,13M | .0060 (10.0) | 21M,12M | .1E-5 ( * )  |
| 1P,2P  | .0150 (2.0)  | 19M,13M | .0030 (80.0) | 22M,12M | .0024 (12.5) |
| 1P,4P  | .0134 (4.5)  | 12M,15M | .0010 (20.0) | 14M,3P  | .0002 (20.0) |
| 1P,5P  | .0179 (4.5)  | 12M,16M | .0033 (24.2) | 15M,4P  | .0027 (22.2) |
| 1P,6P  | .0050 (24.0) | 12M,17M | .0163 (2.4)  | 16M,5P  | .0027 (18.5) |
| 8P,2P  | .0120 (3.3)  | 12M,18M | .4E-4 (150.) | 17M,6P  | .0017 (35.3) |
| 7P,1P  | .9996 (.002) | 12M,21M | .2E-5 (150.) | 18M,7P  | .0020 (2.5)  |
| 1P,7P  | .0049 (8.2)  | 12M,22M | .0007 (100.) | 21M,10P | .0006 (33.3) |
| 9P,1P  | .0095 (3.2)  | 20M,12M | .0026 (15.4) | 22M,11P | .0015 (20.0) |
| 10P,1P | .0130 (1.5)  | 13M,12M | .0230 (1.7)  | 12M,1P  | .1E-5 ( * )  |

\* larger than 999%; (E-a denotes x10<sup>-a</sup>)

Tables 4-16 and 4-19. They enable the comparison of the different models.

It appears that the various goodness of fit criteria do not agree upon one "best" model. Optimum models, for a given compartmental size, are 3a1.2, 5b3, 6b1, 7b1 and 05.

**4.2.4.1 Comparison of the Small-scale Multicompartment Models.** Both the analytic sequential and the numerical simultaneous solution can yield satis-

factory fits to the datapoints. The former method is rather sensitive to the chosen initial parameter values (model 3a1.1 and 3a1.2). The second fit (excretion DAU) depending on the results of the first fit (plasma DAU) also is a disadvantage. The analytic solution predicts an unrealistically large plasma volume. The numerical solution keeps the plasma volume at its physiological value, even if it is free to vary this parameter, but it does not arrive at as good a plasma DAU fit without artificially weighting the observations.

The excretion compartment accounts for plasma concentration dependent urine excretion. The (substantial) bile excretion must either be ignored (thus, disappears in the tissue compartment) or added to urine excretion. In the latter case it is modeled as plasma concentration dependent rather than liver concentration dependent, what would be more realistic. Here, both ways of modeling are equally good. Urine alone yields the lowest AIC, urine + bile yields higher TCC.

**4.2.4.2 Comparison of Intermediate-scale Multicompartment Models.** In the  $5 \times 2$  compartment model the number of transfer rates to be estimated already becomes large. Therefore, it is difficult to find reasonable first estimates to start optimization with. The importance of a good optimization algorithm thus is obvious (model 5b1 and 5b3). Overall fit in the  $5 \times 6$  and  $7 \times 2$  compartment models is comparable. The accuracy of the estimated transfer rates is rather poor, especially for metabolism and metabolite distribution. Improvement with increase of model size is seen in the predicted accuracy of the resulting concentration-time curves.

**4.2.4.3 Comparison of the Large-scale Multicompartment Models.** It was tried to explain the observed data with models that comprise hypotheses about the characteristics of the processes of drug transport and metabolism as listed in Table 4-1. Starting from several sets of arbitrarily chosen initial values the parameters in each model were optimized in the ML sense, using the modified Gauss-Newton algorithm. By evaluating the various optimization criteria the most likely model could be identified as model 05 (Table 4-16).

Based on the total correlation coefficient, TCC, the models 03 (no plasma metabolization) and 05 (plasma metabolization) cannot be discriminated. There also is hardly a difference with model 02 (metabolism at constant rates). The superiority of model 05 is demonstrated by the lowest values of the sum of squared residuals (SSR) and the performance index (PF). Also, the Akaike Information Criterion (AIC) is lowest, despite the highest number of parameters to be estimated.

Figure 4.33 compares the optimal responses (DAU and DOL time-histories) of the various models in four typical compartments, i.e., plasma, urine, liver and heart. As can be seen, the model responses improve considerably with

TABLE 4-18

GOODNESS OF FIT: TCC, SSR AND SDR PER COMPARTMENT FOR OPTIMUM MODEL RESPONSE; (MODEL 05)

| compartment  | TCC  |       | SSR   |      | SDr ( $\mu\text{g}\cdot\text{g}^{-1}$ ) |      |
|--------------|------|-------|-------|------|---|------|
|              | DAU  | DOL   | DAU   | DOL  | DAU                                     | DOL  |
| plasma       | .782 | <.001 | 6.0   | 5.3  | 0.21                                    | 0.25 |
| liver        | .977 | .952  | 8.9   | 9.1  | 2.7                                     | 1.1  |
| spleen       | .989 | .925  | 9.2   | 8.6  | 2.2                                     | 1.6  |
| heart        | .985 | .834  | 7.9   | 11.8 | 1.9                                     | 2.5  |
| kidneys      | .997 | .932  | 7.3   | 7.4  | 1.5                                     | 2.6  |
| lungs        | .901 | .921  | 9.0   | 8.8  | 11.7                                    | 1.6  |
| other_tissue | -    | -     | -     | -    | -                                       | -    |
| urine        | .999 | .981  | 4.8   | 4.5  | 0.3                                     | 2.5  |
| bile         | -    | -     | -     | -    | -                                       | -    |
| bone marrow  | .905 | .942  | 5.0   | 5.2  | 2.6                                     | 0.3  |
| muscles      | .945 | .963  | 8.0   | 80.4 | 0.8                                     | 0.2  |
| overall:     | .948 |       | 207.1 |      |   |      |

N.B., for the definition of TCC and SSR, see Eq.(4.51) through Eq.(4.53);  
SDr is the standard deviation of the residuals

TABLE 4-19 RESULTS OF EVALUATING ADDITIONAL (SMALLER-SCALE) MODELS FOR DAU-DOL PHARMACOKINETICS

| model | n   | w   | m  | NR | method | TCC  | SSR                | PF     | AIC     |
|-------|-----|-----|----|----|--------|------|--------------------|--------|---------|
| 3a1.1 | 2x3 | 2x2 | 9  | 16 | seq.   | .757 | 15.98              | -      | 62.337  |
| 3a1.2 | 2x3 | 2x2 | 9  | 16 | seq.   | .982 | 13.69              | -      | 59.872  |
| 3b1   | 2x3 | 2x2 | 9  | 22 | sim.   | .978 | 19.07              | 109800 | 82.864  |
| 3a2   | 2x3 | 2x2 | 9  | 20 | seq.   | .903 | 11.10              | -      | 66.138  |
| 3b2   | 2x3 | 2x2 | 8  | 22 | sim.   | .999 | 40.33              | 19.66  | 97.334  |
| 3b3   | 2x3 | 2x2 | 9  | 22 | sim.   | .991 | 41.02              | 123600 | 99.708  |
| 5b1   | 2x5 | 2x3 | 15 | 32 | sim.   | .998 | $11.3 \times 10^8$ | 122.83 | 697.079 |
| 5b2   | 2x5 | 2x3 | 14 | 32 | sim.   | .998 | $14.0 \times 10^6$ | 694.00 | 544.664 |
| 5b3   | 2x5 | 2x3 | 15 | 32 | sim.   | .994 | 43.43              | 31.30  | 150.678 |
| 6b1   | 2x6 | 2x5 | 19 | 76 | sim.   | .994 | 65.80              | 91.70  | 356.185 |
| 7b1   | 2x7 | 2x6 | 24 | 94 | sim.   | .995 | 82.53              | 127.58 | 462.836 |

tissues\_1 = well perfused tissues: liver, spleen, heart, kidneys, lungs and skeletal muscles; tissues\_2 = poorly perfused tissues: bone marrow and other tissues that can be reached by the drug, but for which no observations are available. Tissues that cannot be reached by the drug, for example brain tissue and bones, are excluded from the models.

n = number of compartments; w = number of observed compartments; m = number of parameters; NR = number of observations; seq. = sequential fitting (analytical method); sim. = simultaneous solution (numerical method).

for the definition of TCC, SSR, AIC, PF: see Eq.(4.51) through Eq.(4.54), and Eq.(4.24)

Fig. 4.33 Parent drug (DAU) and Metabolite (DOL) concentration-time curves for: A: Plasma, B: Urine, C: Liver and D: Heart; Comparison of optimum responses for models 01 (.....), 02 (----), 03 (————), 04 (-•-•-), 05 (~ ~ ~ ~) and 06 (— — —). Observations are shown as well. Best model, 05, often shows overlap with model 03

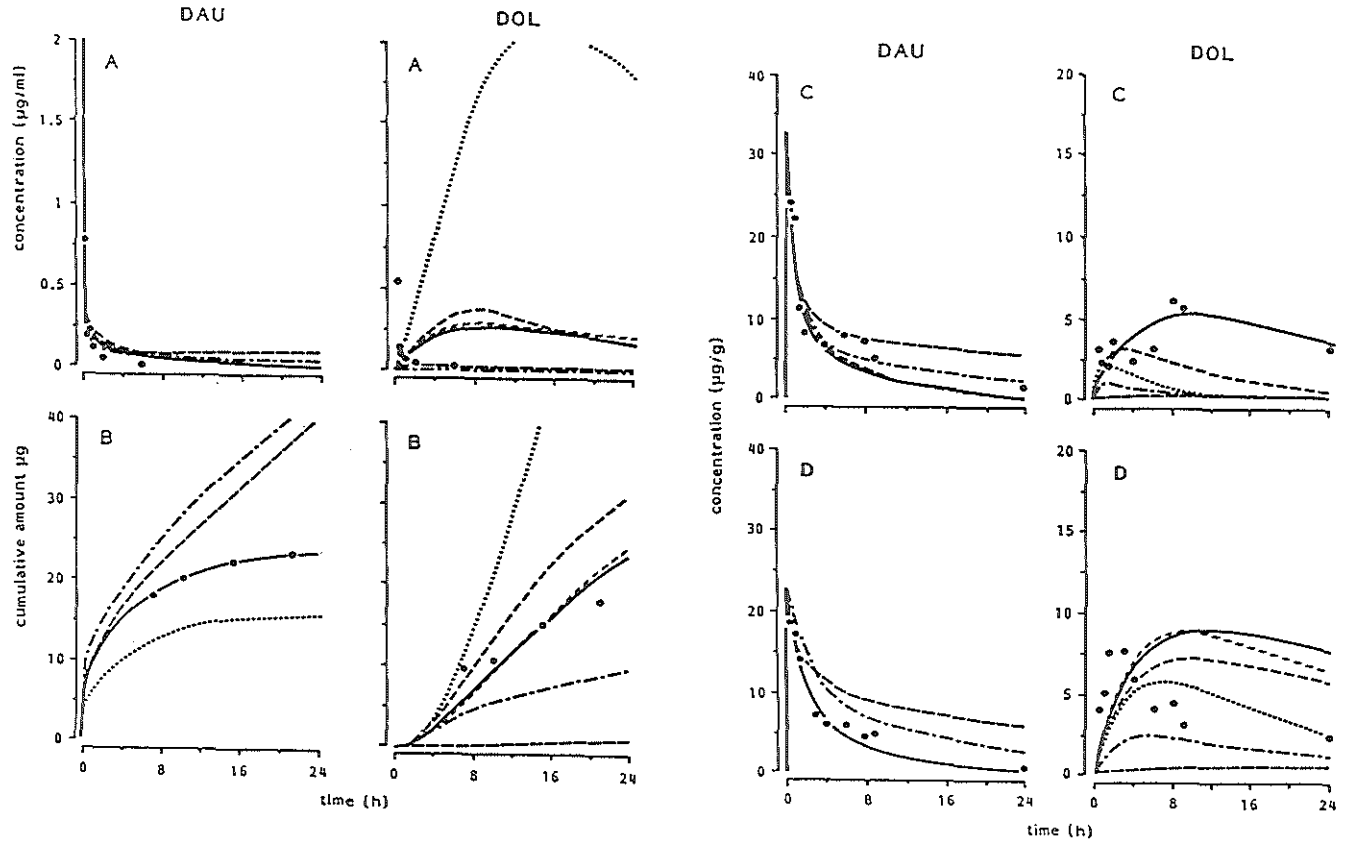
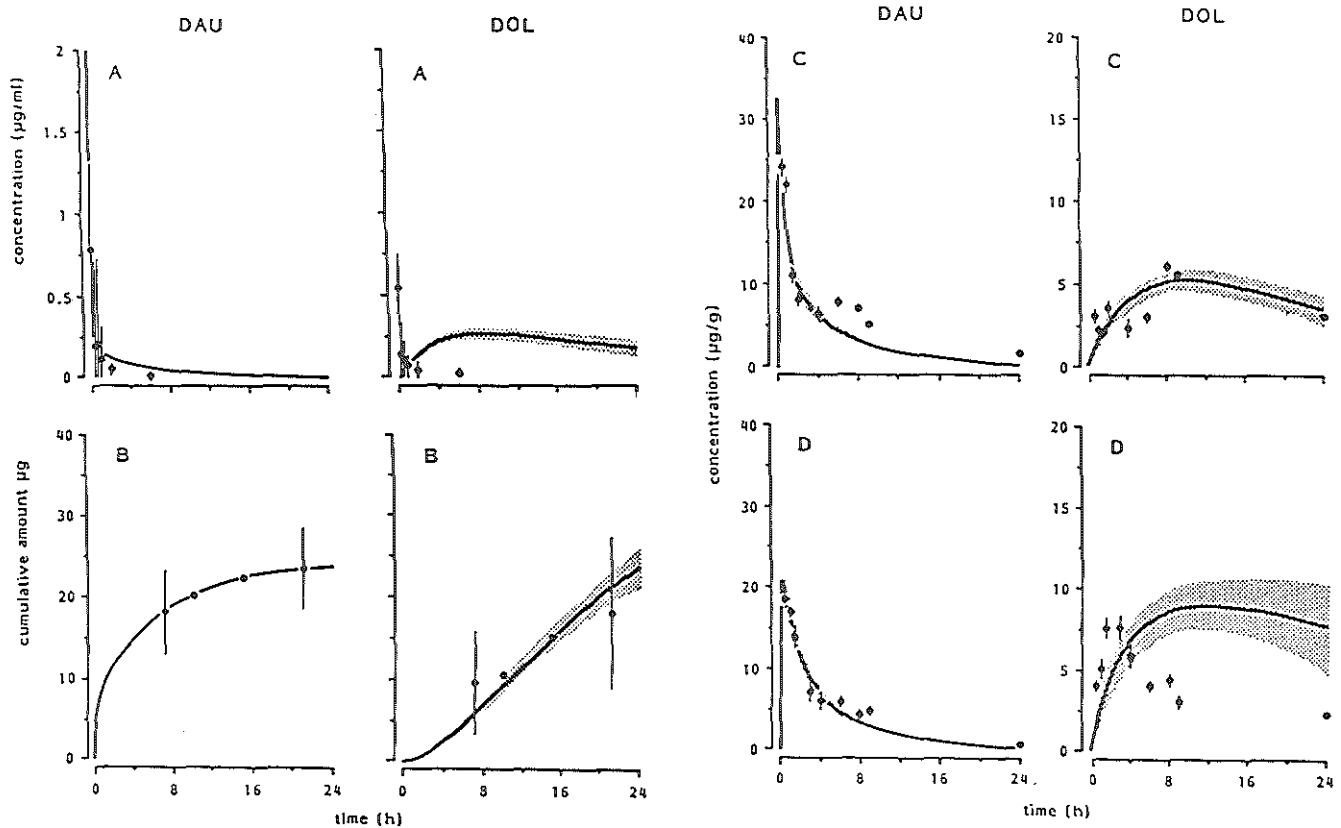




Fig. 4.34 Parent drug (DAU) and Metabolite (DOL) concentration-time curves for: A: Plasma, B: Urine, C: Liver and D: Heart; Optimum response with estimated standard deviation (shaded area) for the best of the tested models (05); observations are shown with their standard deviations



the model structure changing from model 01 to 05 (better agreement with the observed datapoints). In contrast, assuming constant rate metabolic processes (model 06) makes the model response deteriorate.

Figure 4.34 shows again the optimum response of the best model (05) in the mentioned compartments. Also, the calculated accuracy of this model response is indicated. The corresponding parameter values are listed in Table 4-17. The time dependent standard deviations of the state vector elements are very small when the parent drug is concerned, but considerably larger for the metabolite in most compartments.

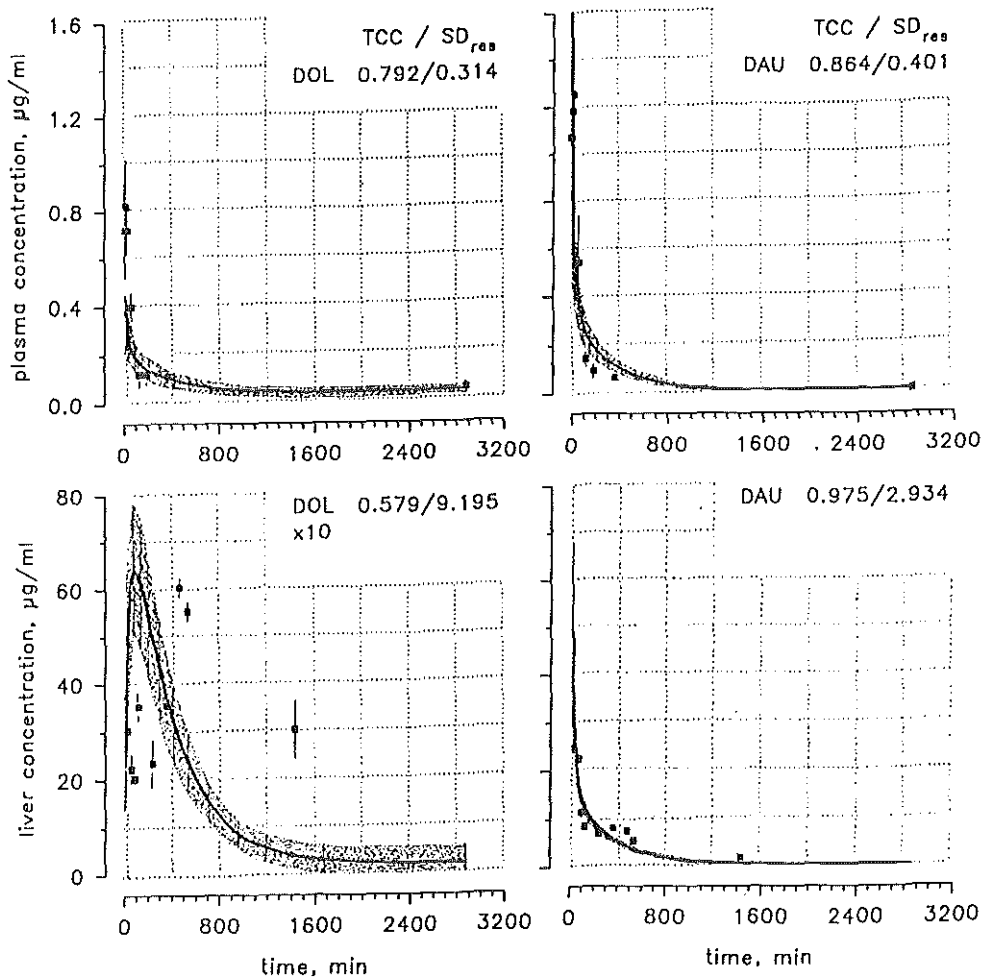
Looking at the optimum parameter values (Table 4-17) it is seen that the DAU distribution parameters, expressed in  $\text{min}^{-1}$ , range from  $6 \times 10^{-3}$  (plasma to spleen) to 1.0 (plasma to other\_tissues) when uptake from plasma is considered, and from  $3 \times 10^{-3}$  (spleen to liver) to  $3 \times 10^{-2}$  (bone marrow to plasma) for organ elimination rates. This clearly shows that there is a considerable variation in the uptake and elimination rates among the organs. Not every parameter value is estimated with equal accuracy, standard deviations from as low as  $2 \times 10^{-3}$  to as large as 24 per cent occur.

In general lower values are found for the DOL distribution parameters, respectively, from  $1 \times 10^{-6}$  (plasma to bone marrow, -other\_tissues, -heart) to  $2 \times 10^{-2}$  (plasma to liver), and from  $2 \times 10^{-6}$  (other\_tissues to plasma) to  $2 \times 10^{-2}$  (spleen to liver). The accuracies, from 0.5 to >1000 per cent, are much poorer than when parent drug is involved.

The rates of metabolism vary a little bit among organs (0.02 - 0.03) with exceptions for spleen (0.002) and bone marrow (0.006). Accuracies are between 2.5 and 35 per cent. As it improves the modeling results, allowing the possibility of metabolism taking place in plasma is essential. Rapid DAU metabolism in blood has been described for human patients [Benjamin et al., 1973] and also is suggested by the observed high initial DOL levels in the rat's plasma. Strangely, quantitatively it hardly plays a role in the model, the rate being  $1 \times 10^{-6} \text{ min}^{-1}$  (though with accuracy poorer than 1000%).

Table 4-18 shows the TCC and SSR of the optimum response of model 05 in each compartment. In general, the TCC values are higher for the parent drug than for the metabolite. Parent drug distribution and elimination are well described by the model and apparently they obey first order linear kinetics. However, possibly the model can still be refined by introducing an enterohepatic recirculation, i.e., DAU might be taken up again, be it with some time delay, from the bile directly into the liver [Tse et al., 1972; Tavoloni and Guarino, 1982; Steiner et al., 1982; Molino et al., 1986]. Raised datapoints around the eight hour time point in the liver, where the present model response decreases monotonously, seem to support this possibility. Visual inspection of the plotted results (Fig. 4.35) indicates that DAU datapoints are fitted better than DOL datapoints. Especially the DOL plasma response, slowly rising and falling, does

Fig. 4.35 Parent drug (DAU) and Metabolite (DOL) concentration-time curves for Plasma and Liver; Optimum response with estimated standard deviation (shaded area) for the best of the tested 2x7 compartment models (7b1); observations are shown with their standard deviations. Peak of DOL in liver is predicted too early, in plasma too low: an effect of ignored saturation? Is rise in liver DAU data at about 400 min due to enterohepatic recirculation? DAU responses fit better than DOL responses do



not correspond to the sharp drop from high initial values that the datapoints exhibit. To a lesser extent the DOL heart response also is delayed with respect to the datapoints. These goodness of fit considerations are also reflected by the SSR values.

**4.2.4.4 Conclusions.** The small scale models do not yield much pharmacokinetic information. Clinically the model is not useful either, as it does not distinguish between relevant tissues. The large scale models allow realistic incorporation of physical phenomena. Therefore, they are useful to investigate basic pharmacokinetic concepts in detail:

- 1) Parent drug distribution and elimination are well described by the model and apparently they obey first order linear kinetics. However, possibly the introduction of an enterohepatic recirculation may still yield an improvement
- 2) The values of the transfer rate constants vary considerably, and so do their estimated accuracies
- 3) Metabolism appears to occur in all organs, including plasma, however at significantly different rates
- 4) The metabolite part of the model response does not explain the datapoints very well, especially for plasma. Probably the model structure is not perfect yet, and nonlinearities should replace the assumed linear first order diffusion kinetics. However, it may still be possible that, searching in the 45 dimensional parameter space by a gradient method, the computational algorithm has not converged to the global minimum, but to some local minimum instead, so that yet another set of parameter values exists that will yield a better response for the same model structure of model 05. Both possibilities should be further investigated, although this will require a considerable computational effort (cpu time consumption) because of the dimension of the model being rather large.

The unwieldiness of the large scale models reduces their suitability for use in the clinic. At present, the  $7 \times 2$  compartment model may be the most useful, though minimal one for clinical purposes. Important is the ability to predict drug concentrations in a few tissues that are susceptible for intoxication, and at the site of the tumor. Further detailing then is less important. Provided that correct pharmacokinetic concepts are obeyed—best investigated through use of the large models—an intermediately sized model that is easily implemented on smaller computers, may be used in the search for efficient drug administration schemes and for patient monitoring purposes.

## 4.3 REFERENCES

- Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Contr AC19:716-723
- Allen DH (1983) Parameter estimation for nonlinear models with emphasis on compartmental models. Biometrics 39:629-637

- Arcamone F (1981) Doxorubicin Anticancer Antibiotics. Academic Press, New York
- Baurain R, Deprez-de Campeneere D and Trouet A (1979) Determination of daunorubicin, doxorubicin, and their fluorescent metabolites by high pressure chromatography. *Cancer Chemother Pharmacol* 2:11-14
- Benjamin RS, Riggs Jr CE and Bachur NR (1973) Pharmacokinetics and metabolism of adriamycin in man *Clin Pharmacol Ther* 14:592-600
- Bevington PhR (1969) Data reduction and error analysis. McGraw Hill, New York
- Bischoff KB and Dedrick RL (1968) Thiopental pharmacokinetics. *J Pharm Sci* 57(8):1346-1351
- Blum RH and Carter SK (1974) Adriamycin: a new antitumour drug with significant clinical activity. *Ann Intern Med* 80:249-259
- Carson ER, Cobelli C and Finkelstein L (1981) The identification of metabolic systems, a review. *Am J Physiol* 240:R120-R129
- Chan KK, Cohen JL, Gross JF, Himmelstein KJ, Bateman JR and Marlis AS (1978) Prediction of adriamycin disposition in cancer patients using a physiologic pharmacokinetic model. *Cancer Treat Rep* 62:1161-1171
- Collins JM, Dedrick RL, King FG, Speyer JL and Myers CE (1980) Nonlinear pharmacokinetic models for 5-fluorouracil in man. *Clin Pharmacol Therapeutics* 28:235-246
- Eksborg S, Stendahl U and Loenroth U (1986) Comparative pharmacokinetic study of adriamycin and 4'epi-adriamycin after their simultaneous intravenous administration. *Eur J Clin Pharmacol* 30:629-631
- Erickson P and Ackerman E (1986) Effects of sampling location on data sensitivity in a three-compartment model. *Comput Biol Med* 16:195-204
- Eijkhoff P (1974) System identification: parameter and state estimation. Wiley, London
- Freireich EJ, Gehan AE, Rall DP, Schmidt LH and Skipper HE (1966) Quantitative comparison of toxicity of anticancer agents in mouse, rat, dog, monkey and man. *Cancer Chemother Rep* 50:219-244
- Garrett ER (1980) Biological responses and pharmacokinetics (part 1), *Pharm Internat*, June 1980:121-125
- Gill PE, Murray W and Wright MH (1981) Practical optimization. Academic Press, London
- Godfrey K (1983) Compartmental models and their application. Academic Press, London
- Golub GH and Reinsch C (1970) Singular value decomposition and least squares solutions. *Numer Math* 14:403-420
- Gottlieb AJ, Weinberg V, Ellison RR, Henderson ES, Terebello H, Rafla S, Cuttner J, Silver RT, Garey RW, Levy RN, Hutchinson JL, Raich P, Cooper MR, Wiernick P, Anderson JR and Holland JF (1984) Efficacy of daunorubicin in the therapy of adult acute lymphocytic leukemia. *Blood* 64:267-274
- Johansen PB, Jensen SE, Rasmussen SN and Dalmark M (1984) Pharmacokinetics of doxorubicin and its metabolite doxorubicinol in rabbits with induced acid and alkaline urine. *Cancer Chemother Pharmacol* 13:5-8
- Kwakernaak H and Sivan R (1972) Linear optimal control systems. Wiley Interscience, New York
- McVie JG (1984) The role of pharmacokinetics in (combination)chemotherapy. *Cancer* 54:1175-1178
- Metzler CM (1971) Usefulness of the two-compartment model in pharmacokinetics. *J Am Stat Assoc* 66:49-52
- Michaelis L and Menten ML (1913) Der Kinetik der Invertinwirkung. *Biochem Z* 49:333-369
- Molino G, Hofmann AF, Cravetto C, Belforte G and Bona B (1986) Simulation of the metabolism and enterohepatic circulation of endogenous chenodeoxycholic acid in man using a physiological pharmacokinetic model. *Eur J Clin Invest* 16:397-414
- Nooter K, Sonneveld P, Martens ACM, Hagenbeek A and Schultz FW (1986) Tissue distribution and myelotoxicity of daunomycin in the rat: rapid bolus injection vs continuous infusion. *Eur*

- J Cancer Clin Oncol 22:801-806
- O'Bryan RM, Baker LH, Gottlieb JE, Rivkin SE, Balcerzak SP and Grummet GN (1977) Dose-response evaluation of adriamycin in human neoplasia. *Cancer* 39:1940-1948
- Peper A, Grimbergen CA, Kraal JW and Engelbart JH (1987) An approach to the modeling of the tolerance mechanism in the drug effect. I: The drug effect as a disturbance of regulations. *J Theor Biol* 127:413-426
- Press WH, Flannery BP, Teukolsky SA and Vetterling WT (1985) Numerical recipes: the art of scientific computing. Cambridge University Press, New Rochelle, NY
- Rossum JM van, Lingen G van and Maes RAA (1988) Analysis of the pharmacokinetic body transport function. *Mathl Comput Modell*, 11:610-614
- Schmidt R (1982) Advances in nonlinear parameter optimization. Lecture notes in control and information sciences, 37, Balakrishnan AV and Thoma M (eds.) Springer Verlag, Berlin
- Schultz FW, Nooter K, Sonneveld P, Mulder JA (1985) Modelling of the in vivo distribution dynamics of daunomycin, an anti-cancer drug. In: Barker HA, Young PC (eds) Identification and system parameter estimation 1985, 2, Proc. 7th IFAC/IFORS Symposium; Pergamon Press, Oxford, 1391-1396
- Schultz FW (1987) Computersimulaties met een multi-compartimenten model voor farmacokinetiek: onderzoek naar mogelijke proefdierbesparing. Rapport, Radiobiologisch Instituut TNO (Feb. 1987)
- Schultz FW, Mulder JA (1988a) Data-sensitivity of estimated parameters in a seven-compartment model for pharmacokinetics. Proc. 6th Internat'l Conference on Mathematical Modelling in Science and Technology, 1987; Rodin EY, Avula XJR (eds) *Mathl Comput Modelling* 11:786-791.
- Schultz FW, Mulder JA, Sonneveld P (1988b) On the identification of a middle-large dynamic system for in vivo drug distribution and metabolism: Comparison of three optimization routines. In: Han-Fu Chen (ed) Identification and Parameter Estimation, Preprints of the 8th IFAC/IFORS Symposium, vol 3, Pergamon Press, Oxford, 1747-1752
- Schweppe FC (1973) Uncertain dynamic systems. Prentice Hall Inc., Englewood Cliffs, NJ
- Sonneveld P (1980) Pharmacokinetics of adriamycin in the rat. Thesis. Radiobiological Institute TNO, Rijswijk, The Netherlands.
- Sonneveld P and Mulder JA (1981) Development and identification of a multicompartment model for the distribution of adriamycin in the rat. *J Pharmacokin Biopharm* 9:577-601
- Steiner JL, Plusquellec Y, Guillaume A and Boisvieux JF (1982) A timelag model for pharmacokinetics of drugs subjected to enterohepatic circulation. *J Pharm Sci* 71:297-302
- Swan GW (1985) Optimal control applications in the chemotherapy of multiple myeloma. *IMA J Math Appl Med Biol* 2:139-160
- Tavoloni N and Guarino AM (1982) Does adriamycin undergo an enterohepatic circulation? *Proc Soc Experim Biol Med* 169:4-46
- Testa B (1987) Pharmacokinetic and pharmacodynamic events: can they always be distinguished? *TIPS* 8:381-383
- Tse FLS, Ballard F and Skinn J (1982) Estimating the fraction of reabsorbed drugs undergoing enterohepatic circulation. *J Pharmacokin Biopharm* 10:455-461
- Villani F, Comazzi R, Genitoni V, Lacaita G, Guindani A, Crippa F, Monti E, Piccinini F, Rossa A, Lanza E, et al. (1985) Preliminary evaluation of myocardial toxicity of 4'-deoxydoxorubicin; experimental and clinical results. *Drugs Exp Clin Res* 11:223-231
- Wagner JG (1975) Fundamentals of clinical pharmacokinetics. Hamilton Press (Drug Intelligence Publication), Hamilton Ill.
- Åström KJ (1979) Maximum likelihood and prediction error methods. In: Isermann R. (ed.) Tutorials on system identification. Technische Hochschule Darmstadt, p53-67

## Chapter 5

### Analysis of DNA Histograms<sup>1</sup>

[Identification of a Flow Cytometer System for Cell Kinetics]

#### 5.1 INTRODUCTION TO DNA HISTOGRAM ANALYSIS

##### 5.1.1 Purpose of Flow Cytometric Data (DNA histogram) Analysis

Most chemotherapeutic agents that are used for cancer treatment interfere with a cell population's proliferative activity. For instance, the cytostatic drug arabinoside-cytosine (ara-C) causes accumulation of cells in the S-phase of the cell cycle, possibly because of a (partial) block of the S-G<sub>2</sub> transition [Karon and Shirakawa, 1970]. The precise cytokinetic action of a drug can be evaluated experimentally by exposing the cell population to the drug for some lengths of time, and by studying the perturbations of the cell cycle, as compared to an untreated population. This implies monitoring of the time courses of various cell kinetic parameters for both the treated and the control populations, like for instance the relative numbers of cells that are in the G<sub>1</sub>-, S-, G<sub>2</sub>- and M-phases of the cell cycle. Constructing (sequential) DNA histograms is one possible technique that serves this purpose (e.g., [Barlogie, 1976; Gray, 1974]).

The construction of a DNA histogram involves counting all the cells in the population that, at the given time instant, contain some, specified amount of DNA. The histogram displays the cell numbers as function of DNA contents. Thus, it yields information on the population's composition at the sampling time, though no distinction can be made between G<sub>2</sub> and M cells. For, each cell in the G<sub>1</sub>-phase of the cell cycle contains 2C DNA, cells in the S-phase possess an amount of DNA between 2C and 4C, and those cells that are in either the G<sub>2</sub> or the M-phase contain 4C DNA (2C corresponds with the amount of DNA in a complete set of chromosomes). So, for instance, the ratio of the counted number of 2C DNA cells and the total number of cells that form the population is identical with the fraction of G<sub>1</sub> cells. How this fraction changes with time can be seen by comparing several, subsequent DNA histograms.

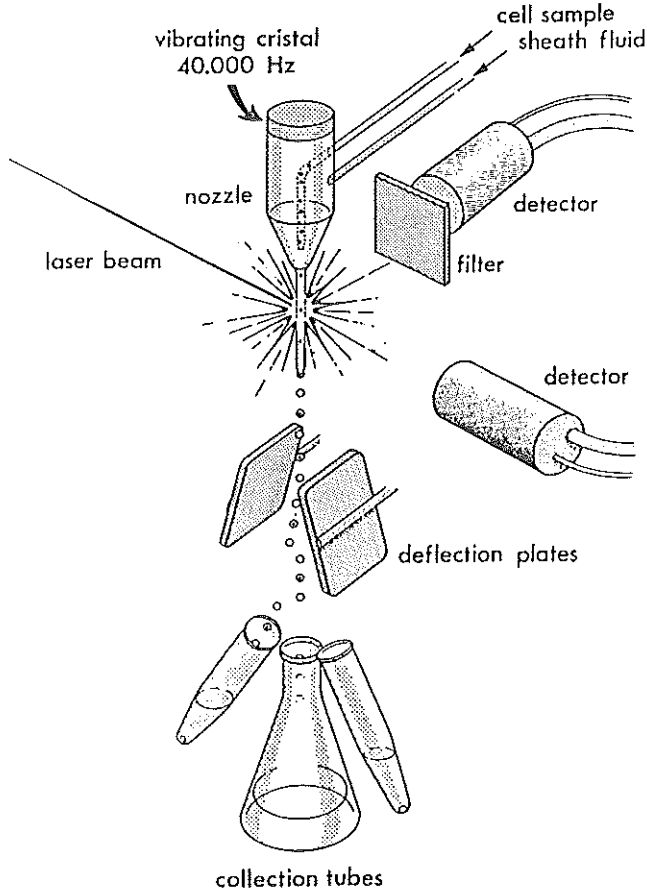
These types of studies lead to conclusions whose reliability greatly depends on how well the appropriate cell kinetic parameters can be determined. A flow cytometer (FCM) is a convenient instrument for measuring DNA histograms. Unfortunately, it is not possible to read the population's composition

---

<sup>1</sup>Parts of Chapter 5 were published before [Schultz et al., 1985; Sonneveld et al., 1988]. The cooperation of prof. dr. H.Baisch of the University of Hamburg and dr. C.B.Bagwell of the University of Miami is gratefully acknowledged.

Fig. 5.1 Major components of a flow cytometer [Trask, 1985]

Detectors register the intensity of light, scattered forwards and sideways when a cell passes through the laser beam; wavelengths to be picked up can be set by filters. A different (or no) cell (type) in a droplet corresponds with a different light scatter pattern. A second function of the FCM is sorting; the electrostatically charged stream of droplets can be continuously deflected towards certain collection tubes, depending on the measured light scatter patterns (FACS, fluorescence activated cell sorting)



directly from such a measured histogram. Some data reduction, or histogram-analysis, is required first. Therefore, an accurate and objective DNA histogram analysis method is a tool of importance.

### 5.1.2 Flow Cytometric Data Acquisition

The process of constructing a DNA histogram can be performed quickly by



means of a flow cytometer (FCM). This instrument can automatically separate, classify and count a population's cells, at rates of a few thousands per second, on the basis of their DNA contents. Or rather based on the intensity of the fluorescent light emission upon laser excitation, which is linearly related to the DNA contents. A detailed description of the principles of the FCM can be found in e.g., [Bagwell, 1979] or [Fulwyler, 1980]. Briefly, the FCM produces a DNA histogram in the following way.

A DNA specific, fluorescent dye (often propidium iodide, PI) is added to the solution in which the population is incubated for some time. The dye molecules attach to the strands of DNA present in the cells. The more DNA a cell contains, the more dye it takes up. The cells are washed and, in suspension, run through the FCM (Fig. 5.1). One after the other, in a stream of droplets, the cells pass through a beam of laser light whose wavelength has been chosen such, that the dye molecules are excited to emit light (fluorescence). The intensity of the event is measured by photo sensors that convert their input into an electrical signal of corresponding magnitude. This, in turn, determines to what class of DNA contents, or channel, the cell is allocated. The channels are numbered 1 through 256 (typically; higher resolutions, e.g., 512 or 1024 channels are also possible) for increasing DNA contents per cell. A count is kept, so eventually the measured histogram, [number of cells] vs [channel number], is obtained.

### 5.1.3 Methods of FCM Data Analysis

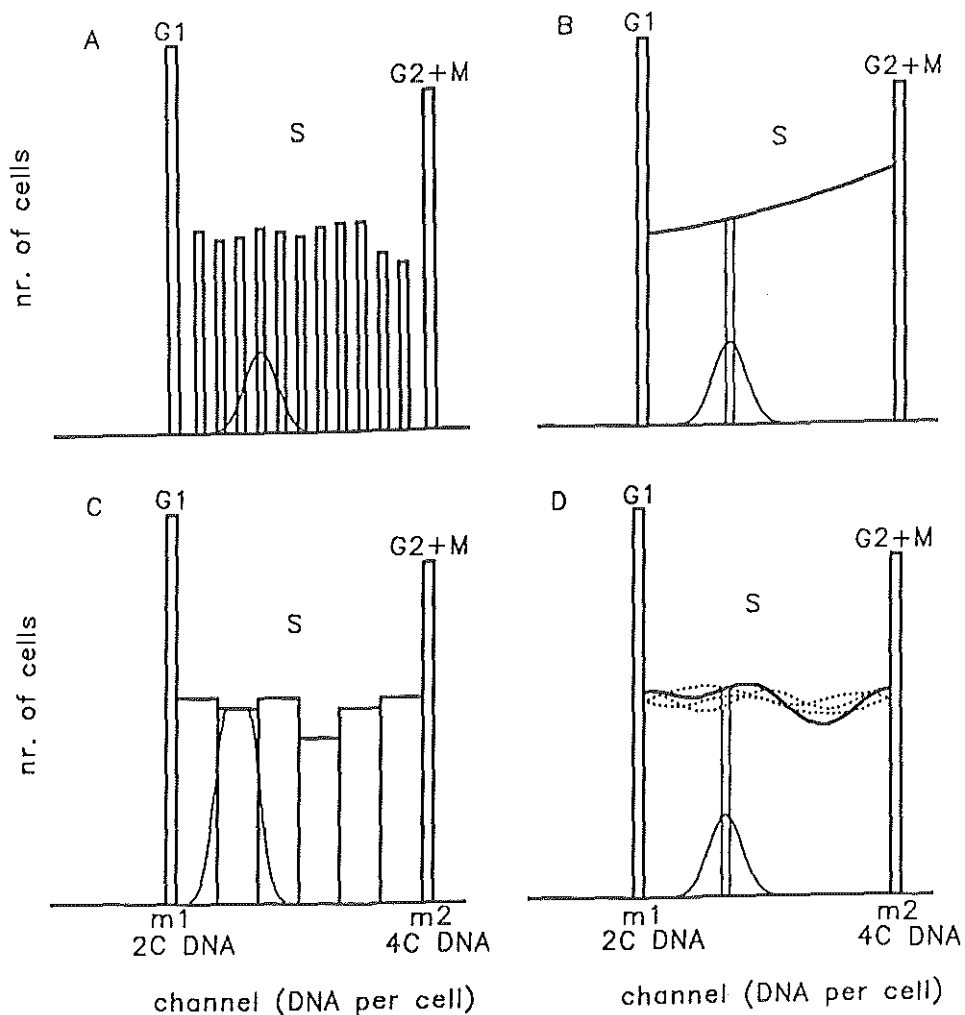
The DNA histogram as produced by the FCM is not suited for direct extraction of the correct percentages of  $G_1$ , S and  $G_2M$  cells. For various reasons the FCM may put a cell into a wrong channel.

A not exactly linear relationship between dye uptake and measured fluorescence may be one error source, though within a certain range linearity can be expected [Bagwell, 1979]. Another error source may be the fact that cells that contain equal amounts of DNA need not necessarily take up equal amounts of fluorescent dye. Due to Brownian movement of the DNA strands dye molecules may miss a potential site of attachment. Therefore, dye uptake can be considered as a random process with a Gaussian probability density function.

Furthermore, cellular debris may cause false registration of an event (background noise) and sometimes two or three cells stick together, forming doublets and triplets that are mistaken for a single event. For example, a doublet of two  $G_1$  cells is easily registered as one  $G_2$  or M cell. To some extent corrections can be made for these types of errors (see below).

Next to these error sources inherent to biochemistry and biophysics, instrumental errors may have their origin in the electronic and optical parts of the FCM, e.g., drift in amplifiers, laser light fluctuation, etc.

Fig. 5.2 Kernel functions  $Z_T(t_i)$ , "true" nr. of cells as function of DNA contents,  $t_i$  (between 2C in channel  $m_1$  and 4C in channel  $m_2$ ), as used in different DNA histogram analysis programs. a) Fried's multiGaussian deconvolution technique, b) Dean/Jett's polynomial curve fitting technique, c) Bagwell's multirectangle technique, d) Present multiharmonic technique



So, data reduction is necessary to extract the true DNA histogram from the measured one. Several DNA histogram analysis methods have been designed to do this. Their principles range from relatively simple graphical methods (e.g., [Baisch, 1975]) to more complicated mathematical procedures that require nu-

merical processing on a digital computer. Examples of the latter are the multiGaussian deconvolution technique [Fried, 1976; Fried and Mandel, 1979], the polynomial curve-fitting technique [Dean and Jett, 1974] and the multirectangular technique [Bagwell, 1979].

Not all techniques exhibit a good performance under all circumstances, nor are they always successfully applicable to all the different classes of DNA histograms (i.e., derived from synchronously or asynchronously growing populations, with very large or small G<sub>2</sub>M fractions, etc.). The graphical methods are quick and simple, but not very accurate. Furthermore, their results depend on the subjective decisions of the analyst. Computerized methods are more objective, but also more complicated. The three numerical methods mentioned have in common that they choose a so-called kernel function to represent the true DNA histogram whose shape is still unknown. This kernel function is broadened to allow for the dispersing effects of erroneous channel allocation occurring in the FCM, i.e., Gaussian curves are superimposed, as the variety and the nature of the error sources make plausible the assumption of Gaussian distributed random errors. Subsequently, using a non-linear least squares method, the broadened kernel function is fitted to the measured DNA histogram.

Fried's *multiGaussian* deconvolution technique chooses as kernel function several spikes (Fig. 5.2a). The spike heights,  $Z_T[i]$ , correspond with the number of cells and the spike location,  $t_i$ , is related to the mean DNA contents per cell. The spikes on the left ( $m_1$ ) and right ( $m_2$ ) flanks represent the G<sub>1</sub> and G<sub>2</sub>M cells respectively, the spikes in between represent the S-phase cells in the population. Thus, for instance, the G<sub>1</sub> fraction is equal to  $Z_T[m_1]/(\sum_{t_i} Z_T[t_i])$ . Each spike is broadened with a Gaussian curve, e.g., for the  $i^{\text{th}}$  spike (mean  $t_i$ , standard deviation  $sd_{t_i}$ ):

$$G(t; t_i, sd_{t_i}) = (2\pi)^{-1/2} \cdot sd_{t_i}^{-1} \cdot \exp\left\{-\frac{1}{2} \frac{(t-t_i)^2}{sd_{t_i}^2}\right\}. \quad (5.1)$$

The resulting function,  $Z_p(t)$ , giving the population's DNA density distribution:

$$Z_p(t) = \sum_{t_i} Z_T[t_i] \cdot G(t; t_i, sd_{t_i}), \quad (5.2)$$

is fitted to the measured DNA histogram, optimizing the free parameter values for  $m_1$ ,  $sd_{m1}$ ,  $m_2$ ,  $sd_{m2}$ , the distance between (i.e., the number and positions of) the spikes, and the spike heights  $Z_T[t_i]$ . It is assumed that the standard deviation of a Gaussian curve depends on the location of its mean, according to:

$$sd_{t_i} = \frac{sd_{m1}}{m_1} \cdot t_i, \quad (5.3)$$

as in the FCM cells with higher DNA contents appear to experience dispersion over a wider range of channels than cells with lower DNA contents do.

A true DNA distribution of S cells should be a continuous one, so obviously Fried's kernel function is not very realistic. Furthermore, a distinct separation between G<sub>1</sub> and early S, as well as between late S and G<sub>2</sub>M cells is not possible. Increasing the number of spikes to reduce this disadvantage leads to severe numerical problems, as it becomes difficult to determine the increased number of free parameters all with sufficient accuracy.

Dean/Jett's *polynomial* technique takes as kernel function (Fig. 5.2b) two spikes, for the G<sub>1</sub>, respectively, the G<sub>2</sub>M cells, while the S-phase distribution in between is simulated with a 2<sup>nd</sup> degree polynomial:

$$P(t) = a + b \cdot t + c \cdot t^2, \quad (m_1 < t < m_2), \quad (5.4)$$

or, in discrete form, the number of cells in channel i, Z<sub>T</sub>[i] becomes—by integration of Eq.(5.4) from t = i-1 to t = i:

$$Z_T[i] = K + L \cdot i + M \cdot i^2, \quad (m_1 < i < m_2), \quad (5.5)$$

where M = c, L = -(c - b) and K = (c/3 - b/2 + a).

Both spikes and polynomial are broadened with Gaussian curves, resulting in the function

$$Z_p(t) = Z_T[m_1] \cdot G(t; m_1, sd_{m1}) + Z_T[m_2] \cdot G(t; m_2, sd_{m2}) + \sum_{t_i} P(t_i) \cdot G(t; t_i, sd_{t_i}). \quad (5.6)$$

The nine parameters Z<sub>T</sub>[m<sub>1</sub>], Z<sub>T</sub>[m<sub>2</sub>], m<sub>1</sub>, m<sub>2</sub>, sd<sub>m1</sub>, sd<sub>m2</sub>, a, b, and c are optimized to fit this function to the measured histogram.

Application of the polynomial technique assures continuity of the S-phase distribution, yet requires few free parameter values to be estimated. But using a 2<sup>nd</sup> degree polynomial may be too stringent a restriction as to turn out a satisfactory fit, especially when drug-perturbed populations are concerned. Trying to increase the kernel's flexibility by allowing a polynomial of higher degree does not improve accuracy, as using more parameters leads to more uncertainty, especially at the polynomial's endpoints, that coincide with the G<sub>1</sub>-early S and late S-G<sub>2</sub> areas.

Bagwell's *multirectangle* technique combines both methods, in that his

kernel function (Fig. 5.2c) approximates the S-phase distribution with a series of contiguous rectangles whose heights, as for Fried's spikes, are allowed to vary independently. Adjacent spikes for  $G_1$  and  $G_2M$  cells are broadened with Gaussian curves and each rectangle is convolved with a Gaussian curve, yielding

$$Z_p(t) = Z_T[m_1] \cdot G(t; m_1, sd_{m1}) + Z_T[m_2] \cdot G(t; m_2, sd_{m2}) + \int_{A_i}^{B_i} Z_T[i] \cdot G(x; 0, sd_i) dx, \quad (5.7)$$

where  $A_i = t - \frac{1}{2} \cdot a_i$ ,  $B_i = t + \frac{1}{2} \cdot a_i$ ,  $a_i$  = width of the  $i^{\text{th}}$  rectangle, and  $sd_i = PC \cdot t_i + CF$  expresses the assumed linear increase of dispersion effects with augmented DNA contents per cell.

$Z_p(t)$  is fitted to the measured DNA histogram by optimizing the parameter values  $m_1$ ,  $m_2$ ,  $PC$ ,  $CF$ ,  $a_i$ 's (number and widths of rectangles) and  $Z_T[i]$ 's (heights of spikes and rectangles).

Thus, Bagwell's approach combines good flexibility for relatively few parameters, and S-phase continuity. Another method of mathematical analysis, which is proposed below, puts emphasis on the same characteristics. This method, called *multiharmonic* technique, comes closer to physical reality, as it tries to describe the process taking place in the FCM. The approach is based on principles originating in the field of general systems theory (e.g., [Eijkhoff, 1974]): a system identification sequence comprises the search for an appropriate set of mathematical formulas, with which the physical reality is described sufficiently well to explain and understand (and predict the future behavior of) the system under investigation. In other words, here, it is attempted to build a mathematical model that simulates the physical processes in the FCM that lead to the measured histogram.

At present, modeling is started with assuming a theoretical histogram, or, equivalently, kernel function, from which the population's composition can be derived at once. A  $G_1$  fraction is assumed and positioned in an appropriate DNA channel, and so is a  $G_2 + M$  fraction. The remaining cells are assumed to be in S-phase. Their true, continuous distribution is approximated with a set of harmonic functions, i.e., a sum of sines and cosines terms (Fig. 5.2d). The advantage is great flexibility, as in this way a large variety of shapes can be formed to simulate the true distribution, while employing only a few free parameters. Table 5-1 compares the number of parameters required by the different methods.

Next, the physical process of dispersion is modeled. It is assumed that errors in channel allocation by the FCM, causing cells that belong to channel  $j$

TABLE 5-1 PARAMETERS IN THE VARIOUS DNA HISTOGRAM ANALYSIS METHODS

| METHOD:                         | multiGaussian  | polynomial   | multirectangular  | multiharmonic   |
|---------------------------------|--|--|---|---|
| <i>peak location:</i>           |  |  |   |   |
| G <sub>1</sub>                  | m <sub>1</sub>   | m <sub>1</sub>   | m <sub>1</sub>  | m <sub>1</sub>  |
| G <sub>2</sub> M                | m <sub>2</sub>   | m <sub>2</sub>   | m <sub>2</sub>  | (m <sub>2</sub> = 2 · m <sub>1</sub> )  |
| nr. of parameters               | 2  | 2  | 2   | 1   |
| <i>kernel:</i>                  |  |  |   |   |
| G <sub>1</sub>                  | Z <sub>T</sub> [m <sub>1</sub> ]                           | Z <sub>T</sub> [m <sub>1</sub> ]                           | Z <sub>T</sub> [m <sub>1</sub> ]                              | Z <sub>T</sub> [m <sub>1</sub> ]  |
| G <sub>2</sub> M                | Z <sub>T</sub> [m <sub>2</sub> ]                           | Z <sub>T</sub> [m <sub>2</sub> ]                           | Z <sub>T</sub> [m <sub>2</sub> ]                              | Z <sub>T</sub> [m <sub>2</sub> ]  |
| S                               | nr. of spikes<br>position of spikes<br>spike heights       | a<br>b<br>c<br>(p = a + b · t' + c · t' <sup>2</sup> )     | nr. of rectangles<br>width of rectangles<br>rectangle heights | nr. of harmonics<br>cosine coefficients, A[k]<br>sine coefficients, B[k] k = 1..N<br>(h = A[0] + 2 · ∑ A[k] · cos(kwt') +<br>B[k] · sin(kwt');<br>A[0] = {NC - Z <sub>T</sub> [m <sub>1</sub> ] - Z <sub>T</sub> [m <sub>1</sub> ]} / M,<br>w = 2π / ((m <sub>1</sub> - 1)) |
| nr. of parameters               | 3 + 2N   | 5  | 3 + 2N - 1  | 3 + 2N  |
| <i>dispersion:</i>              |  |  |   |   |
| G <sub>1</sub>                  | sd <sub>m1</sub>   | sd <sub>m1</sub>   | PC  | PC  |
| G <sub>2</sub> M                | sd <sub>m2</sub>   | sd <sub>m2</sub>   | CF  | CF  |
| S                               | sd <sub>t</sub> = (sd <sub>m1</sub> / m <sub>1</sub> ) · t | sd <sub>t</sub> = (sd <sub>m1</sub> / m <sub>1</sub> ) · t | sd <sub>t</sub> = PC · t + CF                                 | sd <sub>t</sub> = PC · t + CF   |
| nr. of parameters               | 2  | 2  | 2   | 2   |
| <i>total nr. of parameters:</i> | 7 + 2N   | 9  | 6 + 2N  | 6 + 2N  |

histogram of NC cells in M channels

N = nr. of spikes or nr. of rectangles or nr. of cosine coefficients

in S-phase: Z<sub>T</sub>(t) = p or h for polynomial or multiharmonic method, respectively

t is the measure of fluorescence intensity (DNA contents per cell); t' = t - m<sub>1</sub>

to turn up in channels  $j$ ,  $j\pm 1$ ,  $j\pm 2$ , etc., occur at random according to a Gaussian probability density function. This adds two more parameters to the model.

By substituting the "best" values for the free parameters in the model equations, the model output, or predicted histogram, can be calculated. If the way of modeling has been correct, this predicted histogram should show an optimum degree of accordance with the actually measured histogram. If not, obviously the model structure ought to be improved. Parameter estimation techniques deal with the problems of finding the correct parameter values.

The proposed multiharmonic method is discussed in more detail below. Subsequently, examples of its application will be given.

## 5.2 THE MULTIHARMONIC TECHNIQUE

### 5.2.1 Construction of the Mathematical Model

Let the population under investigation consist of a total of  $NC$  cells. The population is assumed to be homogeneous, i.e., it consists of cells of a single cell line only. This restriction is necessary, because  $G_1$  cells—and other cells of corresponding maturity—from different cell lines may contain different amounts of DNA (different value of  $C$  in the  $2C$  DNA quantity); thus, a heterogeneous mixture will yield a DNA histogram of (partially) overlapping distributions. They are difficult to decompose, although new techniques are being developed (See Chapter 7.4.1). By means of an FCM, that can distinguish  $M$  channels or classes of DNA contents, a measured DNA histogram has been produced:  $Z_M[i]$  cells of a single type have been counted in channel  $i$ . No cells are lost in the process:  $\sum_i Z_M[i] = NC$ . Scaling is such that cells that have been put into channel  $2\cdot i$  must have twice the nominal DNA contents of cells in channel  $i$ . (This means that in the linear relationship between fluorescence  $F$  and DNA contents  $i$ ,  $F = a\cdot i + b$ , the term  $b$  must be zero or very small compared to the factor  $a$ . For, if  $F_1 = a\cdot i_1 + b$ ,  $F_2 = a\cdot i_2 + b$ ,  $F_2 = 2\cdot F_1$ , then  $i_2 = 2\cdot i_1 + b/a$ ). The problem is to find the unknown true histogram,  $Z_t$ , from the measured one,  $Z_M$ .

**5.2.1.1 Creating the Theoretical Histogram.** According to the proposed systems theoretical approach a model must be built that describes, in mathematical terms, how the measured histogram has been achieved. Therefore, first a theoretical histogram is constructed. Suppose there are  $Z_T[m_1]$   $G_1$  cells in the population. These cells are located in the  $2C$  DNA channel  $m_1$ . Further suppose that the amount of  $G_2M$  cells present comes to  $Z_T[m_2]$ . They must be located in the  $4C$  DNA channel  $m_2 = 2\cdot m_1$ . Thus, the fractions of  $G_1$  and  $G_2M$  cells amount to, respectively:

$$F_1 = \frac{Z_T[m_1]}{NC} \cdot 100\% \quad \text{and} \quad F_2 = \frac{Z_T[m_2]}{NC} \cdot 100\%. \quad (5.8)$$

This leaves  $NC - Z_T[m_1] - Z_T[m_2]$  cells in S-phase (fraction  $F_S = 100 - F_1 - F_2$  %), to be distributed among the channels between  $m_1$  and  $m_2$ . The mathematical expression that describes the true shape of the S-phase distribution curve, is approximated with a set of harmonic functions, in a way as shown below.

**5.2.1.2 Fourier expansion.** Provided that certain conditions (e.g., single valuedness and integrateability) are fulfilled, any periodical function  $Z(t)$  with period  $T$  can be expanded into a converging series of harmonic functions (Fourier expansion, e.g., [Tuma, 1979]):

$$Z(t) = \sum_{k=-\infty}^{\infty} \{A[k] \cdot \cos(k \cdot \omega \cdot t) + B[k] \cdot \sin(k \cdot \omega \cdot t)\}, \quad \omega = 2\pi/T. \quad (5.9)$$

For the amplitudes, or Fourier coefficients,  $A[k]$  and  $B[k]$  the following expressions are valid:

$$A[k] = \frac{1}{T} \cdot \int_0^T Z(t) \cdot \cos(k \cdot \omega \cdot t) dt$$

$$B[k] = \frac{1}{T} \cdot \int_0^T Z(t) \cdot \sin(k \cdot \omega \cdot t) dt$$

$$k = 0, \pm 1, \pm 2, \dots \quad (5.10)$$

The function  $Z(t)$  can be approximated by considering only the first few, say  $N$ , terms of the series. Then,  $Z(t)$  can be written as:

$$Z(t) \approx A[0] + 2 \cdot \sum_{k=1}^N \{A[k] \cdot \cos(k \cdot \omega \cdot t) + B[k] \cdot \sin(k \cdot \omega \cdot t)\}. \quad (5.11)$$

Of course, the larger  $N$  is chosen, the better the approximation. A fairly small  $N$ , however, may already result in a good enough approximation, especially if  $Z(t)$  does not exhibit sharp changes, but fluctuates gradually.



TABLE 5-2 THE THEORETICAL DNA DISTRIBUTION IN THE MODEL: DNA CONTENTS PER CELL = t, NR. OF CELLS = Z<sub>T</sub>(t)

|                         |   |
|-------------------------|---|
| $0 \leq t \leq m_1 - 1$ | $Z_T(t) = 0$  |
| $m_1 - 1 < t \leq m_1$  | $Z_T(t) = Z_T[m_1]$   |
| $m_1 < t \leq m_2 - 1$  | $Z_T(t) = A[0] + 2 \cdot \sum \{ (A[k] \cdot \cos(k \cdot w \cdot (t - m_1)) + (B[k] \cdot \sin(k \cdot w \cdot (t - m_1))) \}; k = 1, N$ |
| $m_2 - 1 < t \leq m_2$  | $Z_T(t) = Z_T[m_2]$   |
| $m_2 < t \leq M$        | $Z_T(t) = 0$  |

NC: given nr. of cells

M: given nr. of channels

$m_1$ : chosen G<sub>1</sub> channel (2C DNA)

N: chosen nr. of Fourier pairs (sin, cos terms)

$$m_2 = 2 \cdot m_1$$

$$T = m_2 - 1 - m_1 = m_1 - 1$$

$$w = 2\pi/T$$

M

$$\int_0^M Z_T(t) dt = NC$$

TABLE 5-3 THE DISCRETIZED THEORETICAL DNA DISTRIBUTION: Z<sub>T</sub>[i] CELLS IN DNA CHANNEL i

| i                           | Z <sub>T</sub> [i]   | $\left\{ \sum_{i=1}^M Z_T[i] = NC \right\}$ |
|-----------------------------|--|---|
| 1 through $m_1 - 1$         | 0  |   |
| $m_1$                       | $Z_T[m_1]$   |   |
| $m_1 + 1$ through $m_2 - 1$ | $(NC - Z_T[m_1] - Z_T[m_2]) / (m_1 - 1) +$ $+ 2 \cdot \sum_{k=1}^N \left\{ \frac{A[k]}{k \cdot w} \cdot \left( 2 \cdot \sin^2 \left( \frac{k \cdot w}{2} \right) \cdot \sin(k \cdot w \cdot (i - m_1)) + \right. \right.$ $\left. \left. + \sin(k \cdot w) \cdot \cos(k \cdot w \cdot (i - m_1)) \right) \right\} +$ $- 2 \cdot \sum_{k=1}^N \left\{ \frac{B[k]}{k \cdot w} \cdot \left( 2 \cdot \sin^2 \left( \frac{k \cdot w}{2} \right) \cdot \cos(k \cdot w \cdot (i - m_1)) + \right. \right.$ $\left. \left. - \sin(k \cdot w) \cdot \sin(k \cdot w \cdot (i - m_1)) \right) \right\}$ |   |
| $m_2$                       | $Z_T[m_2]$   |   |
| $m_2 + 1$ through M         | 0  |   |

### 5.2.1.3 Application of Fourier Expansion to the S-Phase Distribution.

The true S-phase distribution is likely to show gradual fluctuations, and therefore is expected to be approximated with sufficient accuracy if the above mentioned decomposition into N pairs of harmonic functions is carried out. By writing t for DNA contents and  $Z_T(t)$  for the theoretical DNA distribution, the resulting expression is listed in Table 5-2. Table 5-3 shows the same expression for the theoretical distribution, but now converted to a discrete form by calculating the number of cells (nr. of cells) in channel i with:

$$Z_T[i] = \int_{i-1}^i Z_T(t) dt, \quad i = 1, 2, \dots, M. \quad (5.12)$$

The mean DNA contents per cell in channel i comes to:

$$\{2 \cdot (i - 1/2) + 1\} \cdot C/m_1. \quad (5.13)$$

In this way the modeling of the theoretical histogram, shown in Fig. 5.3, has been completed. For given  $m_1$  and N it is fully determined by the  $2 \cdot N + 2$  parameters:  $Z_T[m_1]$ ,  $Z_T[m_2]$ ,  $A[1]$ , ...,  $A[N]$ ,  $B[1]$ , ...,  $B[N]$ .

It should be noted that the first Fourier coefficient,  $A[0]$  alone is responsible for a uniform distribution of S cells among the S channels:

$$A[0] = \frac{1}{T} \cdot \int_0^T Z_T(t) \cdot \cos(0 \cdot w \cdot t) dt = \frac{1}{m_2 - 1 - m_1} \cdot \int_{m_1}^{m_2 - 1} Z_T(t) dt = \frac{NC - Z_T[m_1] - Z_T[m_2]}{m_1 - 1} = \frac{\text{nr. of S cells}}{\text{nr. of S channels}}. \quad (5.14)$$

The additional Fourier terms may alter the shape of the S-phase distribution, but they do not contribute to the area under the curve (nr. of cells):

$$\int_0^T C_a \cdot \cos(k \cdot w \cdot t) dt = \int_0^T C_b \cdot \sin(k \cdot w \cdot t) dt = 0 \quad \text{for any } C_a, C_b, k. \quad (5.15)$$

Therefore, considering more sine and/or cosine terms does not alter the magnitude of the S fraction.

**5.2.1.4 Modeling the Dispersion Effects.** Next, the dispersion occurring in the FCM is modeled. It is assumed that the random channel allocation, by which

cells, that due to their DNA contents belong to channel  $j$ , turn up in channel  $j$  as well as in the channels in the neighborhood of  $j$ , can be modeled with a Gaussian density distribution function, whose parameters are:

$$\text{mean value} = j - 1/2; \quad \text{standard deviation} = PC \cdot (j - 1/2) + CF. \quad (5.16)$$

The latter expression accounts for the linear increase of the dispersion effects with increasing channel number, which is an observed FCM characteristic. Occasionally, however, the dispersion effects appear to increase in proportion to the square root of the channel number, for which phenomenon theoretical support exists. Equation(5.16) should then be modified accordingly. Often the constant factor,  $CF$ , is taken to be zero; a non-zero value, however, is not uncommon (e.g., [Brunsting et al., 1979]). The factor  $PC$  should not be confused with the so-called coefficient of variation,  $CV = \text{sd}_j / j$ , relating the standard deviation to the channel number, although the two variables are nearly equal if  $CF$  is zero.

The dispersion function states how the number of cells,  $Z_T[j]$ , theoretically belonging to channel  $j$ , is redistributed among channel  $j$  and its neighbor channels  $j \pm 1, j \pm 2$ , etc. Channel  $j$  contributes to the number of cells that end up in channel  $jp$  by "losing" a quantity (Fig. 5.3; 5.4):

$$\Delta_j Z_P[jp] = Z_T[j] \cdot \int_{jp-1}^{jp} (2\pi)^{-1/2} \cdot \{PC \cdot (j - 1/2) + CF\}^{-1} \cdot \exp\left\{-1/2 \cdot \frac{(x - (j - 1/2))^2}{(PC \cdot (j - 1/2) + CF)^2}\right\} dx + r_j. \quad (5.17)$$

The first term on the right hand side of this equation represents the Gaussian dispersion, the second term is a rest term that can be explained as follows. In principle the Gaussian function is defined from  $jp = -\infty$  until  $jp = +\infty$ , so  $Z_T[j]$  should be spreaded over this range. However, at a certain channel number  $jp = jc$  the corresponding contribution  $\Delta_j Z_P[jc]$  (without the result added yet) falls below a preset small number  $\text{eps}$  (e.g.,  $\text{eps} = 0.00001$ ). Then further dispersion is stopped. But the remaining tails of the Gaussian should not be ignored, if the number of cells in the population is to be kept constant (especially for larger  $PC$  values negligence will lead to cumulative error and considerable cell loss). Therefore, the tails are divided into equal parts  $r_j$  to be added to the previous channels',  $jp = j, j \pm 1, j \pm 2, \dots, j \pm jc$ , contents. The uniformity of this second redistribution is a matter of convenience; only a very minor error is introduced this way.

The total number of cells that eventually will be found in channel  $jp$

Fig. 5.3 Model of the theoretical distribution:  $Z_T(t)$ , number of cells as function of DNA contents,  $t$ , can be represented by the discrete theoretical histogram  $\{j, Z_T[j]\}$ , or  $Z_T[j]$  cells in DNA channel  $j$ . The total number of cells in the population is given by

$$NC = \int_0^{\infty} Z_T(t) dt = \sum_j Z_T[j].$$

The relations between channel number, mean DNA contents per cell and fluorescence intensity are also shown. Cells with DNA contents  $2C \cdot j/m_1$ , in theory corresponding with channel  $j$ , may be redistributed in surrounding channels, e.g.,  $\Delta_j Z_P[i]$  of these cells will end up in channel  $i$

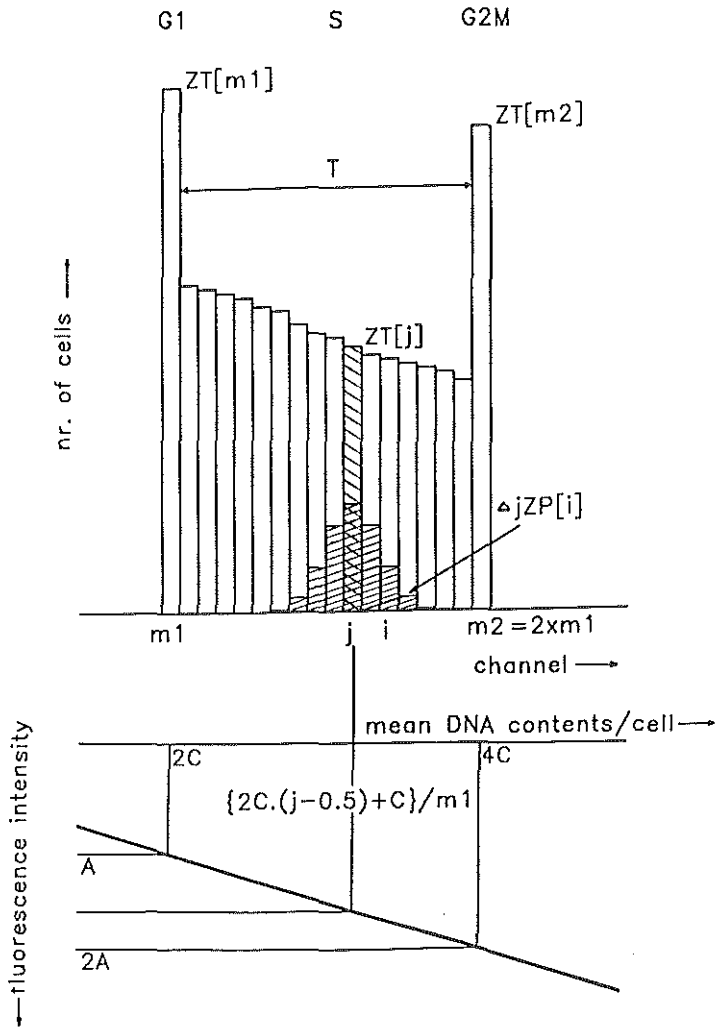
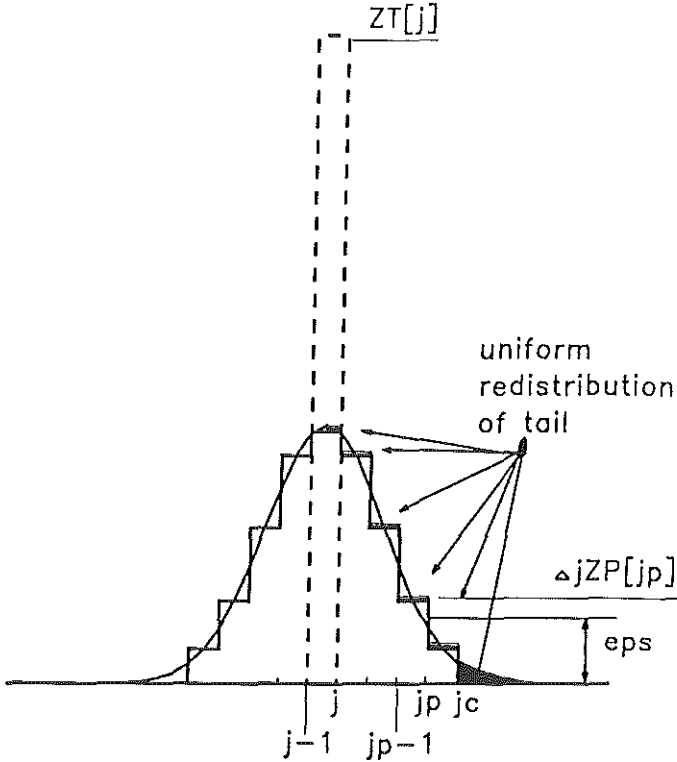


Fig. 5.4 Model of the dispersion:  $Z_T[j]$  cells pertaining to DNA channel  $j$  are redistributed in the surrounding channels, according to a Gaussian probability density function. Tails, beyond channel  $j_c$  for which  $\Delta_j Z_p[j_c] < \text{eps}$  (a preset value), are redistributed once more, now uniformly



amounts to the sum of the contributions that it receives from all channels  $j$  in the theoretical histogram:

$$Z_p[j_p] = \sum_{j=m_1}^{m_2} \Delta_j Z_p[j]. \quad (5.18)$$

Now the full model has been accomplished and a predicted histogram,  $Z_p$ , has been obtained. This predicted histogram is the output of the model that simulates the FCM system. The precise model output depends on the values of the parameters previously mentioned, which determine the theoretical histogram, and on the dispersion parameters PC and CF, which determine how the theoretical histogram is spreaded to yield the predicted histogram. So, given values of  $m_1$

and  $N$ , the output of the mathematical model depends on  $2 \cdot N + 4$  parameters (for  $1 \leq i \leq M$ ):

$$Z_P[i] = f\{Z_T[m_1], Z_T[m_2], PC, CF, A[1], \dots, A[N], B[1], \dots, B[N]\}. \quad (5.19)$$

In this way, four different DNA histograms can be distinguished for the same population:

- 1) the true histogram  $Z_T$ .
- 2) its error-polluted version, produced by the FCM, the measured histogram  $Z_M$ .
- 3) a version, as estimated by the model, the theoretical histogram  $Z_T$ .
- 4) the predicted histogram  $Z_P$ , the model output to be compared to the measured histogram.

Note, that in all four histograms the population size must remain the same, i.e., NC cells.

## 5.2.2 Parameter Estimation

**5.2.2.1 Introduction.** When values are substituted for the parameters in the model equations, a specific model output or predicted histogram  $\{i, Z_P[i]\}$  is obtained. The purpose of a parameter estimation procedure is to find those parameter values that lead to an optimum resemblance of the predicted and actually measured histograms.

It should be noted that even if the model equations, with the proper parameter values substituted, describe the physical system exactly, differences between measurement and prediction are sooner the rule than the exception.

For example, Ohm's law models the relationship of electrical current  $I$  and voltage  $V$  in a resistor  $R$  (units of ampère, volt and ohm, respectively) with a straight line:

$$V = I \cdot R. \quad (5.20)$$

Experiment shows that observed datapoints  $\{I, V\}$  often will not lie exactly on this line, due to the limited accuracy of the measurements caused by imperfection of the instrumentation (and the observers). Yet, for all practical purposes this electrical system is represented very well by the straight line relationship: datapoints will never lie at large distance from the line and, with increasing number of observations, they will—at random—lie above the line as often as below it. In fact, the straight line is supposed to reflect the physical reality perfectly, and the observed deviations are called random measurement errors.

Obviously, if the discrepancies between observed points and optimum model output had appeared to be systematic, and/or (too) large, the postulated

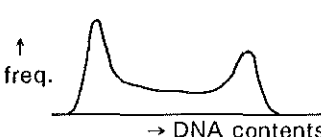
Fig. 5.5 Diagram of the histogram analysis procedure

SYSTEM

INPUT  
 NC cells with different DNA contents marked with some fluorescent dye yield:  
 $Z_T$ , TRUE histogram (not known)

FCM  
 allocates NC cells to M channels based on fluorescence intensity (physical process)

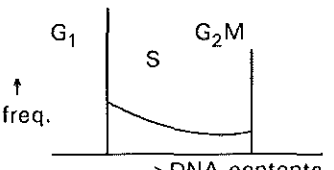
OUTPUT  
 $Z_M$ , OBSERVED histogram



↑ freq.  
 → DNA contents

MODEL

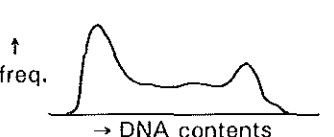
1a INPUT FUNCTION  
 or Kernel Function and Parameter Values determine the distribution of NC cells in:  
 $Z_T$ , THEORETICAL histogram



↑ freq.  
 → DNA contents

1b TRANSFER FUNCTION  
 or Spreading Function and Parameter Values allocate NC cells to M channels (mathematical process)

1c MODEL RESPONSE  
 $Z_P$ , PREDICTED histogram



↑ freq.  
 → DNA contents

to minimize the difference:

- II adjust Parameter Values (Maximum Likelihood criterion) and if necessary,
- III adjust model equations (model matching)

$$\Rightarrow Z_t = Z_T$$

straight line relationship would not have held. In that case the model should have been revised for incorrectness or incompleteness. Finding the correct model (equations) to describe the physical reality is called the process of system identification.

A similar procedure is followed in the present case of modeling the FCM (Fig. 5.5). The *à priori* postulated model is assumed to be correct, i.e., the model equations describe the physical system exactly, once the optimum parameter values have been substituted in them. Then, the calculated model output is identical with the true system output, which differs from the actually observed system output only because (random) measurement errors have occurred. Decisions about acceptance, adaptation or rejection of the model must be made, *à posteriori*, depending on how satisfactorily the model output fits the observed data.

To solve the problem of finding the correct parameter values, in the present case a maximum likelihood (ML) technique is applied [Eijkhoff, 1974; Åström, 1979], which regards the parameters as random variables rather than as unknown constants. Experience with this method was already obtained in solving a similar problem of estimating pharmacokinetic parameters [Sonneveld and Mulder, 1981]. With the ML technique not only a fair estimate of the optimum parameter values can be obtained (it can be proven that the ML estimates equal the true parameter values if the number of observations rises to infinity, which property is called asymptotic unbiasedness), but also an estimate of the errors to be expected in the resulting parameter values can be calculated. The technique is discussed in more detail below.

**5.2.2.2 Maximum Likelihood Optimization.** For a certain set of parameter values, combined in the parameter vector  $p_k$ , the model predicts  $Z_p[i]$  cells to be found in channel  $i$ . The number of cells actually measured in this channel amounts to  $Z_M[i]$  (Fig. 5.6).

$Z_p[i]$  is composed of contributions from the channels  $m_1, \dots, m_2$  in the theoretical histogram, according to Eq.(5.18).

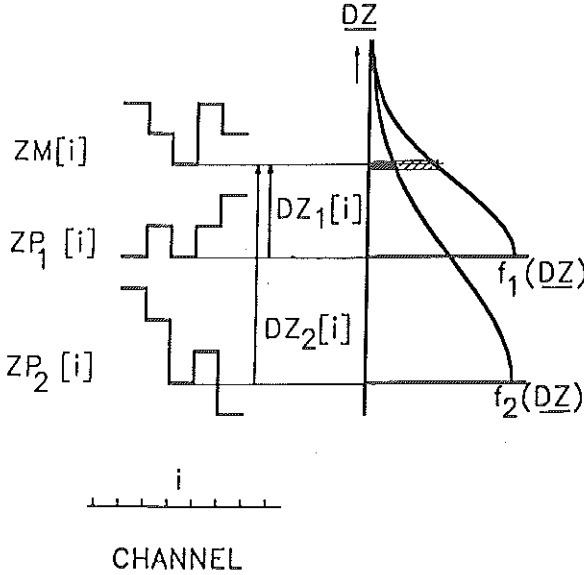
The pure Gaussian dispersion term in Eq.(5.17) is the product of the theoretical number of cells in a channel  $j$ ,  $Z_T[j]$ , and the (Gaussian) probability,  $Pr_{i,j}$ :

$$Pr_{i,j} = \int_{j^{p-1}}^{j^p} (2\pi)^{-1/2} \cdot [PC \cdot (j-1/2) + CF]^{-1} \cdot \exp \left\{ -1/2 \cdot \frac{[x - (j-1/2)]^2}{[PC \cdot (j-1/2) + CF]^2} \right\} dx \quad (5.21)$$

—valid for each of these cells—that a cell from this channel  $j$  is put into channel  $i$ . As such, the term expresses a mean value for  $\Delta_j Z_p[i]$ , the contribution of



Fig. 5.6 Maximum likelihood parameter estimation. Measured system output  $Z_M$ , model output  $Z_P$ , and probability density function  $f_i\{\underline{DZ}\}$  of the measurement error  $\underline{DZ}$  for two ( $j = 1, 2$ ) different sets of parameter values. For channel  $i$ , the probability of the measurement error  $\underline{DZ}[i]$  being approx. equal to the residual  $DZ_j[i] = Z_M[i] - Z_{P_j}[i]$  is:  $Pr\{\underline{DZ}[i] = DZ_j[i]\} = f_j\{DZ_j[i]\} \cdot dDZ_j[i] = \text{shaded area}$ . The greater this probability, the better the model accounts for the observation; in this example, if  $Z_{P_1}$  were true it would make measuring  $Z_M[i]$  more probable than  $Z_{P_2}$  being true. Therefore, to find the optimum parameter values, maximize  $\sum f\{DZ[i]\} \cdot dDZ[i]$ , for all  $i$  simultaneously



channel  $j$  to the predicted number of cells in channel  $i$ . In reality, if the experiment would be repeated with the same cell population and the same model and the same model parameters ( $Z_T[i]$ , PC and CF),  $\Delta_j Z_p[i]$  will assume different values, according to a binomial distribution. In fact, the probability that  $Z_T[j]$  cells in channel  $j$  contribute  $K$  cells to channel  $i$  ( $0 \leq K \leq Z_T[j]$ ) is given by:

$$Pr\{\Delta_j Z_p[i]=K\} = \frac{Z_T[j]!}{K! \cdot (Z_T[j] - K)!} \cdot Pr_{i,j}^K \cdot (1 - Pr_{i,j})^{Z_T[j] - K} \quad (5.22)$$

( $K!$  is pronounced  $K$  factorial and is shorthand for  $1 \cdot 2 \cdot 3 \cdot \dots \cdot K$ ). Thus, the expected value of  $\Delta_j Z_p[i]$  is  $Z_T[j] \cdot Pr_{i,j}$  and its standard deviation is:

$$sd_{\Delta_j Z_p[i]} = \{Z_T[j] \cdot Pr_{i,j} \cdot (1 - Pr_{i,j})\}^{1/2} \quad (5.23)$$

From this equation and Eq.(5.17) it can be seen that  $\Delta_j Z_p[i]$  increases proportionally with  $Z_T[j]$  and  $sd_{\Delta_j Z_p[i]}$  increases proportionally with  $Z_T[j]^{1/2}$ . The

relative deviation  $sd_{\Delta_j Z_P[i]} / \Delta_j Z_P[i]$ , therefore, decreases with increasing  $Z_T[j]$ . If  $Z_T[j]$  is large, the experimental law of large numbers applies, i.e., the actual  $\Delta_j Z_P[i]$  will tend to its expected value of  $Z_T[j] \cdot Pr_{i,j}$  and the relative deviation will tend to zero. This is why—next to the vanishing relative influence of the rest term in Eq.(5.17)—it is important that the cell population run through the FCM is as large as (practically) possible.

From Eq.(5.18) and Eq.(5.23) it is derived that the standard deviation of the predicted number of cells in channel  $i$  will be:

$$sd_{Z_P[i]} = \left\{ \sum_{j=m_1}^{m_2} Z_T[j] \cdot Pr_{i,j} \cdot (1 - Pr_{i,j}) \right\}^{1/2}. \quad (5.24)$$

If the model, thus (mean)  $Z_P[i]$ , were correct, then the residual process error or the difference  $D_Z[i] = Z_M[i] - Z_P[i]$ , would be characterized by the standard deviation in Eq.(5.24) if no further measurement errors have occurred.

Let the residual process error,  $D_Z[i]$  be a realization of the random variable  $\underline{D}_Z[i]$  (measurement error). As has been argued above, the distribution of this random variable will basically be a composition of binomial distributions. For large numbers of cells a binomial distribution can be approximated well by a normal distribution. The sum of several normal distributions is, itself, normally distributed. Therefore, let the random variable be a normally distributed with mean value and variance according to:

$$mean_{D_Z[i]} = 0 \quad \text{and} \quad var_{D_Z[i]} = Q[i]. \quad (5.25)$$

The probability density function of this random variable then is given by:

$$pdf\{\underline{D}_Z[i]\} = (2\pi)^{1/2} \cdot Q[i]^{1/2} \cdot \exp \left\{ -1/2 \cdot \frac{(D_Z[i] - \sum_i D_Z[i])^2}{Q[i]} \right\}. \quad (5.26)$$

This expression holds for all  $i=1,2,\dots,M$ .

Assume independence of  $D_Z[i]$  and  $D_Z[j]$  ( $i \neq j$ ). Further, assume  $Q[i]$  and  $Q[j]$  both equal to  $Q$ . The former assumption will be more reasonable than the latter that is taken as a simplification. The joint probability density function  $pdf\{\underline{D}_Z[i], \underline{D}_Z[j], \dots\}$ , or the likelihood function  $L$ , becomes:

$$L = \prod_{i=1}^M pdf\{\underline{D}_Z[i]\} = (2\pi)^{-1/2 \cdot M} \cdot Q^{-1/2 \cdot M} \cdot \exp \left\{ -1/2 \cdot Q^{-1} \cdot \sum_{i=1}^M D_Z[i]^2 \right\}. \quad (5.27)$$

This function  $L$  reflects the likelihood of measuring the sequence  $Z_M[i]$ ,

( $i=1,2,\dots,M$ ), where the sequence  $Z_p[i]$  should have been measured. The purpose of optimisation, in the ML sense, is to find those parameter values, and corresponding model output, that result in the highest likelihood of the occurrence of the actual observations. In other words, the ML estimates of the parameter values lead to the model output that is the best representation of the true system output, one can achieve with this model.

To find the combination of optimum parameter values, the likelihood function  $L$  must be maximized with respect to the parameters. Instead of maximizing  $L$ , for convenience  $-\ln(L)$  can be minimized for the same results, as the latter function exhibits a monotone decrease with increasing  $L$ .

$$-\ln(L) = \frac{1}{2} \cdot M \cdot \{\ln(2\pi) + \ln(Q)\} + \frac{1}{2} \cdot \sum_{i=1}^M \{D_Z[i] \cdot Q^{-1} \cdot D_Z[i]\}. \quad (5.28)$$

The sum in the last right-hand term is called the performance index. It tends to  $M$ , for  $Q$  can be approximated by:

$$Q \approx \frac{1}{M} \cdot \sum_{i=1}^M D_Z[i]^2 = \frac{SSQ}{M}. \quad (5.29)$$

(SSQ is the sum of the squared errors). Therefore, in turn, minimization of  $-\ln(L)$  is equivalent to minimizing  $Q$ , for instance by finding the simultaneous zero of the derivatives of  $Q$  with respect to the parameters. Hence, set the following gradient (gdt) to zero:

$$gdt = \frac{\partial Q}{\partial p} = 0. \quad (5.30)$$

By using a gradient method the solution of this equation can be numerically searched for. Starting with an initial estimate, parameter vector  $p_0$ , iteratively for  $k = 0, 1, 2, \dots$  a parameter step vector can be calculated according to:

$$\Delta p_k = p_{k+1} - p_k = \theta \cdot \frac{\partial Q}{\partial p_k}, \quad (5.31)$$

until the step  $\Delta p_k$  no longer exceeds a small, preset magnitude  $\delta$ . The last calculated parameter vector  $p_{\text{final}}$ , then is taken as solution, i.e., as optimum estimate of the true combination of parameter values.

The well-known Newton-Raphson algorithm, for instance, uses for  $\Theta$  the Hessian matrix of second order derivatives:

$$H = \left[ \frac{\partial^2 Q}{\partial p \partial p^T} \right]. \quad (5.32)$$

Thus:

$$\Delta p_k = -H^{-1} \cdot \frac{\partial Q}{\partial p_k}. \quad (5.33)$$

However, calculation of the second order derivatives usually is very time consuming. Therefore, it may be advantageous to employ the Gauss-Newton (GN) algorithm that employs the so-called Fisher Information matrix, E:

$$E = \left[ \frac{\partial Q}{\partial p} \cdot \frac{\partial Q}{\partial p^T} \right], \quad (5.34)$$

which involves first order derivatives only. It can be shown that E tends to -H when the number of observations increases.

Explicit expressions for the gdt and the Information matrix can be derived by actually differentiating Q with respect to the parameters. This results in:

$$\frac{\partial Q}{\partial p} = -\frac{2}{M} \cdot \sum_{i=1}^M S[i] \cdot D_Z[i] = -\frac{2}{M} \cdot \sum_{i=1}^M \frac{\partial Z_p[i]}{\partial p} \cdot D_Z[i], \quad (5.35)$$

respectively, with second order terms neglected:

$$E = \frac{2}{M} \cdot \sum_{i=1}^M \left[ \frac{\partial Z_p[i]}{\partial p} \right]^T \cdot \frac{\partial Z_p[i]}{\partial p}, \quad (5.36)$$

after which the parameter step vector can be calculated with:

$$\Delta p_k = E^{-1} \cdot gdt. \quad (5.37)$$

Another characteristic of the Information matrix is that its inverse can serve as a lower bound on the variance-covariance matrix,  $V_p$  of the estimated parameter values,

$$V_p \geq E^{-1}. \quad (5.38)$$

It can be proven that  $E^{-1}$  and  $V_p$  tend to be identical when the number of observations increases to infinity. This is called the property of asymptotic efficiency of the ML estimation procedure. In practice,  $E^{-1}$  is supposed to represent  $V_p$  adequately, even for a limited number of observations. Thus, the standard deviation of the ML estimate of parameter  $j$  is given by:

$$sd_{p_j} = ( E^{-1}[j,j] )^{1/2}, \quad (5.39)$$

and the correlation coefficient for parameters  $j$  and  $l$  comes to:

$$cor(p_j, p_l) = \frac{E^{-1}[j,l]}{( E^{-1}[j,j] \cdot E^{-1}[l,l] )^{1/2}}. \quad (5.40)$$

The variance-covariance matrix of the predicted histogram  $\{i, Z_p[i]\}$  can be calculated next, from:

$$V_{Z_p} = \left[ \frac{\partial Z_p[i]}{\partial p} \cdot E^{-1} \cdot \left( \frac{\partial Z_p[i]}{\partial p} \right)^T \right], \quad (5.41)$$

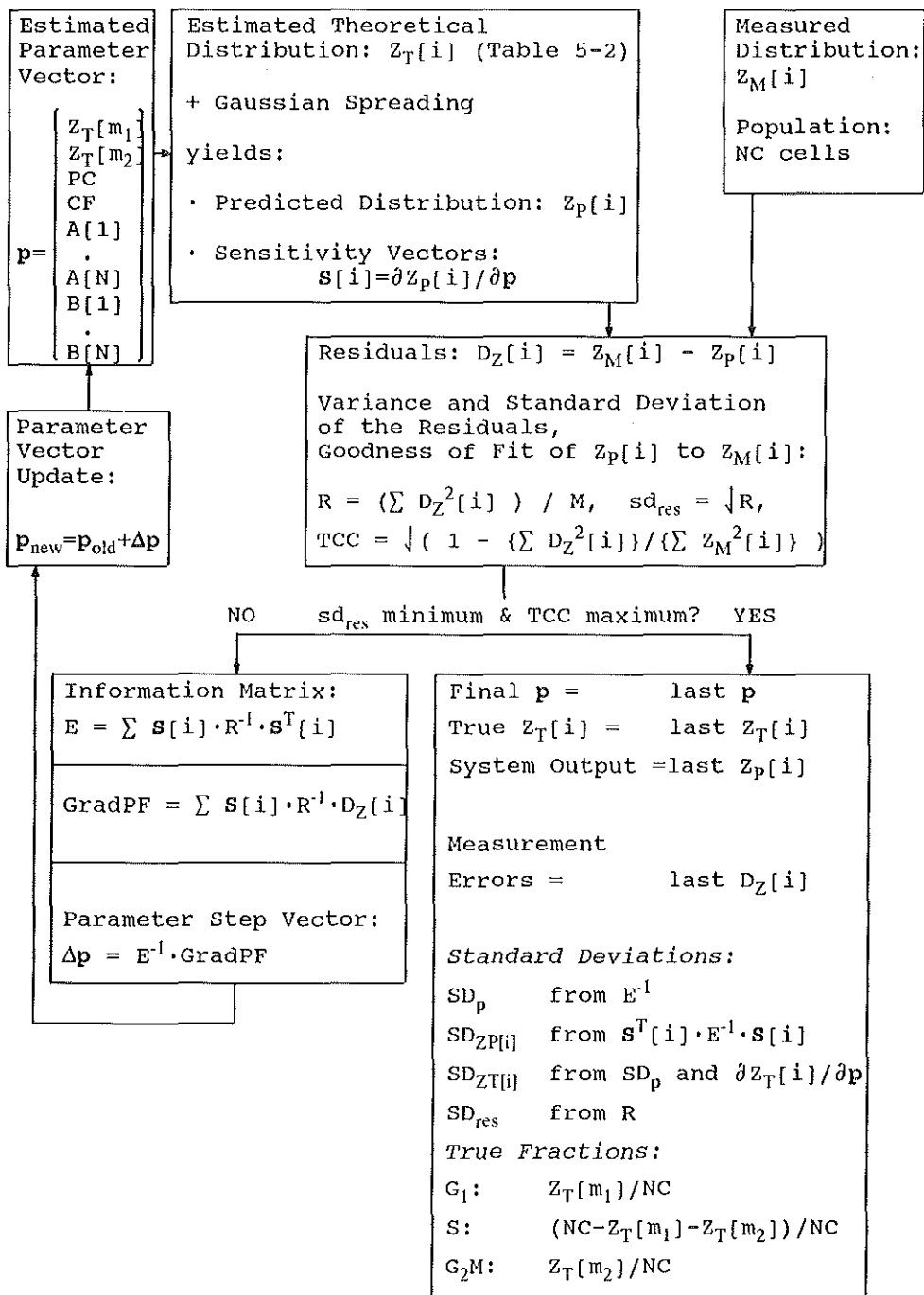
and the standard deviation of the predicted  $Z_p[i]$  becomes:

$$sd_{Z_p[i]} = V_{Z_p}^{1/2}[i,i]. \quad (5.42)$$

A goodness of fit criterion provides the possibility to get an impression of how well the calculated model output fits the observed data. When several curves must be compared with respect to the same set of observations, the total correlation coefficient is a good measure for this purpose. It is defined by:

$$TCC = \left\{ 1 - \frac{\sum_{i=1}^M (Z_M[i] - Z_p[i])^2}{\sum_{i=1}^M (Z_M[i])^2} \right\}^{1/2}. \quad (5.43)$$

TABLE 5-4 COMPUTATIONAL FLOW CHART



For a perfect fit TCC takes on the value of 1, smaller values meaning poorer fit to no fit at all. In practice it was found that for the present case TCC must be greater than 0.985 to indicate a reasonable, graphically confirmed, goodness of fit.

Another measure of goodness of fit is, of course, the standard deviation of the residual process errors,  $sd_{rpe} = (SSQ/M)^{1/2}$  (see Eq.(5.29)).

**5.2.2.3 Computation.** The sequence of calculations necessary to perform the ML parameter estimation procedure as described above is given schematically on the computational flow chart (Table 5-4). Taking this scheme as a guide, a computer program was designed for making DNA histogram analysis runs, using the multiharmonic technique (program HISTANFOU). The program has been written in FORTRAN V, to be used on a DATA GENERAL ECLIPSE minicomputer system, operating under AOS. The program occupies 16K bytes of core memory and needs about 0.5 s CPU time to execute 10 iterations in BATCH mode. (This does not include input file preparation).

Beside the initial estimated parameter values ( $Z_T[m_1]$ ,  $Z_T[m_2]$ , PC, CF, A's and B's) the program asks for an identification name, the total number of channels, the  $G_1$  peak location  $m_1$ , the number ( $N \leq 7$ ) of Fourier coefficient pairs to be considered, and a maximum number of iteration steps allowed. These variables may be given either interactively from the keyboard/screen combination, or they are read from an input file if the program is to be run in batch mode. Program output is written to screen and/or to an output file for each subsequent iteration (new estimate of the parameter values, TCC,  $sd_{rpe}$ ). If desired, plots of the various histograms can be made using a Versatec plotter.

When working interactively, the actual time consumed will often not be less than half an hour, owing to the relatively slow processes of writing to screen and file and reading keyed-in data. It should be noted that for each histogram it may be necessary to run the program a few times with different  $m_1$  and/or  $N$  values to see whether the results improve. A presently available routine that optimizes the  $G_1$  peak channel position  $m_1$  as a free parameter is too time consuming to be of practical use (as  $m_1$  occurs very frequently in the model equations; letting its value float means that many function evaluations must be performed, stretching required computation time beyond 15 minutes per iteration).

### Practical Points

- 1) *INPUT.* The input to the program is the raw histogram produced by the FCM. It is read from file as an array of 256 (typical resolution) real numbers, representing the numbers of events registered in the subsequent channels. The total number of cells involved is found by adding the 256 numbers.
- 2) *INITIAL PARAMETER VALUES.* The iterative optimization procedure

requires a set of initial parameter values to start from. Although in principle any such set will do, in practice the best results are achieved when choosing parameter values that are as close as possible to the true values.

The position of the peak values in the raw histogram are looked for to estimate  $m_1$ . A consideration is that the second peak should be at about twice the channel number of the first one,  $m_2 = 2 \cdot m_1$ . Occasionally, it may be difficult to find distinct peaks. Several trial runs then must identify the best fitting theoretical histogram.

An initial estimate of the number of  $G_1$  cells may be obtained by adding the cells in the channels to the left of  $m_1$ , and multiplying the result by two. Likewise, the number of  $G_2M$  cells may be estimated by adding and doubling all cells to the right of the channel  $m_2$ . Another way of estimating  $Z_T[m_1]$  and  $Z_T[m_2]$  is by taking 2.5 times the contents of the respective peak channels, the ordinate at the center of a standard normal distribution being approximately 0.4 (Fig. 5.7).

The dispersion parameters, PC and CF, may be estimated as follows. In the standard normal distribution the ordinate at a distance of twice the standard deviation from the mean equals 13.5% of the ordinate at the mean. So, by finding the channel to the left of  $m_1$ , whose contents equals (approximately) 0.135 times the number of cells observed in channel  $m_1$ , the  $2 \cdot sd_{m1}$  value is obtained. In a similar way, looking at the right-hand side of  $m_2$ , the  $2 \cdot sd_{m2}$  value is found (Fig. 5.7). Knowing  $m_1$ ,  $m_2$ ,  $sd_{m1}$  and  $sd_{m2}$ , by substituting them for  $j$  and  $sd$  in Eq.(5.16), two algebraic equations result from which the two unknown values, PC and CF can easily be solved.

The coefficients of the Fourier terms may initially be set to zero. The number of Fourier terms,  $N$ , should be chosen between 1 and 5.

3) *OPTIONAL PREPARATORY PROCEDURES*. The raw histogram may be subjected to one or more of the following manipulations before the actual ML optimization.

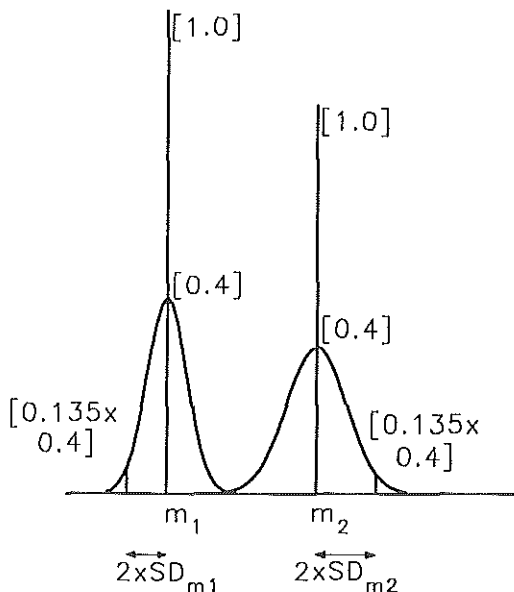
- *RESET*. The number of cells in any channel may be changed interactively. This option may be used, for instance, to remove debris from the histogram. Events registered in channels outside the  $3 \cdot sd_{m1}$  range on the left, and the  $3 \cdot sd_{m2}$  range to the right, most probably do not belong to the histogram and can be ignored by resetting the contents of the channels concerned to zero.

- *BACKGROUND CLEAN-UP*. Raw histograms are polluted by background noise, i.e., some fluorescence is observed that is not associated with DNA in a regular cell. Such background fluorescence is of low intensity. The chance of it being observed is, therefore, much bigger in lower than in higher channel ranges. For the decrease with channel number, generally, a power function is assumed:  $y = a \cdot x^b$  ( $y$  = fluorescence,  $x$  = channel). This enables a correction of the raw histogram as follows.

The power function changes into a straight line if a log-log conversion is



Fig. 5.7 A rough estimate of  $sd_{m_1}$  and  $sd_{m_2}$  can be found from the raw histogram. By considering the observed peak ( $G_1$  or  $G_2M$ ) channel contents as the result of a Gaussian redistribution, it follows from the standardized normal distribution curve that the channel contents is about 0.4 of the contents without redistribution. Also, the amplitude at a distance of 2 standard deviations from the mean ( $m_1$  or  $m_2$ , respectively) will be about 0.135x the observed peak channel contents

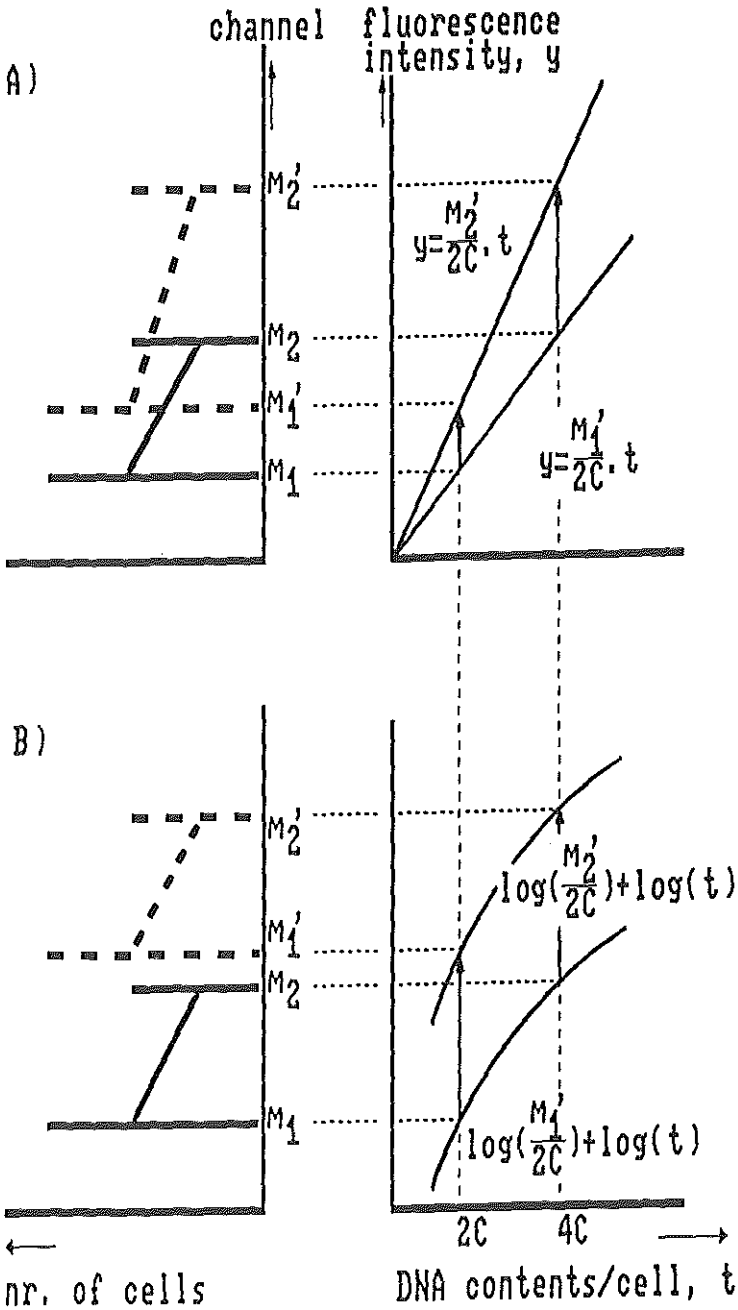


applied:  $\log(y) = \log(a \cdot x^b) = \log(a) + b \cdot \log(x)$ , or  $Y = A + b \cdot X$ . Therefore, by fitting a straight line (linear least squares method) to the datapoints that form the base line in the log-log transformed histogram, the contribution of the background to the fluorescence signal is identified. This straight line is subtracted from the log-log histogram and an inverse transformation is performed, after which a "clean" histogram remains.

- **SMOOTHING.** High frequency fluctuations in the contents of neighbor channels may be considered as artifacts, i.e., a phenomenon due to the instrumentation that has nothing to do with the studied cell population. It is possible to get rid of such high frequency disturbances by applying a filter (smoothing). A standard routine from the IMSL-package was used. It performs a binomial smoothing with endpoint correction, resulting in a less "ragged" histogram.

- **LOG-LIN SCALE CONVERSION.** The FCM can be switched between linear and logarithmic amplification. The multiharmonic analysis method is suited only for a linear scale of cellular DNA contents (fluorescence). So, if a histogram was measured in log-amplification mode, a log-lin scale conversion must

Fig. 5.8 Influence of shifting a histogram (increase in amplification) in case of a) a linear and b) a logarithmic scale of fluorescence intensity (channel number). Log amplification does not alter the shape of the histogram, lin amplification does



be applied first.

The Rijswijk FCM is adjusted such that the fluorescence increases logarithmically from A in channel 1 to  $10 \cdot A$  in channel 200 (in general: fluorescence  $k \cdot A$  is found in  $\log$  channel  $199 \cdot \log(k) + 1$ , i.e., a doubling of intensity ( $k = 2$ ) in 61 channels). This scaling is converted, first, such that the fluorescence intensity increases linearly from  $B = A/20$  in channel 1 to  $200 \cdot B = 10 \cdot A$  in channel 200. Having the same fluorescence intensity in  $\log$  channel  $i$  and  $\log$  channel  $j$ , the relation between  $i$  and  $j$  is given by

$$j = \{ 199 \cdot 10^{(i-1)/199} - 190 \} / 9. \quad (5.44)$$

Now, on the  $\log$  scale fluorescence A is found in channel 20 and  $10 \cdot A = 200 \cdot B$  is found in channel 200. Next, a new conversion scales A to be found in  $\log$  channel 1 and  $10 \cdot A$  still in  $\log$  channel 200. The new  $\log$  channel,  $j'$ , relates to the old  $\log$  channel,  $j$ , as follows:

$$j' = \{ 199 \cdot j - 3800 \} / 180. \quad (5.45)$$

Thus, fluorescence intensity A is found in  $\log$  channel  $j = 20$ ,  $j' = 1$ ; and  $10 \cdot A$  is found in  $\log$  channel  $j = j' = 200$ .

- **RIGHT/LEFT SHIFT.** The histogram can be shifted several channels to the right or to the left. This option is particularly useful for  $\log$ -amplified histograms, because then a horizontal shift does not result in a change of shape of the histogram (Fig 5.8). Thus, the peaks of sequential histograms may be shifted to the same channel number simplify direct visual inspection.

## 5.3 APPLICATION OF THE MULTIHARMONIC ANALYSIS TECHNIQUE

### 5.3.1 Test of the Multiharmonic Analysis Method

**5.3.1.1 Test Histograms.** To evaluate the performance of the present DNA histogram analysis method the computer program has been run with different measured histograms, for which results of other applied methods of analysis were known.

The first histogram, FRAMA, pertains to a population of HELA S3 cells in exponential growth. The second histogram represents a fictitious population of the same type of cells. For this synthesized histogram, BGWLL2, the exact percentages of  $G_1$ , S and  $G_2M$  cells, thus, the true histogram, are known. The true histogram was spreaded (Gaussian with constant CV), and statistical noise was added. A population of mouse lymphoma L5178Y cells has yielded a third measured histogram, SMH. The measured histograms are shown

in Figs. 5.9a, 5.10a and 5.11a, respectively.

A series of simulated histograms (I through V) analyzed with different methods was made available by Dr. H. Baisch of the University of Hamburg. Published analysis results enable direct comparison with results from the present method. The five simulated histograms differ mainly in CV value (CV, or coefficient of variation, is defined as the standard deviation for dispersion over mean channel number; in the present case it is constant in each histogram, consecutively: 2.5, 4.6, 6.0, 9.7 and 14.6%). The population size was always about 13,000 cells and the fractions of cells in the  $G_1$ -, S- and  $G_2M$ -phases of the cell cycle were approximately 45, 35 and 15%, respectively. The simulated histograms were constructed using the following assumptions:

- 1) the population was growing asynchronously, without cell loss, resulting in an exponentially decreasing maturity distribution;
- 2) the DNA content per cell increased from constant  $2C$  for  $G_1$  cells to constant  $4C$  for  $G_2M$  cells. The rate of DNA synthesis was low in early and late S-phase and high in mid S-phase;
- 3) due to measurement and staining variability, Gaussian dispersion with constant CV throughout a histogram was assumed;
- 4) statistical noise, generated with a Monte Carlo method, was added to the histograms constructed for the three above assumptions.

For further details, see [Baisch et al., 1982]. The histograms are shown as "observed histogram  $Z_M$ " in Figs. 5.12 through 5.16.

To explore the ability of the multiharmonic method to analyze unusual DNA histograms, i.e., from cell populations that are not exponentially growing, the method was applied to simulated histograms—again provided by dr. Baisch—containing disproportional amounts of phase fractions (Histograms BA35, BA27 and BA21; Table 5-5).

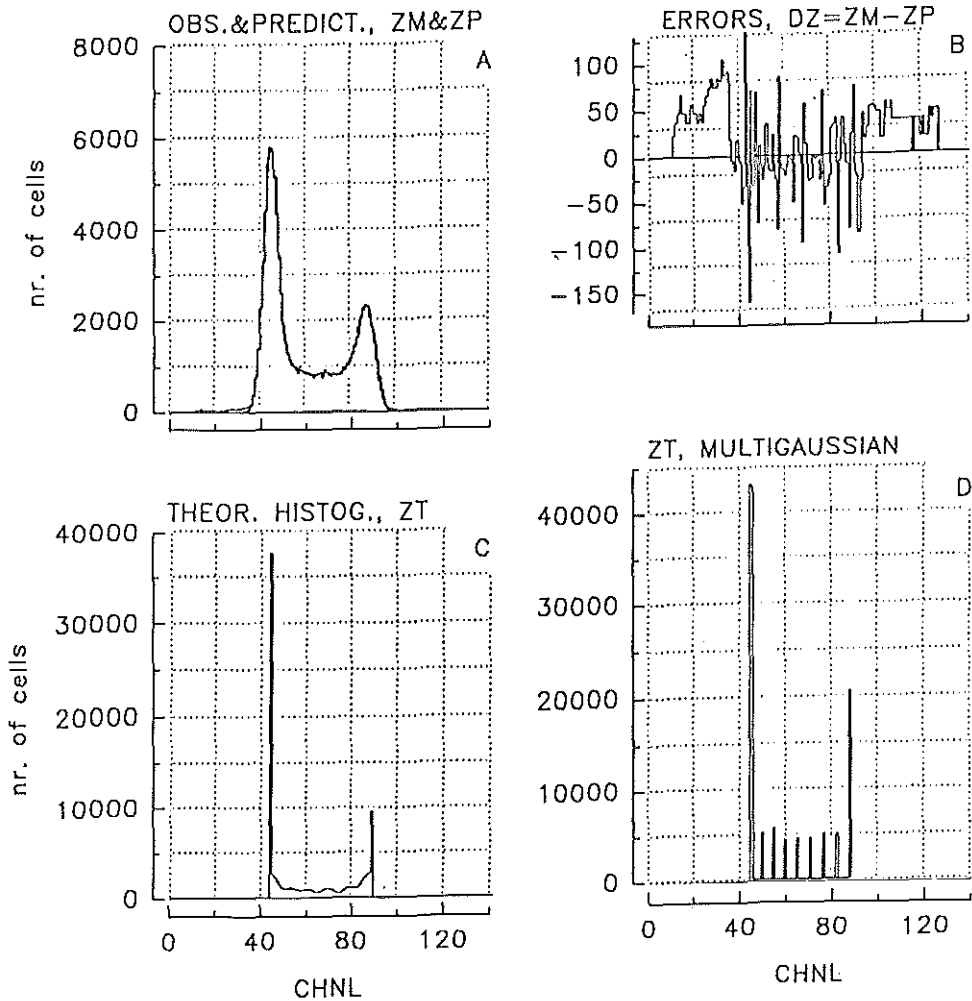
**5.3.1.2 Results for FRAMA, BGWLL2 and SMH Histograms.** By looking at the measured histograms initial values for the various model parameters were roughly estimated. The Fourier coefficients were initially set to zero. No preprocessing of raw data was performed.

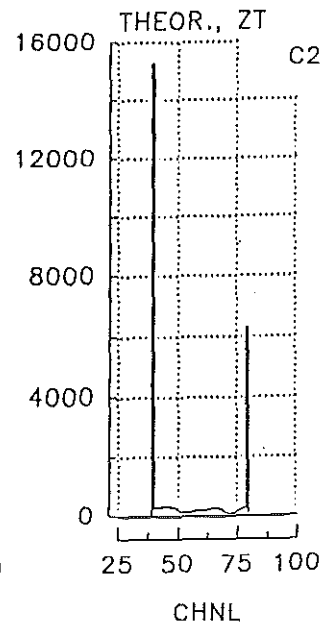
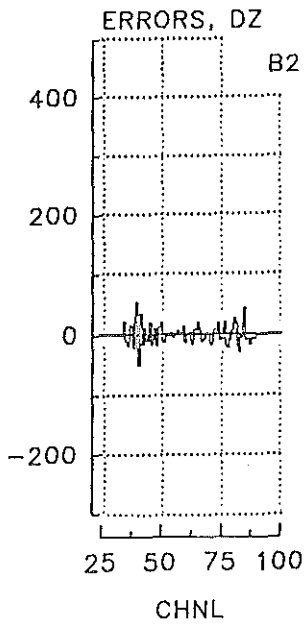
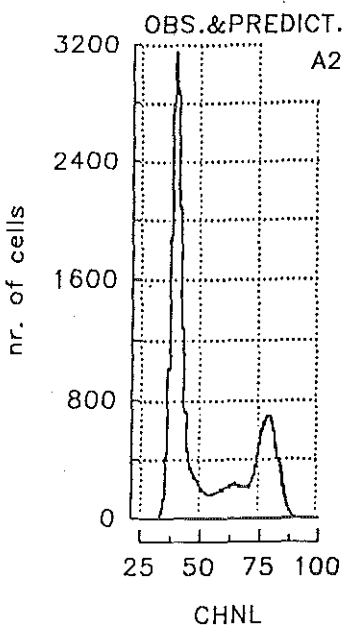
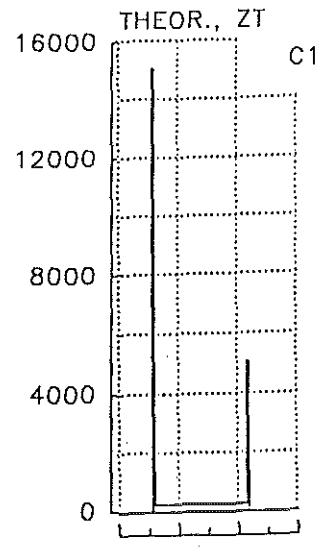
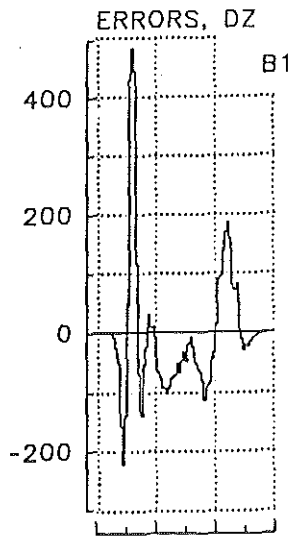
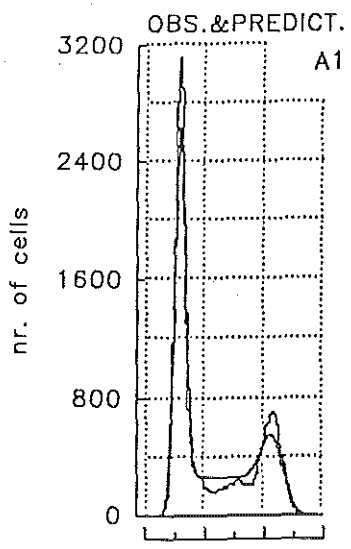
In all three cases the multiharmonic analysis program converged to yield optimum parameter values, with which a considerably improved model response could be obtained. The results of the analyses are listed in Table 5-6 and shown in Figs. 5.9, 5.10 and 5.11.

For example, take the FRAMA histogram. The goodness of fit of the predicted histogram (model output) to the measurements increased from 0.98904 to 0.99967 in 7 iteration steps, while the standard deviation of the residuals, i.e., the differences between measurements and prediction, went down from 264.2 to 38.2.

According to the model matching principle in system identification

Fig. 5.9 FRAMA histogram, HELA S3 cells; a) measured histogram (system output) and predicted histogram (optimum model output); b) deviations between measured and predicted histograms (measurement errors); c) corresponding theoretical histogram, assumed to be identical to the true histogram; d) true histogram according to multiGaussian analysis





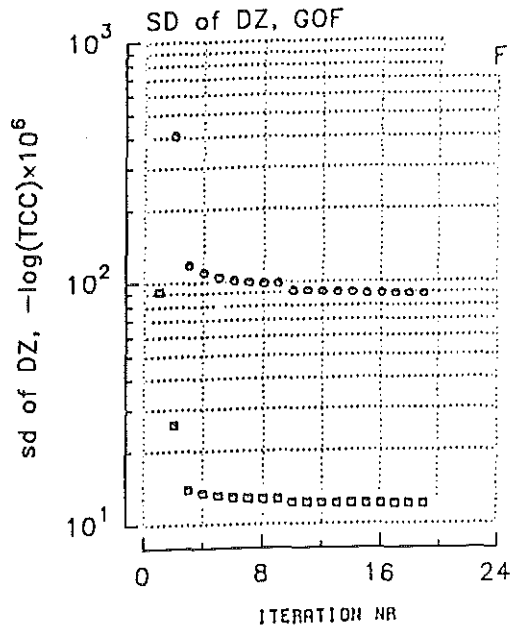
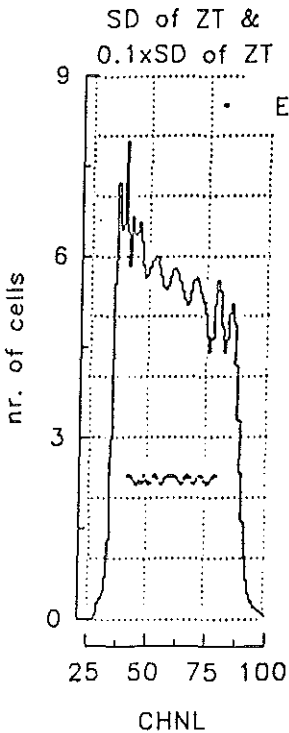
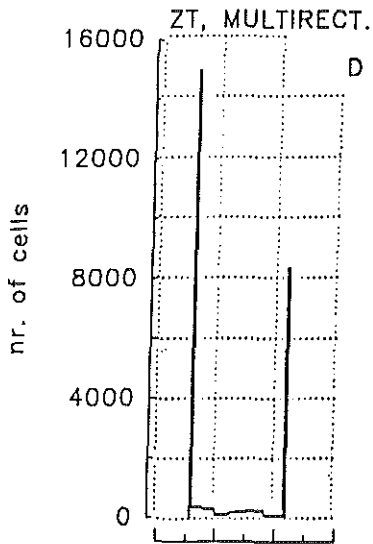


Fig. 5.10 BGWLL2 histogram simulated HELA S3 cells; a) measured histogram (system output) and predicted histogram ( $m_1 = 40$ ,  $N = 4$ ; 1: for initial parameter values; 2: optimum model output); b) deviations between measured and predicted histograms (measurement errors); c) corresponding theoretical histogram, at the optimum assumed to be identical to the true histogram; d) true histogram according to multirectangle analysis; e) standard deviations in the predicted histogram and in the theoretical histogram; f) courses of the standard deviation of the residual processing errors and of the goodness of fit criterion with iteration number

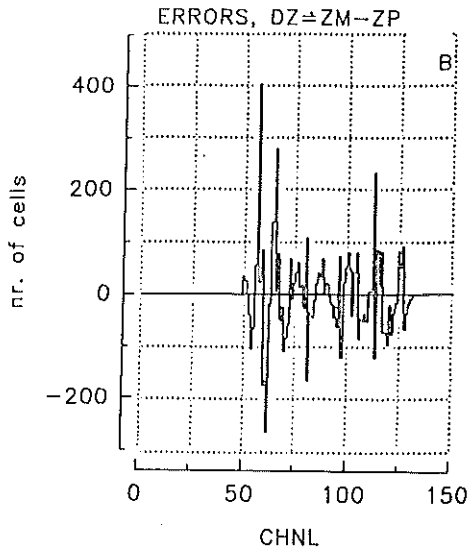
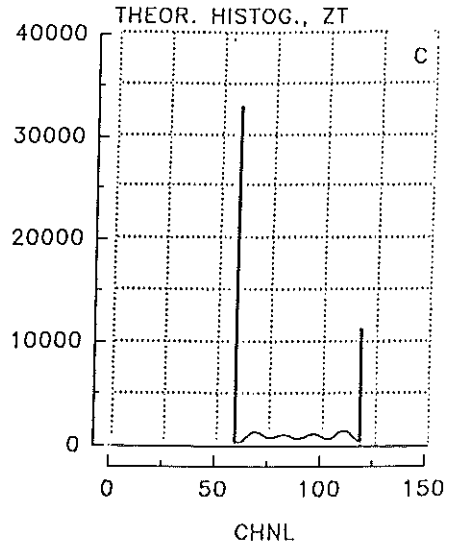
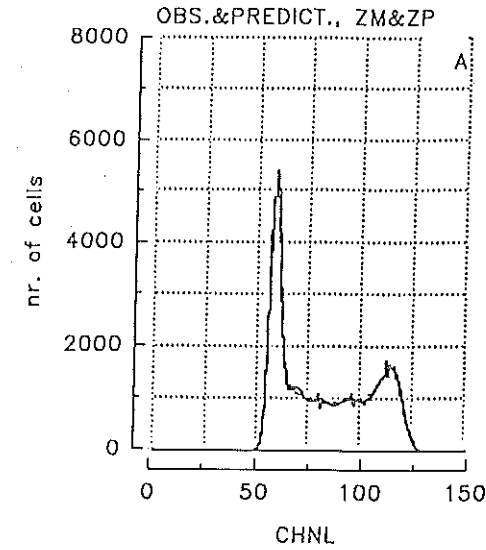


Fig. 5.11 SMH histogram, L5178Y cells; a) measured histogram (system output) and predicted histogram (optimum model output); b) deviations between measured and predicted histograms (measurement errors); c) corresponding theoretical histogram, assumed to be identical to the true histogram



TABLE 5-5 SIMULATED HISTOGRAMS WITH DISPROPORTIONAL PHASE FRACTIONS

| name | fraction:      |            |                  | dispersion:      |                  |
|------|----------------|------------|------------------|------------------|------------------|
|      | G <sub>1</sub> | S          | G <sub>2</sub> M | SD <sub>m1</sub> | SD <sub>m2</sub> |
| BA35 | high 79.1%     | low 22.0%  | low 19.9%        | 3.00             | 6.00             |
| BA27 | low 13.5%      | high 76.3% | medium 30.2%     | 3.04             | 6.08             |
| BA21 | low 7.4%       | low 1.7%   | high 49.9%       | 2.40             | 4.80             |

procedures, no satisfactory results being obtained—even if the optimum parameter values have been determined—indicates that the model should be adapted. The truth of this principle can be seen from the fact that an earlier version of the program could not achieve such good results. In that previous version the dispersion effects in the FCM were modeled with a Gaussian function of which the standard deviation was assumed to increase at a constant rate with increasing channel number,  $sd = PC \cdot (j - 1/2)$ , instead of  $sd = PC \cdot (j - 1/2) + CF$ , as in the present version. The earlier model could not achieve a better goodness of fit than 0.99633 nor bring the standard deviation of the residuals below 123.8.

In all three cases the present program version produced estimates of the population's composition, %G<sub>1</sub>, %S and %G<sub>2</sub>M, with an accuracy of  $\pm 0.5\%$  or better. The accuracy of the parameters that determine the dispersion ranges from 0.1% to 2.5%. The accuracy of the estimated Fourier coefficients, that determine the shape of the theoretical S-phase distribution, shows large fluctuations, suggesting that there is great uncertainty about the actual shape, which, to be diminished, requires information beyond the scope of a single measured histogram. This is also illustrated, for the BGWLL2 histogram, in Fig. 5.10e: the accuracy of the final predicted histogram is much better than that of the corresponding theoretical histogram, less than 8 cells against about 25 cells per S-phase channel.

Table 5-7 shows model matching results. Shifting  $m_1$  one channel right or left from true position dramatically decreases the maximum goodness of fit and increases the minimum attainable  $SD_{rpe}$ . Increasing the number of Fourier coefficients yields improvement at first. Then, for  $N > 4$  improvement becomes marginal, while uncertainty in the parameter estimates increases (not shown). Best results are obtained with  $m_1 = 40$  and  $N = 4$ . For this case, Fig. 5.10f shows that the GN algorithm first converged fast, then slowly toward a flat optimum.

Table 5-8 reveals for the SMH histogram that pre-smoothing the raw data results in better fits (compare runs 1 and 2). It also shows that the choice of a proper start point is not a serious problem here. For different initial parameter estimates the GN algorithm has converged to nearly the same optimum (runs 2 and 3); the final estimated phase fractions do not differ significantly.

TABLE 5-6 RESULTS OF THE MULTIHARMONIC ANALYSIS TECHNIQUE; HISTOGRAMS FRAMA, BGWLL2 AND SMH

| Histogram:                         |  | FRAMA [Fried & Mandel, 1979] |                                 |        | BGWLL2 [Bagwell, 1979] |                                  |         | SMH [Sheck et al., 1980] |                                 |         |        |
|------------------------------------|--|------------------------------|---------------------------------|--------|------------------------|----------------------------------|---------|--------------------------|---------------------------------|---------|--------|
| Population:                        |  | 97658                        | HELA S3 cells in                |        | 30024                  | HELA S3 cells in                 |         | 99924                    | mouse lymphoma                  |         |        |
| nr. of channels, M=                |  | 150                          | exponential growth              |        | 120                    | exponential growth               |         | 150                      | L5178Y cells                    |         |        |
| $G_1$ fraction in channel, $m_1=$  |  | 45                           |                                 | 40     |                        |                                  | 59      |                          |                                 |         |        |
| $G_2M$ fraction in channel, $m_2=$ |  | 90                           |                                 | 80     |                        |                                  | 118     |                          |                                 |         |        |
| <u>Parameter values:</u>           |  |                              |                                 |        |                        |                                  |         |                          |                                 |         |        |
|                                    |  | initial                      | final $\pm$ sd<br>(iteration 7) |        | initial                | final $\pm$ sd<br>(iteration 18) |         | initial                  | final $\pm$ sd<br>(iteration 9) |         |        |
| nr. of $G_1$ cells, $Z_T[m_1]=$    |  | 34000.0                      | 37452.8                         | 272.6  | 15000.0                | 15230.8                          | 59.8    | 34177.0                  | 32709.6                         | 362.1   |        |
| nr. of $G_2M$ cells $Z_T[m_2]=$    |  | 13000.0                      | 9467.9                          | 288.3  | 5000.0                 | 6275.4                           | 85.0    | 12776.0                  | 11069.2                         | 459.5   |        |
| Gaussian spreading                 |  | PC=                          | 0.0600                          | 0.0541 | 0.0004                 | 0.0600                           | 0.0529  | 0.0002                   | 0.0450                          | 0.0600  | 0.0005 |
|                                    |  | CF=                          | 0.0000                          | 2.8053 | 0.0033                 | 0.0000                           | -0.1006 | 0.0024                   | 0.0000                          | -0.0842 | 0.0046 |
| Fourier coefficients               |  | A[1]=                        | 0.0                             | 284.45 | 7.90                   | 0.0                              | 15.04   | 2.56                     | 0.0                             | 6.3     | 9.52   |
|                                    |  | A[2]=                        | 0.0                             | 184.83 | 10.52                  | 0.0                              | 9.64    | 3.70                     | 0.0                             | -58.80  | 11.68  |
|                                    |  | A[3]=                        | 0.0                             | 106.96 | 14.82                  | 0.0                              | 5.53    | 6.13                     | 0.0                             | -78.50  | 15.59  |
|                                    |  | A[4]=                        | 0.0                             | 94.23  | 25.43                  | 0.0                              | 2.25    | 9.57                     | 0.0                             | -144.72 | 21.80  |
|                                    |  | A[5]=                        | 0.0                             | 105.80 | 50.04                  | 0.0                              | -       | -                        | 0.0                             | -       | -      |
|                                    |  | B[1]=                        | 0.0                             | -4.27  | 4.84                   | 0.0                              | 14.26   | 1.67                     | 0.0                             | -2.47   | 7.12   |
|                                    |  | B[2]=                        | 0.0                             | -35.02 | 7.15                   | 0.0                              | 45.17   | 2.58                     | 0.0                             | -17.30  | 9.37   |
|                                    |  | B[3]=                        | 0.0                             | -55.20 | 11.68                  | 0.0                              | 4.58    | 5.00                     | 0.0                             | -60.39  | 13.14  |
|                                    |  | B[4]=                        | 0.0                             | 19.18  | 21.90                  | 0.0                              | -15.78  | 9.07                     | 0.0                             | -51.22  | 20.56  |
|                                    |  | B[5]=                        | 0.0                             | -57.21 | 52.18                  | 0.0                              | -       | -                        | 0.0                             | -       | -      |
| <u>yield:</u>                      |  |                              |                                 |        |                        |                                  |         |                          |                                 |         |        |
| Goodness of fit                    |  |                              |                                 |        |                        |                                  |         |                          |                                 |         |        |
| $Z_p$ to $Z_M$ , TCC=              |  | 0.98404                      | 0.99967                         |        | 0.98790                | 0.99980                          |         | 0.99424                  | 0.99830                         |         |        |
| SD <sub>residuals</sub> :          |  | 264.2                        | 38.2                            |        | 92.1                   | 12.0                             |         | 125.9                    | 68.4                            |         |        |
| Fourier coefficient, A[0]=         |  | 1151.32                      | 1153.12                         | 9.02   | 257.02                 | 218.40                           |         | 913.29                   | 968.02                          |         |        |
| Composition of population:         |  |                              |                                 |        |                        |                                  |         |                          |                                 |         |        |
| fraction $G_1$ (%), $F_1=$         |  | 34.8                         | 38.4                            | 0.3    | 50.0                   | 50.8                             | 0.2     | 34.2                     | 32.7                            | 0.4     |        |
| fraction S (%), $F_S=$             |  | 61.9                         | 51.9                            | 0.4    | 33.3                   | 28.3                             | 0.3     | 53.0                     | 56.2                            | 0.6     |        |
| fraction $G_2M$ (%), $F_2=$        |  | 13.3                         | 9.7                             | 0.3    | 16.7                   | 20.9                             | 0.3     | 12.8                     | 11.1                            | 0.5     |        |
| Dispersion:                        |  |                              |                                 |        |                        |                                  |         |                          |                                 |         |        |
| SD in channel $m_1$                |  | 2.70                         | 3.08                            |        | 2.40                   | 2.01                             |         | 2.66                     | 2.51                            |         |        |
| SD in channel $m_2$                |  | 5.40                         | 3.29                            |        | 4.80                   | 4.13                             |         | 5.31                     | 4.25                            |         |        |

TABLE 5-7 BGWLL2 HISTOGRAM; MODEL MATCHING: THE INFLUENCE OF CHOSEN VALUES OF  $m_1$  ( $G_1$  CHANNEL POSITION) AND N (EXTENT OF THE FOURIER EXPANSION)

| run   | 1     | 2     | 3     | 4     | 5     | 6     |
|---|-------|-------|-------|-------|-------|-------|
| <i>with chosen:</i>   |       |       |       |       |       |       |
| $m_1$   | 39    | 40    | 41    | 40    | 40    | 40    |
| N   | 3     | 3     | 3     | 2     | 4     | 6     |
| <i>analysis results:</i>  |       |       |       |       |       |       |
| iteration nr.   | 12    | 14    | 16    | 5     | 18    | 8     |
| TCC initial   | 0.946 | 0.988 | 0.954 | 0.988 | 0.988 | 0.988 |
| TCC maximum   | 0.996 | 0.999 | 0.983 | 0.999 | 0.999 | 0.999 |
| init. $SD_{rpe}$  | 191.6 | 92.1  | 178.1 | 92.1  | 92.1  | 92.1  |
| min. $SD_{rpe}$   | 50.5  | 12.0  | 108.3 | 13.3  | 12.0  | 12.0  |
| <i>estimated phase fractions (%) and SD:</i>  |       |       |       |       |       |       |
| $G_1$   | 22.1  | 50.9  | 80.3  | 50.6  | 50.8  | 50.2  |
|   | 0.6   | 0.2   | 2.2   | 0.2   | 0.2   | 0.3   |
| S   | 60.7  | 27.7  | -1.9  | 28.8  | 28.3  | 32.6  |
|   | 1.5   | 0.3   | 3.0   | 0.3   | 0.3   | 0.5   |
| $G_2M$  | 17.2  | 21.4  | 21.6  | 20.6  | 20.9  | 17.2  |
|   | 1.4   | 0.2   | 2.1   | 0.2   | 0.4   | 0.4   |
| <i>(true percentages: <math>G_1=50.0</math>, <math>S=30.0</math>, <math>G_2M=20.0</math>)</i> |       |       |       |       |       |       |

**Performance of Multiharmonic Technique in Comparison to Other Analysis Methods.** When comparing the present multiharmonic technique with other mathematical methods of analysis, (Tables 5-9, 5-10, 5-11), the following conclusions can be drawn.

The results of the analyses of the BGWLL2 histogram show that the present method arrives at the best fit (lowest SSQ) and its estimated fraction sizes come close to their true values. The results obtained with the multirectangular technique are about equally good, with perhaps a slight overestimation of the  $G_2M$  fraction and a slight underestimation of the S fraction, but with the advantage of fast computation (3 iteration steps). Comparison of Figs. 5.11c and 5.11d shows that the surplus of  $G_2M$  in the case of the multirectangular analysis appears mainly as late S cells in the multiharmonic analysis. The multiGaussian technique yields results that clearly overestimate both the  $G_1$  and the  $G_2M$  fractions at the expense of the S fraction, which could be expected, as, due to the spikes kernel function representation, early and late S cells tend to be seen as belonging to the  $G_1$  and  $G_2M$  peaks, respectively. Finally, the polynomial technique leads to the worst fit (highest SSQ), probably due to the assumed restriction in the shape of the S-phase distribution. The S-phase fraction is overestimated, and the method predicts too few  $G_2M$  cells.

If always the different methods are assumed to retain their individual

TABLE 5-8 SMH HISTOGRAM. INFLUENCE ON THE ANALYSIS RESULTS OF THE CHOSEN INITIAL PARAMETER VALUES AND OF SMOOTHING OF RAW DATA<sup>a</sup>

| run data                                     | 1 raw | 2 smoothed | 3 smoothed |      |       |      |
|--|-------|------------|------------|------|-------|------|
| <i>with chosen initial parameter values:</i> |       |            |            |      |       |      |
| Z <sub>T</sub> [m <sub>1</sub> ]             | 34177 | 32710      | 50000      |      |       |      |
| Z <sub>T</sub> [m <sub>2</sub> ]             | 12776 | 11069      | 20000      |      |       |      |
| PC   | .045  | .028       | .036       |      |       |      |
| CF   | .000  | .902       | .000       |      |       |      |
| A[1]   | .000  | 6.3        | .000       |      |       |      |
| A[2]   | .000  | -58.8      | .000       |      |       |      |
| A[3]   | .000  | -78.5      | .000       |      |       |      |
| A[4]   | .000  | -144.7     | .000       |      |       |      |
| B[1]   | .000  | -2.5       | .000       |      |       |      |
| B[2]   | .000  | -17.3      | .000       |      |       |      |
| B[3]   | .000  | -60.4      | .000       |      |       |      |
| B[4]   | .000  | -51.2      | .000       |      |       |      |
| <i>analysis results:</i>                     |       |            |            |      |       |      |
| iteration nr.                                | 9     | 3          | 13         |      |       |      |
| TCC initial                                  | .9943 | .9989      | .7885      |      |       |      |
| TCC maximum                                  | .9983 | .9992      | .9992      |      |       |      |
| initial SD <sub>r<sub>pa</sub></sub>         | 125.9 | 55.5       | 714.7      |      |       |      |
| minimum SD <sub>r<sub>pe</sub></sub>         | 68.4  | 47.0       | 47.7       |      |       |      |
| <i>estimated phase fractions (%) and SD:</i> |       |            |            |      |       |      |
| G <sub>1</sub>                               | 32.7  | .4         | 32.9       | .3   | 33.1  | .3   |
| S  | 56.2  | .6         | 56.1       | .4   | 55.5  | .6   |
| G <sub>2</sub> M                             | 11.1  | .5         | 11.0       | .3   | 11.4  | .4   |
| <i>final dispersion parameters (and SD):</i> |       |            |            |      |       |      |
| PC   | .028  | .001       | .030       | .001 | .026  | .001 |
| CF   | .902  | .005       | .912       | .003 | 1.114 | .003 |

<sup>a</sup> binomial smoothing over 3 elements with side correction; m<sub>1</sub> = 40, N = 3

TABLE 5-9 FRAMA HISTOGRAM: COMPARISON OF TWO ANALYSIS TECHNIQUES

| Histogram: FRAMA, 97658 HELA S3 cells |                 |      |                   |                |                |                     |                     |
|---------------------------------------|-----------------|------|-------------------|----------------|----------------|---------------------|---------------------|
|                                       | %G <sub>1</sub> | %S   | %G <sub>2</sub> M | m <sub>1</sub> | m <sub>2</sub> | SD[m <sub>1</sub> ] | SD[m <sub>2</sub> ] |
| a                                     | 44.0            | 35.0 | 21.0              | 45.2           | 88.5           | 3.1                 | 3.0                 |
| b                                     | 38.3            | 52.0 | 9.7               | 45.0           | 90.0           | 3.1                 | 3.3                 |

a: Fried's multiGaussian technique [Fried and Mandel, 1979],  
M = 150 channels, 7 S-phase compartments assumed.

b: Present multiharmonic technique,  
M = 150 channels, N = 5 sin, cos terms for S-phase description

TABLE 5-10 BGWLL2 HISTOGRAM: COMPARISON OF METHODS OF ANALYSIS

Histogram: BGWLL2, 30024 simulated HELA S3 cells

|   | %G <sub>1</sub> | %S   | %G <sub>2</sub> M | m <sub>1</sub> | m <sub>2</sub> | SD[m <sub>1</sub> ] | SD[m <sub>2</sub> ] | SSQ   | #it |
|---|-----------------|------|-------------------|----------------|----------------|---------------------|---------------------|-------|-----|
| a | 53.5            | 24.0 | 22.5              | 40.1           | 79.9           | 2.01                | 3.96                | 25323 | 13  |
| b | 50.2            | 33.8 | 16.0              | 40.0           | 80.6           | 2.01                | 3.66                | 45589 | 5   |
| c | 49.5            | 27.6 | 22.9              | 40.0           | 79.8           | 1.99                | 4.11                | 18474 | 3   |
| d | 50.8            | 28.3 | 20.9              | 40.0           | 80.0           | 2.01                | 4.13                | 17278 | 18  |
| t | 50.0            | 30.0 | 20.0              | 40.0           | 80.0           | 2.00                | 4.00                | -     | -   |

- a: *Fried's multiGaussian technique [Bagwell, 1979, p 235],  
M = 120 channels, 6 S-phase compartments (spikes) assumed.*
- b: *Dean/Jett's polynomial technique [Bagwell, 1979, p 235],  
M = 120 channels, 2nd order polynomial for S-phase description.*
- c: *Bagwell's multirectangular technique [Bagwell, 1979, p 235],  
M = 120 channels, 6 S-phase compartments (rectangles) assumed.*
- d: *Present multiharmonic technique,  
M = 120 channels, N=4 sin,cos terms for S-phase description.*
- t: *True values; synthesized histogram simulates NC=30024 HELA S3  
cells; constant CV, non-constant rate of DNA-synthesis,  
+ statistical noise [Bagwell, 1979, p 235].*

SSQ = sum of squared deviations between calculated and measured histograms.  
#it = number of iterations required

properties of over/under estimating as noticed above for the BGWLL2 histogram, then, for the SMH histogram the conclusion is valid, that the present multiharmonic technique again has arrived at very probable population fraction sizes, also, if compared to the outcome of the autoradiobiography experiment. And the same thing can be said as far as the FRAMA histogram is concerned, though here the discrepancy in G<sub>2</sub>M size found with the multiGaussian technique (probably too large) and the multiharmonic technique (less than half the former value) seems rather (too?) large.

The multiGaussian analysis of the FRAMA histogram, using seven S spikes in the kernel function, has required the estimation of 20 model parameters. Only 15 parameters were needed for a continuous kernel function in the multiharmonic analysis, using five pairs of Fourier coefficients. The polynomial technique would have required as few as nine parameters, but it is doubtful whether a good fit would have been achieved with a second degree curve.

**5.3.1.3 Results for the Simulated Histograms I through V.** The results are shown in Figs. 5.12 through 5.16. Some numerical results are given in Tables 5-12.

TABLE 5-11 SMH HISTOGRAM: COMPARISON OF METHODS OF ANALYSIS

Histogram: SMH, 99924 mouse lymphoma L5178Y cells.

|   | %G <sub>1</sub> | %S    | %G <sub>2</sub> M | m <sub>1</sub> | m <sub>2</sub> | SD[m <sub>1</sub> ] | SD[m <sub>2</sub> ] |
|---|-----------------|-------|-------------------|----------------|----------------|---------------------|---------------------|
| a | 32.09           | 51.04 | 16.83             | -              | -              | -                   | -                   |
| b | 27.45           | 63.05 | 9.50              | -              | -              | -                   | -                   |
| c | 26.15           | 59.56 | 13.62             | -              | -              | -                   | -                   |
| d | 32.73           | 56.19 | 11.08             | 59             | 118            | -                   | -                   |
| e | -               | 60.4  | -                 | -              | -              | -                   | -                   |

- a: *Fried's multiGaussian technique [Sheck et al., 1980, p109], M=150 channels, 6 to 10 S-phase compartments assumed, spacing linear or non-linear, m<sub>2</sub>/m<sub>1</sub> specified or free.*  
 b: *Dean/Jett's polynomial technique [Sheck et al., 1980, p109], M=150 channels, 2nd order polynomial for S-phase description, m<sub>2</sub>/m<sub>1</sub> free.*  
 c: *Bagwell's multirectangular technique [Sheck et al., 1980, p109], M=150 channels, 6 to 9 S-phase compartments assumed, spacing non-linear, m<sub>2</sub>/m<sub>1</sub> specified or free.*  
 d: *Present multiharmonic technique [Sheck et al., 1980, p109], M=150 channels, N=4 sin,cos terms for S-phase description, m<sub>2</sub>/m<sub>1</sub> specified.*  
 e: *obtained by autoradiography [Sheck et al., 1980, p109].*

a,b,c: percentages are mean values resulting from several runs

No precleaning of the measured histograms was performed. Rough estimates for various parameter values required as initial input for the computer program could be obtained by visual inspection of each observed histogram. The Fourier coefficients were initially set to zero.

The choice of the necessary number of Fourier coefficients, N, and of the best G<sub>1</sub> peak position, m<sub>1</sub>, was determined in several runs by inspection of the residuals. This is illustrated for histogram II in Table 5-15.

This table shows that a wrong choice of G<sub>1</sub> peak channel clearly corresponds with a decreased goodness of fit and a large variation in the deviations between predicted and observed histograms. Therefore, it is not difficult to choose the correct value of m<sub>1</sub>.

Once the proper m<sub>1</sub> has been set, increasing the number of Fourier terms for S-phase description appears to improve the goodness of fit. However, beyond a certain value of N—and in practice, at N = 5 to 7, this value is rather low—further increase in goodness of fit becomes insignificant. Thus, only a few

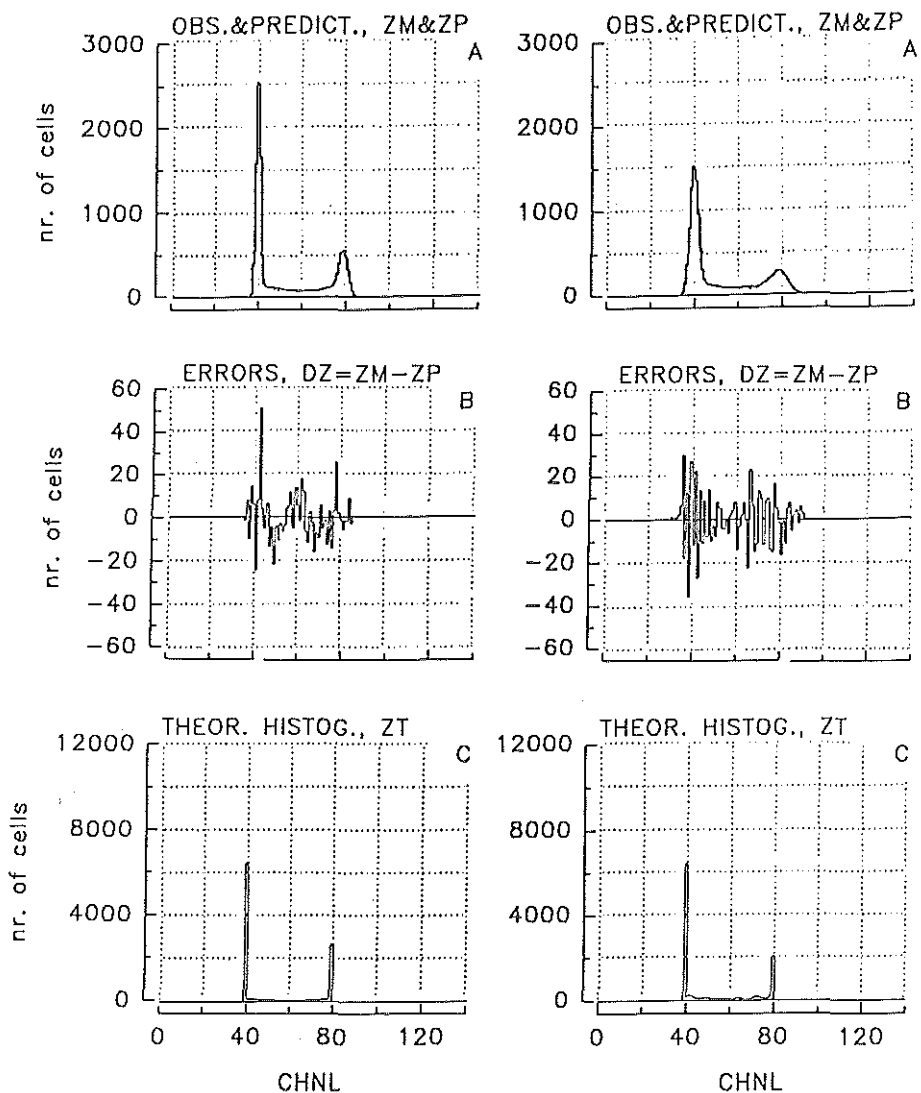
Fourier terms suffice for the description of the population's S-phase fraction.

An improper choice of  $m_1$  immediately becomes obvious. A shift of only one channel to either side of the correct channel, 40, causes a dramatic increase in the minimum standard deviation of the residuals,  $\sigma_r$ , and decreases the maximum attainable goodness of fit,  $TCC_{max}$ , of the predicted histogram to the observations. Keeping  $m_1 = 40$  and doing analysis runs with increasing  $N$ , thus increasing the number of terms in the approximation of the S-phase distribution, shows that at first  $TCC_{max}$  increases considerably, indicating the improvement of the model. Then, at a certain moment (in this case  $N = 6$ ), a further increase in  $N$  apparently no longer adds significantly to the goodness of fit.  $Sd_{rpe}$  may still slightly decrease, but the uncertainty in the estimated parameter values increases at a faster rate (not shown).

Apparently, a near constant predicted histogram  $Z_P$  can be the result of different theoretical histograms  $Z_T$ . The uncertainty lies mainly in the shape of the S-phase distribution. Estimated phase fraction sizes are fairly constant (compare runs 6 and 7 in Table 5-15). But the (relative) accuracy of the Fourier coefficients that determine the theoretical S-phase distribution shows large fluctuations, as can be seen in Table 5-14(II). In absolute value, the SD of consecutive Fourier coefficients, both A and B, also increases. Therefore, from a given moment on, taking more terms into consideration is rather useless, as it does not contribute to greater accuracy. On the contrary, parameters start to develop high correlations, i.e., their values can no longer be varied independently. Furthermore, negative cell numbers are sometimes calculated for several channels in the corresponding theoretical histogram  $Z_T$ .

For histogram II, run 6 yields the best results. Table 5-14(II) shows that starting from the initial estimates the multiharmonic technique converges to yield the optimum parameter values in 8 iteration steps, improving the fit from 0.99344 to 0.99946. The corresponding predicted histogram  $Z_P$  fits the measured data well (Fig. 5.13); deviations are small and positive or negative at random. However, these deviations are somewhat too large to confirm the complete correctness of the present model. For example, in the S channels, the calculated standard deviation  $sd_P$  amounts to about 4.5 cells, which means that the residuals ought to lie, for 99%, within the  $\pm 3 \cdot 4.5 = \pm 13.5$  cells boundaries. Clearly, some deviations exceed 15 cells or more. As could be expected, the uncertainty in the estimated theoretical histogram is greater:  $Z_T$  can be varied to some extent without much influence on the resulting  $Z_P$ . In S-phase, the standard deviation  $sd_T$  amounts to about 23 cells. In other words, there is uncertainty about the actual shape of the S-phase distribution, which, to be diminished, requires information beyond the scope of a single observed histogram.

In spite of this observation, the model can be said to be adequate as far as the determination of the fraction sizes is concerned. For the simulated histo-

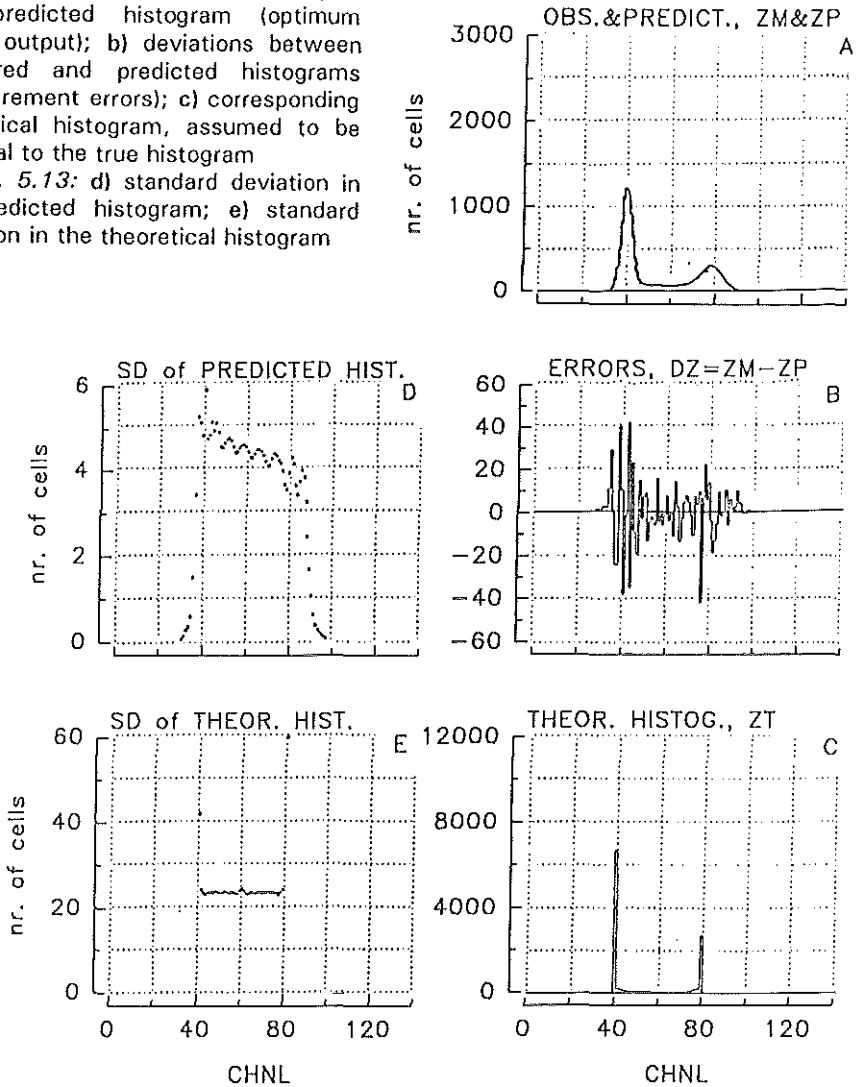


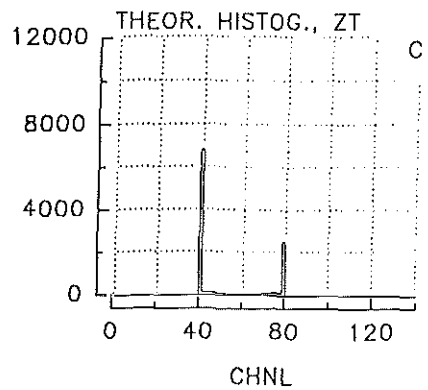
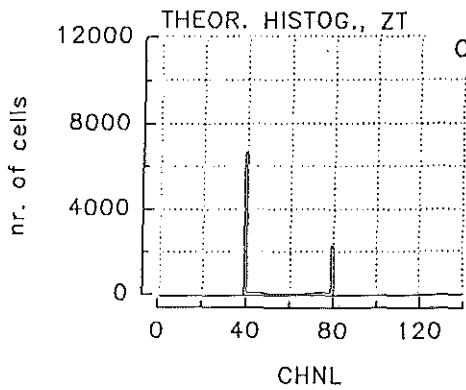
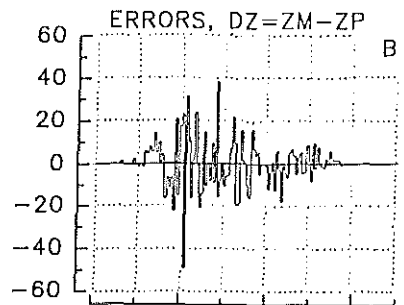
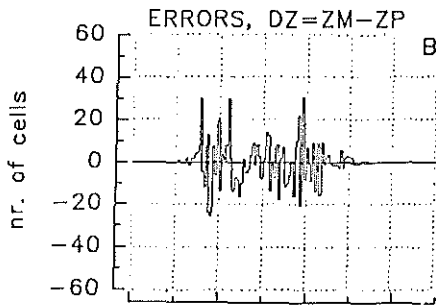
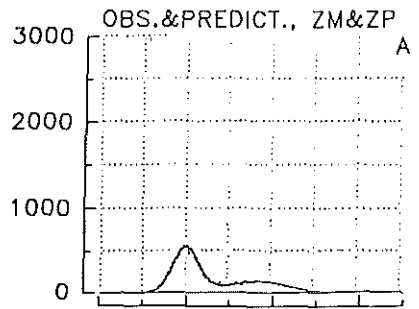
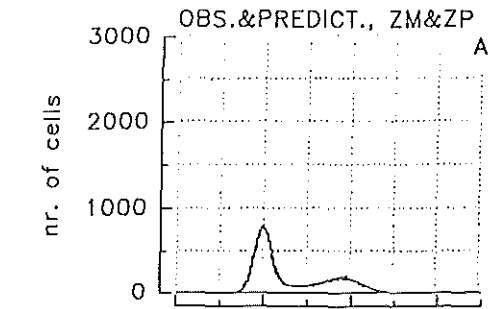
i

ii



Fig. 5.12 through 5.16: Simulated histograms I through V (Baisch); a) measured histogram (system output) and predicted histogram (optimum model output); b) deviations between measured and predicted histograms (measurement errors); c) corresponding theoretical histogram, assumed to be identical to the true histogram for Fig. 5.13; d) standard deviation in the predicted histogram; e) standard deviation in the theoretical histogram





iv

v

TABLE 5-14 RESULTS OF THE MULTIHARMONICAL ANALYSIS TECHNIQUE; HISTOGRAMS I THROUGH V (Baisch)

| Histogram:  | I (Baisch 25)         |               |        | II (Baisch 32)        |               |        | III (Baisch 18)       |               |        |
|---|-----------------------|---------------|--------|-----------------------|---------------|--------|-----------------------|---------------|--------|
| Population:   | 13000 simulated cells |               |        | 13000 simulated cells |               |        | 13015 simulated cells |               |        |
| nr. of channels, M=   | 128                   |               |        | 128                   |               |        | 128                   |               |        |
| G <sub>1</sub> fraction in channel, m <sub>1</sub> =            | 40                    |               |        | 40                    |               |        | 40                    |               |        |
| G <sub>2</sub> M fraction in channel, m <sub>2</sub> =          | 80                    |               |        | 80                    |               |        | 80                    |               |        |
| <u>Parameter values:</u>  | initial               | final ±       | sd     | initial               | final ±       | sd     | initial               | final ±       | sd     |
|   |                       | (iteration 7) |        |                       | (iteration 8) |        |                       | (iteration 9) |        |
| nr. of G <sub>1</sub> cells, Z <sub>T</sub> [m <sub>1</sub> ]=  | 6500.0                | 6388.4        | 17.4   | 6500.0                | 6362.2        | 41.6   | 7200.0                | 6652.7        | 53.5   |
| nr. of G <sub>2</sub> M cells Z <sub>T</sub> [m <sub>2</sub> ]= | 2600.0                | 2614.1        | 26.2   | 2300.0                | 2016.4        | 59.4   | 2900.0                | 2713.4        | 74.8   |
| Gaussian PC=  | 0.0175                | 0.0258        | 0.0005 | 0.0525                | 0.0449        | 0.0024 | 0.0525                | 0.0592        | 0.0005 |
| spreading CF=   | 0.0000                | -0.0275       | 0.0220 | 0.0000                | 0.0361        | 0.0961 | 0.0000                | 0.0110        | 0.0060 |
| Fourier A[1]=   | 0.0                   | 14.49         | 0.89   | 0.0                   | 20.82         | 1.80   | 0.0                   | 23.56         | 2.40   |
| coefficients A[2]=  | 0.0                   | 5.29          | 0.97   | 0.0                   | 7.41          | 2.46   | 0.0                   | 13.74         | 3.81   |
| A[3]=   | 0.0                   | 0.16          | 1.13   | 0.0                   | 0.28          | 3.60   | 0.0                   | 16.47         | 5.67   |
| A[4]=   | 0.0                   | -0.46         | 1.34   | 0.0                   | -4.96         | 6.01   | 0.0                   | -             | -      |
| A[5]=   | 0.0                   | -             | -      | 0.0                   | 10.72         | 9.58   | 0.0                   | -             | -      |
| B[1]=   | 0.0                   | 2.82          | 0.75   | 0.0                   | 4.05          | 1.13   | 0.0                   | 1.46          | 2.40   |
| B[2]=   | 0.0                   | 1.56          | 0.85   | 0.0                   | 4.99          | 1.58   | 0.0                   | 1.30          | 3.81   |
| B[3]=   | 0.0                   | 1.72          | 1.02   | 0.0                   | 2.57          | 2.79   | 0.0                   | 0.09          | 5.76   |
| B[4]=   | 0.0                   | 2.46          | 1.26   | 0.0                   | 15.63         | 5.67   | 0.0                   | -             | -      |
| B[5]=   | 0.0                   | -             | -      | 0.0                   | 21.58         | 9.31   | 0.0                   | -             | -      |
| <u>yield:</u>   |                       |               |        |                       |               |        |                       |               |        |
| Goodness of fit   |                       |               |        |                       |               |        |                       |               |        |
| Z <sub>p</sub> to Z <sub>M</sub> , TCC=                         | 0.95183               | 0.99981       |        | 0.99344               | 0.99946       |        | 0.98983               | 0.99902       |        |
| SD <sub>residuals</sub> :                                       | 99.3                  | 6.4           |        | 29.4                  | 8.4           |        | 33.9                  | 10.5          |        |
| Fourier coefficient, A[0]=                                      | 100.00                | 102.50        | 0.81   | 107.69                | 118.50        |        | 74.74                 | 93.56         |        |
| Composition of population:                                      |                       |               |        |                       |               |        |                       |               |        |
| fraction G <sub>1</sub> (%), F <sub>1</sub> =                   | 50.0                  | 49.1          | 0.1    | 50.0                  | 48.9          | 0.3    | 55.3                  | 51.1          | 0.4    |
| fraction S (%), F <sub>S</sub> =                                | 30.0                  | 30.8          | 0.2    | 32.3                  | 35.6          | 0.6    | 22.4                  | 28.0          | 0.7    |
| fraction G <sub>2</sub> M (%), F <sub>2</sub> =                 | 20.0                  | 20.1          | 0.2    | 17.7                  | 15.5          | 0.5    | 22.3                  | 20.9          | 0.6    |
| Dispersion:   |                       |               |        |                       |               |        |                       |               |        |
| SD in channel m <sub>1</sub>                                    | 0.70                  | 1.01          |        | 2.10                  | 1.84          |        | 2.10                  | 2.38          |        |
| SD in channel m <sub>2</sub>                                    | 1.40                  | 2.04          |        | 4.20                  | 3.63          |        | 4.20                  | 4.75          |        |

TABLE 5-14 CONTINUED

| Histogram:  |               | IV (Baisch 13)        |         |        | V (Baisch 03)         |         |        |
|---|---------------|-----------------------|---------|--------|-----------------------|---------|--------|
| Population:   |               | 12999 simulated cells |         |        | 12998 simulated cells |         |        |
| nr. of channels, M=   |               | 128                   |         |        | 128                   |         |        |
| G <sub>1</sub> fraction in channel, m <sub>1</sub> =            |               | 40                    |         |        | 40                    |         |        |
| G <sub>2</sub> M fraction in channel, m <sub>2</sub> =          |               | 80                    |         |        | 80                    |         |        |
| <u>Parameter values:</u>  |               | initial               | final ± | sd     | initial               | final ± | sd     |
|   |               | (iteration 13)        |         |        | (iteration 15)        |         |        |
| nr. of G <sub>1</sub> cells, Z <sub>T</sub> [m <sub>1</sub> ]=  |               | 6500.0                | 6642.3  | 47.0   | 8200.0                | 6814.7  | 78.3   |
| nr. of G <sub>2</sub> M cells Z <sub>T</sub> [m <sub>2</sub> ]= |               | 2200.0                | 2259.3  | 63.5   | 2300.0                | 2482.2  | 104.8  |
| Gaussian  | PC=           | 0.0700                | 0.1011  | 0.0009 | 0.1200                | 0.1426  | 0.0021 |
| spreading   | CF=           | 0.0000                | -0.1956 | 0.0162 | 0.0000                | 0.0902  | 0.0444 |
| Fourier   | A[1]=         | 0.0                   | 21.55   | 2.43   | 0.0                   | 22.33   | 4.90   |
| coefficients  | A[2]=         | 0.0                   | -       | -      | 0.0                   | -       | -      |
|   | A[3]=         | 0.0                   | -       | -      | 0.0                   | -       | -      |
|   | A[4]=         | 0.0                   | -       | -      | 0.0                   | -       | -      |
|   | A[5]=         | 0.0                   | -       | -      | 0.0                   | -       | -      |
|   | B[1]=         | 0.0                   | -0.03   | 1.76   | 0.0                   | 0.32    | 4.07   |
|   | B[2]=         | 0.0                   | -       | -      | 0.0                   | -       | -      |
|   | B[3]=         | 0.0                   | -       | -      | 0.0                   | -       | -      |
|   | B[4]=         | 0.0                   | -       | -      | 0.0                   | -       | -      |
|   | B[5]=         | 0.0                   | -       | -      | 0.0                   | -       | -      |
|   | <u>yield:</u> |                       |         |        |                       |         |        |
| Goodness of fit   |               |                       |         |        |                       |         |        |
| Z <sub>P</sub> to Z <sub>M</sub> ,                              | TCC=          | 0.97306               | 0.99904 |        | 0.96667               | 0.99812 |        |
| SD <sub>residuals</sub> :                                       |               | 45.1                  | 8.5     |        | 43.5                  | 10.4    |        |
| Fourier   |               |                       |         |        |                       |         |        |
| coefficient, A[0]=  |               | 110.23                | 105.06  | 2.03   | 64.05                 | 94.90   |        |
| Composition of population:                                      |               |                       |         |        |                       |         |        |
| fraction G <sub>1</sub> (%), F <sub>1</sub> =                   |               | 50.0                  | 51.1    | 0.3    | 63.1                  | 52.4    | 0.6    |
| fraction S (%), F <sub>S</sub> =                                |               | 33.1                  | 31.5    | 0.6    | 19.2                  | 28.5    | 1.0    |
| fraction G <sub>2</sub> M (%), F <sub>2</sub> =                 |               | 16.9                  | 17.4    | 0.5    | 17.7                  | 19.1    | 0.8    |
| Dispersion:   |               |                       |         |        |                       |         |        |
| SD in channel m <sub>1</sub>                                    |               | 2.80                  | 3.85    |        | 4.80                  | 5.79    |        |
| SD in channel m <sub>2</sub>                                    |               | 5.60                  | 7.89    |        | 9.60                  | 11.49   |        |

TABLE 5-15 SIMULATED HISTOGRAM II; MODEL MATCHING: INFLUENCE OF THE CHOICE OF  $G_1$  CHANNEL POSITION,  $m_1$ , AND NUMBER OF TERMS IN THE S-PHASE DISTRIBUTION'S FOURIER EXPANSION, N

Histogram: II, Baisch 32, 13000 cells

| run   | 1      | 2      | 3      | 4      | 5      | 6      | 7 <sup>a</sup> |
|---|--------|--------|--------|--------|--------|--------|----------------|
| <i>with chosen:</i>   |        |        |        |        |        |        |                |
| $m_1$   | 39     | 40     | 41     | 40     | 40     | 40     | 40             |
| N   | 3      | 3      | 3      | 1      | 4      | 5      | 6              |
| <i>analysis results:</i>  |        |        |        |        |        |        |                |
| nr. of iterations   | 12     | 7      | 6      | 3      | 8      | 8      | 9              |
| TCC <sub>max</sub>  | .99449 | .99944 | .97904 | .99917 | .99944 | .99946 | .99946         |
| SD <sub>min</sub> (residuals)   | 26.98  | 8.60   | 52.42  | 10.47  | 8.58   | 8.42   | 8.37           |
| <i>estimated phase fractions (%) and SD:</i>  |        |        |        |        |        |        |                |
| $G_1$   | 20.8   | 49.1   | 74.1   | 50.7   | 50.0   | 48.9   | 48.7           |
|   | .7     | .3     | 2.2    | .3     | .3     |        | .3.4           |
| S   | 78.5   | 37.2   | 10.9   | 34.7   | 36.7   | 35.6   | 35.6           |
|   | 1.8    | .5     | 3.0    | .5     | .5     |        | .6             |
| $G_2M$  | 0.7    | 13.7   | 15.0   | 14.6   | 14.3   | 15.5   | 15.7           |
|   | 1.6    | .4     | 2.0    | .4     | .4     | .5     | .5             |
| <i>(true percentages: <math>G_1=45.7</math>, <math>S=42.6</math>, <math>G_2M=11.7</math>)</i> |        |        |        |        |        |        |                |

<sup>a</sup>)negative cell numbers occur in several channels of the corresponding theoretical histogram

grams I through V, the magnitudes of the various phase fractions are exactly known. This makes it possible to test to what extent the method is able to produce accurate estimates of the composition of the population. As can be seen in Table 5-16, the multiharmonic technique yields very acceptable phase fraction sizes. A slight overestimation of the  $G_1$  and  $G_2M$  fractions occurred; for histogram II, there is a 3.5% difference between true and estimated values. Too few S cells were predicted (7% discrepancy). Due to the limited accuracy of the S-phase distribution approximation, calculated S-phase subfraction sizes will obviously be less reliable.

Similar results were obtained with the multiharmonic analyses of the other simulated histograms (Figs 5.12 and 5.14 through 5.16). It was observed that the larger the CV, the fewer Fourier terms are necessary to approximate the true S distribution shape (N = 5 for CV = 5%, N = 1 for CV = 15%).

**Performance of Multiharmonic Technique in Comparison to Other Analysis Methods.** The series of simulated histograms has been previously analyzed with various mathematical methods. The results for the multiGaussian and polynomial

TABLE 5-16 SIMULATED HISTOGRAMS I THROUGH V: COMPARISON OF RESULTS OF THREE ANALYSIS TECHNIQUES

| histogram | method | %G <sub>1</sub> | %S   | %G <sub>2</sub> M | CV <sub>m1</sub> (%) | CV <sub>m2</sub> (%) |
|-----------|--------|-----------------|------|-------------------|----------------------|----------------------|
| I         | a      | 51              | 28   | 21                | 2.8                  | 2.8                  |
|           | b      | 49              | 32   | 19                | -                    | -                    |
|           | c      | 50              | 30   | 20                | 2.5                  | 2.8                  |
|           | t      | 45.4            | 37.1 | 17.4              | 2.5                  | 2.5                  |
| II        | a      | 53              | 31   | 16                | 4.7                  | 4.7                  |
|           | b      | 53              | 30   | 17                | -                    | -                    |
|           | c      | 49              | 36   | 15                | 4.6                  | 4.5                  |
|           | t      | 45.7            | 42.6 | 11.7              | 4.6                  | 4.6                  |
| III       | a      | 54              | 22   | 24                | 6.0                  | 6.0                  |
|           | b      | 55              | 21   | 24                | -                    | -                    |
|           | c      | 51              | 28   | 21                | 5.9                  | 5.9                  |
|           | t      | 49.6            | 30.5 | 19.9              | 6.0                  | 6.0                  |
| IV        | a      | 56              | 25   | 19                | 9.7                  | 9.7                  |
|           | b      | 54              | 25   | 21                | -                    | -                    |
|           | c      | 51              | 32   | 17                | 9.6                  | 9.9                  |
|           | t      | 47.7            | 36.5 | 15.8              | 9.7                  | 9.7                  |
| V         | a      | 57              | 24   | 19                | 12.0                 | 12.0                 |
|           | b      |                 |      | not done          |                      |                      |
|           | c      | 52              | 29   | 19                | 14.5                 | 14.4                 |
|           | t      | 50.5            | 31.3 | 18.2              | 14.6                 | 14.6                 |

a: Fried's multiGaussian technique [Baisch et al., 1982]

b: Dean and Jett's polynomial technique [Baisch et al., 1982]

c: Present multiharmonic technique

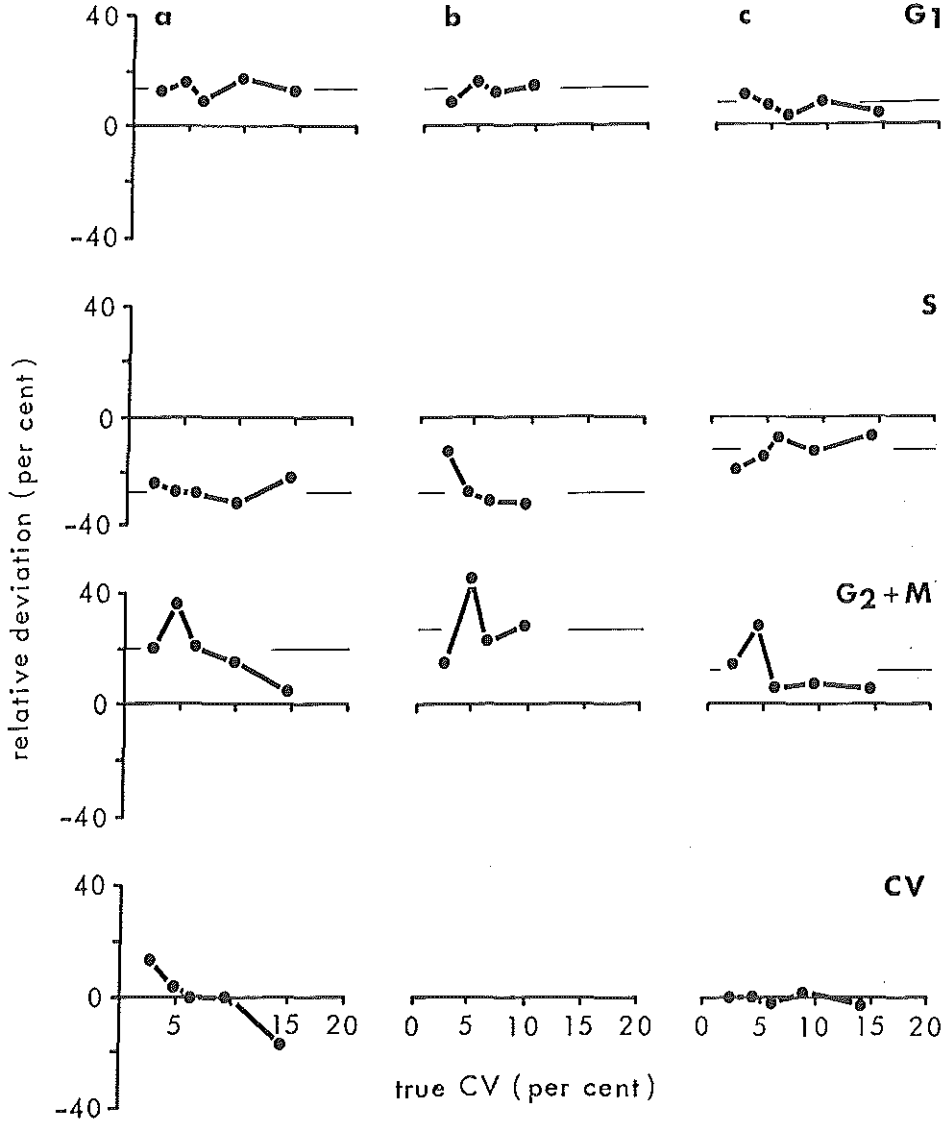
t: True values [Baisch et al., 1982]

techniques, which are the most common methods of analysis, are repeated in Table 5-16 and Fig. 5.17 so as to compare them with those of the present multiharmonic technique. Figure 5.17 reveals at once that, for each histogram, all three methods overestimate the amount of G<sub>1</sub> cells (the minimum and maximum score of the relative deviation, defined as

$$\text{relative deviation} = 100 \cdot \frac{\% \text{estimated} - \% \text{true}}{\% \text{true}}, \quad (5.46)$$

amounts to 3% and 17% respectively). This is also the case with the number of G<sub>2</sub>M cells (4-45%), at the expense of the S fraction (7-32%). At average the

Fig. 5.17 Histograms I through V: comparison of different methods of analysis. Relative deviation between estimated and true population composition (%G<sub>1</sub>, %S, %G<sub>2</sub>M) and between estimated and true coefficient of variation, as function of the true value of the latter parameter for spread. Mean values of relative deviation over the entire true CV range are indicated (thin solid lines). Rel.dev. = 100 • (estimated value - true value) / true value (%). a) multiGaussian, b) polynomial, c) multiharmonic technique



multiharmonic technique shows the best results, with the smallest deviations between actual and estimated population compositions. The average relative deviations over the entire CV range amount to  $+6.0 \pm 3.1\%$ ,  $-12.5 \pm 5.0\%$  and  $+11.9 \pm 10.0\%$  for the  $G_1$ -, S- and  $G_2M$ -phases, respectively. The polynomial technique failed in analyzing the histogram with the highest CV. With respect to the estimation of CV, the multiGaussian and multiharmonic techniques do not differ much, except at very high and low CV values where the former technique becomes unreliable. The average relative deviation amounts to  $-0.7 \pm 10.8\%$  as compared to  $-0.3 \pm 1.0\%$  for the multiharmonic technique. It should be noted that the multiharmonic technique does not consider CV to be necessarily constant throughout a histogram. As mentioned above, by modeling the Gaussian dispersion of the theoretical histogram with  $sd_j = PC \cdot (j^{-1/2}) + CF$  instead of  $sd_j = CV \cdot j$ , better fits could often be achieved (although PC and CF appear to be highly correlated: correlation coefficient,  $> 0.9$ ). To enable comparison, in Table 5-20,  $CV_{m1}$  and  $CV_{m2}$  values have always been calculated as:  $CV_j = sd_j/j$ . Thus, it can be concluded that the fact that the simulated histograms should have a constant true CV value was not always recognized by the multiharmonic technique, but diversions remain very small. The CV estimate by the polynomial technique was not evaluated.

#### **5.3.1.4 Results for the Simulated Histograms BA35, BA27 and BA21.**

With respect to the histograms analyzed above, corresponding to asynchronous exponential growth, these histograms contain disproportionately sized phase fractions. Still, in general the multiharmonic technique succeeded rather well in fitting the observed histograms and finding the proper fraction sizes. Results (shown in Table 5-17 and in Fig.s 5.18 through 5.20) are somewhat poorer for the case in which both the true  $G_1$  and the  $G_2M$  fractions are very small.

### **5.3.2 Sequential Histograms**

By analyzing sequential histograms, i.e., a series of consecutive histograms measured to monitor the development of a cell population as function of time, changes in cell kinetics can be detected and quantified in terms of time histories of the cell cycle phase fractions. For instance, in this way the influence of adding growth factors or exposure to drugs can be studied quantitatively.

**5.3.2.1 Influence of Growth Factor Presence on the Cell Kinetics of B-lymphocytes.** The following experiment was conducted by Lansdorp et al. [1986]. B1329 cells cultured with or without growth factor (1% v/v HGF) were harvested after fixed time intervals by centrifugation. The cell pellet (approximately  $10^6$  cells) was resuspended in 1 ml ice-cold phosphate buffered saline (PBS) and this suspension was added dropwise to 2 ml ethanol of  $-20^\circ\text{C}$  under



Fig. 5.18 Histogram Ba35; a) measured histogram (system output) and predicted histogram (optimum model output); b) deviations between measured and predicted histograms (measurement errors); c) corresponding theoretical histogram, assumed to be identical to the true histogram

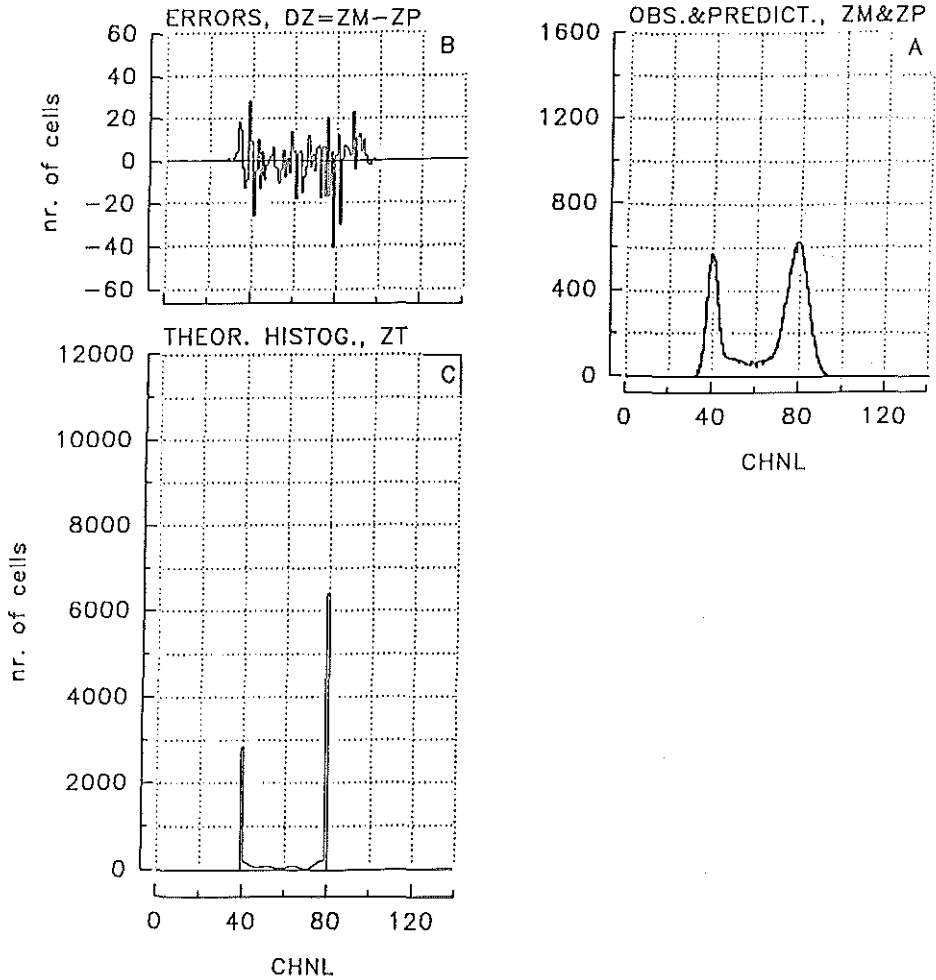


Fig. 5.19 Histogram Ba27; a) measured histogram (system output) and predicted histogram (optimum model output); b) deviations between measured and predicted histograms (measurement errors); c) corresponding theoretical histogram, assumed to be identical to the true histogram

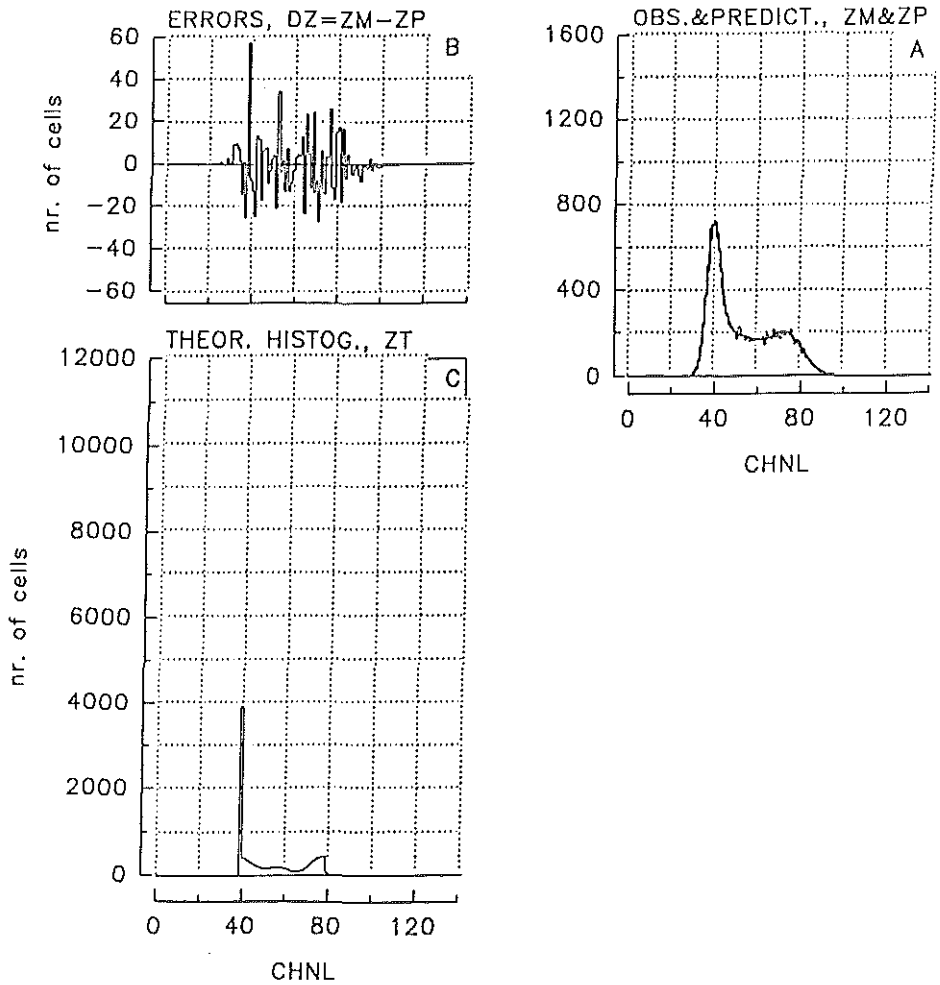


Fig. 5.20 Histogram Ba21; a) measured histogram (system output) and predicted histogram (optimum model output); b) deviations between measured and predicted histograms (measurement errors); c) corresponding theoretical histogram, assumed to be identical to the true histogram

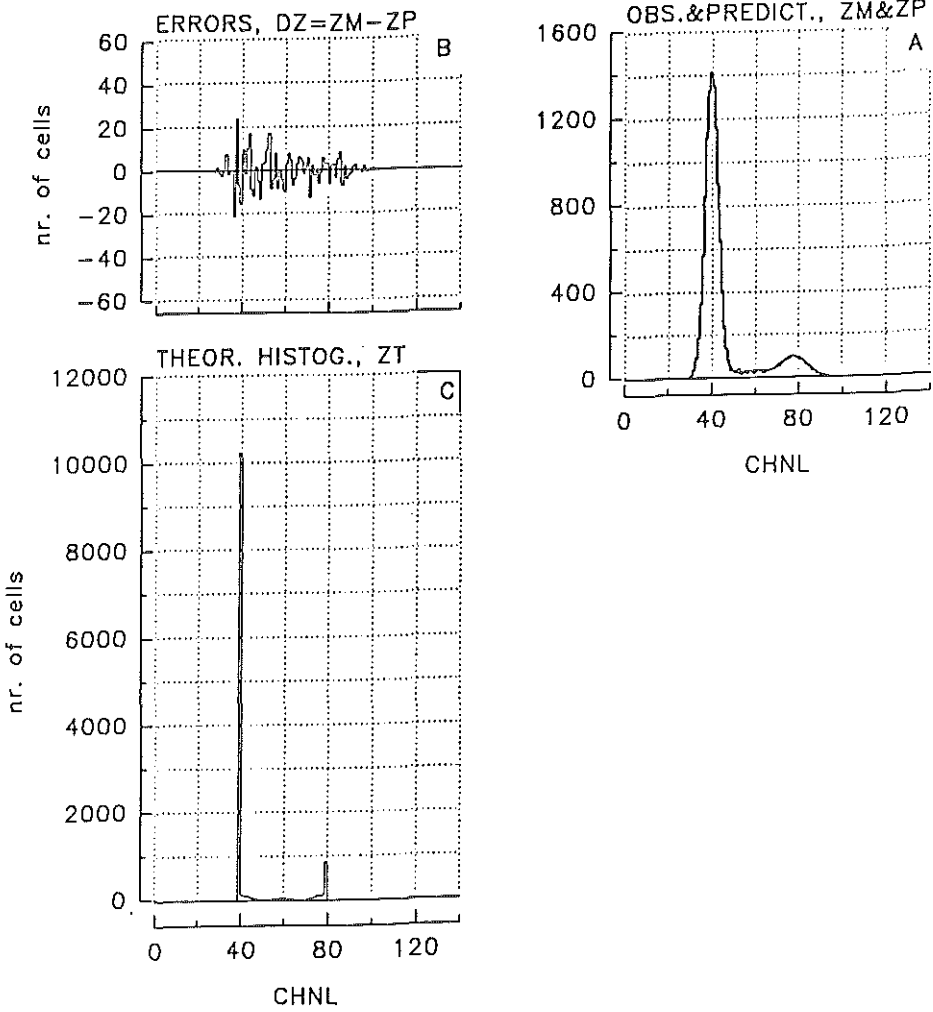


TABLE 5-17 RESULTS OF THE MULTIHARMONIC ANALYSIS TECHNIQUE; HISTOGRAMS WITH DISPROPORTIONAL PHASE FRACTIONS

| Histogram:                         | BA35; $G_1$ high, S low, |               |         | BA27; $G_1$ low, S high, |               |         | BA21; $G_1$ low, S medium, |               |         |        |
|------------------------------------|--------------------------|---------------|---------|--------------------------|---------------|---------|----------------------------|---------------|---------|--------|
| Population:                        | 13000 cells              |               |         | 13070 cells              |               |         | 13000 cells                |               |         |        |
| nr. of channels, M=                | 128                      |               |         | 128                      |               |         | 128                        |               |         |        |
| $G_1$ fraction in channel, $m_1=$  | 40                       |               |         | 40                       |               |         | 40                         |               |         |        |
| $G_2M$ fraction in channel, $m_2=$ | 80                       |               |         | 80                       |               |         | 80                         |               |         |        |
| <u>Parameter values:</u>           | initial                  | final $\pm$   | sd      | initial                  | final $\pm$   | sd      | initial                    | final $\pm$   | sd      |        |
|                                    |                          | (iteration 8) |         |                          | (iteration 6) |         |                            | (iteration 6) |         |        |
| nr. of $G_1$ cells, $Z_T[m_1]=$    | 8500.0                   | 10226.3       | 29.7    | 2200.0                   | 3884.7        | 53.9    | 3200.0                     | 2823.0        | 41.8    |        |
| nr. of $G_2M$ cells $Z_T[m_2]=$    | 2600.0                   | 857.5         | 37.9    | 800.0                    | 65.0          | 91.3    | 6600.0                     | 6399.7        | 58.4    |        |
| Gaussian spreading                 | PC=                      | 0.0525        | 0.0761  | 0.0050                   | 0.0525        | 0.1148  | 0.0072                     | 0.0350        | 0.0600  | 0.0013 |
|                                    | CF=                      | 0.0000        | -0.0204 | 0.1972                   | 0.0000        | -1.6283 | 0.2928                     | 0.0000        | -0.0842 | 0.0730 |
| Fourier coefficients               | A[1]=                    | 0.0           | 19.24   | 1.43                     | 0.0           | 62.22   | 3.33                       | 0.0           | 25.42   | 1.90   |
|                                    | A[2]=                    | 0.0           | 20.46   | 2.33                     | 0.0           | 33.21   | 5.33                       | 0.0           | 18.94   | 3.09   |
|                                    | A[3]=                    | 0.0           | -       | -                        | 0.0           | -       | -                          | 0.0           | 17.45   | 4.47   |
|                                    | A[4]=                    | 0.0           | -       | -                        | 0.0           | -       | -                          | 0.0           | -       | -      |
|                                    | A[5]=                    | 0.0           | -       | -                        | 0.0           | -       | -                          | 0.0           | -       | -      |
|                                    | B[1]=                    | 0.0           | 0.65    | 0.92                     | 0.0           | 5.64    | 2.22                       | 0.0           | 0.97    | 1.30   |
|                                    | B[2]=                    | 0.0           | -1.55   | 2.12                     | 0.0           | -20.51  | 5.33                       | 0.0           | -0.68   | 2.13   |
|                                    | B[3]=                    | 0.0           | -       | -                        | 0.0           | -       | -                          | 0.0           | -14.94  | 4.56   |
|                                    | B[4]=                    | 0.0           | -       | -                        | 0.0           | -       | -                          | 0.0           | -       | -      |
|                                    | B[5]=                    | 0.0           | -       | -                        | 0.0           | -       | -                          | 0.0           | -       | -      |
| <u>yield:</u>                      |                          |               |         |                          |               |         |                            |               |         |        |
| Goodness of fit                    |                          |               |         |                          |               |         |                            |               |         |        |
| $Z_p$ to $Z_M$ , TCC=              | 0.96856                  | 0.99983       |         | 0.93670                  | 0.99845       |         | 0.89925                    | 0.99906       |         |        |
| SD <sub>residuals</sub> :          | 72.6                     | 5.4           |         | 65.2                     | 10.4          |         | 85.8                       | 8.5           |         |        |
| Fourier coefficient, A[0]=         | 87.18                    | 49.13         | 1.24    | 258.20                   | 233.85        |         | 82.05                      | 96.85         |         |        |
| Composition of population:         |                          |               |         |                          |               |         |                            |               |         |        |
| fraction $G_1$ (%), $F_1=$         | 65.4                     | 78.7          | 0.2     | 16.8                     | 29.7          | 0.4     | 24.6                       | 21.7          | 0.3     |        |
| fraction S (%), $F_S=$             | 26.2                     | 14.7          | 0.4     | 77.1                     | 69.8          | 0.8     | 24.6                       | 29.1          | 0.6     |        |
| fraction $G_2M$ (%), $F_2=$        | 8.4                      | 6.6           | 0.3     | 6.1                      | 0.5           | 0.7     | 50.8                       | 49.2          | 0.5     |        |
| Dispersion:                        |                          |               |         |                          |               |         |                            |               |         |        |
| SD in channel $m_1$                | 2.10                     | 3.02          |         | 2.10                     | 3.26          |         | 1.40                       | 2.31          |         |        |
| SD in channel $m_2$                | 4.20                     | 6.07          |         | 4.20                     | 7.55          |         | 2.80                       | 4.72          |         |        |

TABLE 5-18 EFFECT OF GROWTH FACTOR ON THE CELL CYCLE OF A B-CELL HYBRIDOMA

| time (h)   | B-lymphocytes cultured |       |       |       |       |       |       |       |       |       |
|--|------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|  | WITHOUT                |       |       |       |       | WITH  |       |       |       |       |
|  | Growth Factor          |       |       |       |       |       |       |       |       |       |
|  | 2                      | 4     | 8     | 12    | 24    | 2     | 4     | 8     | 12    | 24    |
| <i>composition:</i>  |                        |       |       |       |       |       |       |       |       |       |
| %G <sub>1</sub>  | 28.8                   | 28.9  | 45.9  | 49.3  | 71.7  | 30.2  | 40.3  | 50.2  | 35.8  | 28.5  |
| sd <sub>%G<sub>1</sub></sub>                                   | 0.7                    | 0.4   | 0.7   | 0.5   | 0.6   | 0.5   | 0.6   | 1.1   | 0.5   | 0.4   |
| %S   | 57.5                   | 53.3  | 37.8  | 33.8  | 9.8   | 48.5  | 45.5  | 40.0  | 54.9  | 64.4  |
| sd <sub>%S</sub>   | 1.2                    | 0.8   | 1.1   | 0.8   | 1.1   | 0.8   | 0.9   | 1.4   | 0.8   | 0.5   |
| early  | 20.5                   | 20.2  | 12.8  | 11.4  | 3.3   | 13.3  | 13.8  | 8.4   | 16.3  | 20.3  |
| medium   | 19.4                   | 16.1  | 12.3  | 11.8  | 3.3   | 18.2  | 17.8  | 20.2  | 22.4  | 22.2  |
| late   | 17.6                   | 17.0  | 12.7  | 10.6  | 3.2   | 17.0  | 13.9  | 11.4  | 16.2  | 21.9  |
| %G <sub>2</sub> M  | 13.7                   | 17.8  | 16.3  | 16.9  | 18.5  | 21.3  | 14.2  | 9.8   | 9.3   | 7.1   |
| sd <sub>%G<sub>2</sub>M</sub>                                  | 0.9                    | 0.6   | 0.8   | 0.6   | 0.9   | 0.6   | 0.6   | 0.9   | 0.5   | 0.4   |
| <i>dispersion; in channel j, sd<sub>j</sub> = PC · j + CF:</i> |                        |       |       |       |       |       |       |       |       |       |
| PC   | 0.06                   | 0.07  | 0.06  | 0.06  | 0.12  | 0.05  | -0.01 | -0.02 | 0.01  | 0.01  |
| sd <sub>PC</sub>   | 0.01                   | <0.01 | <0.01 | <0.01 | 0.02  | <0.01 | <0.01 | 0.01  | 0.01  | <0.01 |
| CF   | 0.68                   | -1.04 | 1.24  | 0.18  | -3.88 | 1.68  | 7.36  | 12.01 | 7.38  | 3.95  |
| sd <sub>CF</sub>   | 0.52                   | 0.29  | 0.01  | 0.36  | 0.90  | 0.01  | 0.37  | 0.85  | 0.60  | 0.26  |
| <i>other variables:</i>  |                        |       |       |       |       |       |       |       |       |       |
| G <sub>1</sub> peak, m <sub>1</sub> =                          | 78                     | 74    | 75    | 77    | 76    | 75    | 88    | 82    | 82    | 72    |
| S-phase description,<br>N =                                    | 3                      | 4     | 3     | 3     | 0     | 2     | 2     | 2     | 3     | 4     |
| nr. iterations   |                        | 16    | 13    | 26    | 6     | 3     | 5     | 34    | 29    | 5     |
| Goodness of Fit,<br>TCC =                                      | .9929                  | .9975 | .9956 | .9977 | .9947 | .9950 | .9915 | .9868 | .9954 | .9976 |
| SD <sub>residuals</sub>  | 12.3                   | 8.1   | 9.0   | 14.1  | 19.6  | 8.3   | 22.7  | 19.1  | 11.2  | 10.6  |

continuous stirring. Cells were kept in fixative for several hours before treatment with RNase (SIGMA) for 15 minutes at room temperature (1 mg/ml in PBS). The cells then were washed and incubated with Propidium-Iodide at a concentration of 10 µg/ml PBS for at least 15 minutes before cytofluorometric analysis on a FACS-II from Becton-Dickinson. DNA histograms were prepared and analyzed with the computer program using the multiharmonic technique (Table 5-18). For comparison, a few histograms were also analyzed using the polynomial technique.

Except for the 2h histogram without HGF, for which both techniques yield similar results, the polynomial techniques estimates larger S fractions than the multiharmonic technique, at the expense of the G<sub>1</sub> and G<sub>2</sub>M fractions (Table 5-19).

The influence of the growth factor can clearly be seen. Especially beyond 8h the population incubated with growth factor shows a much larger

TABLE 5-19 COMPARISON OF MULTIHARMONIC (MH) AND POLYNOMIAL (P) ANALYSES

| time (h)   | B-lymphocytes cultured |       |      |      |       |      |       |       |      |      |
|--|------------------------|-------|------|------|-------|------|-------|-------|------|------|
|  | WITHOUT                |       |      |      |       | or   |       |       |      |      |
|  | Growth Factor          |       |      |      |       | WITH |       |       |      |      |
|  | 2                      | 4     | 8    | 12   | 24    | 2    | 4     | 8     | 12   | 24   |
| <i>composition:</i>  |                        |       |      |      |       |      |       |       |      |      |
| MH: %G <sub>1</sub>  | 28.8                   | 28.9  | 45.9 | 49.3 | 71.7  | 30.2 | 40.3  | 50.2  | 35.8 | 28.5 |
| %S   | 57.5                   | 53.3  | 37.8 | 33.8 | 9.8   | 48.5 | 45.5  | 40.0  | 54.9 | 64.4 |
| %G <sub>2</sub> M  | 13.7                   | 17.8  | 16.3 | 16.9 | 18.5  | 21.3 | 14.2  | 9.8   | 9.3  | 7.1  |
| <i>composition:</i>  |                        |       |      |      |       |      |       |       |      |      |
| P: %G <sub>1</sub>   | 29.4                   | -     | 40.6 | -    | -     | 23.9 | -     | -     | 26.4 | -    |
| %S   | 56.4                   | -     | 47.5 | -    | -     | 63.4 | -     | -     | 67.1 | -    |
| %G <sub>2</sub> M  | 14.2                   | -     | 12.0 | -    | -     | 12.7 | -     | -     | 6.5  | -    |
| MH: <i>dispersion; in channel j, sd<sub>j</sub> = PC·j + CF:</i> |                        |       |      |      |       |      |       |       |      |      |
| PC   | 0.06                   | 0.07  | 0.06 | 0.06 | 0.12  | 0.05 | -0.01 | -0.02 | 0.01 | 0.01 |
| CF   | 0.68                   | -1.04 | 1.24 | 0.18 | -3.88 | 1.68 | 7.36  | 12.01 | 7.38 | 3.95 |
| P: <i>dispersion; in channel j, sd<sub>j</sub> = CV·j:</i>       |                        |       |      |      |       |      |       |       |      |      |
| CV   | 0.08                   | -     | 0.08 | -    | -     | 0.06 | -     | -     | 0.10 | -    |

fraction of cells in S-phase. Without growth factor, the cells tend to accumulate in G<sub>1</sub>-phase (resting; Fig. 5.21).

**5.3.2.2 Influence of Imposing a Temporary Block in Early S-Phase on the Kinetics of a Simulated Exponentially Growing Cell Population.** Cells in exponential growth were subjected to a block in early S-phase. After 9 hours the block was released. Just before and at several time points during and after the block, the cell population's composition was observed by flow cytometric measurement of DNA histograms. Table 5-20, Figs 5.22 through 5.25 and 5.26 show the results after multiharmonic analyses. This sheds light on the cell kinetics. Just before the block there were 46% G<sub>1</sub> cells, 37% S cells and 17% G<sub>2</sub>M cells. During the block cells accumulated in G<sub>1</sub>-phase, leaving few S and very few G<sub>2</sub>M cells. Immediately after release of the block, proliferation continues: G<sub>1</sub> cells convert to S cells; the G<sub>1</sub> fraction decreases, the S fraction increases fast and the G<sub>2</sub>M fraction follows more slowly. Six hours after block release the G<sub>1</sub> fraction has passed its minimum and has started to increase again. The G<sub>2</sub>M fraction is still increasing, now at a more rapid rate. The S fraction, therefore, is growing smaller again. Nine hours after block release the original stationary phase has not yet been restored. This will take more time, during which, probably, the population's composition will show some damped oscillation.

Fig. 5.21 Influence of growth factor (HGF) on the cell kinetics of hybridoma B-lymphocytes. Cells cultured in the presence of HGF have larger S-phase fractions (proliferation), while without HGF the G<sub>1</sub>-phase is larger (resting cells)

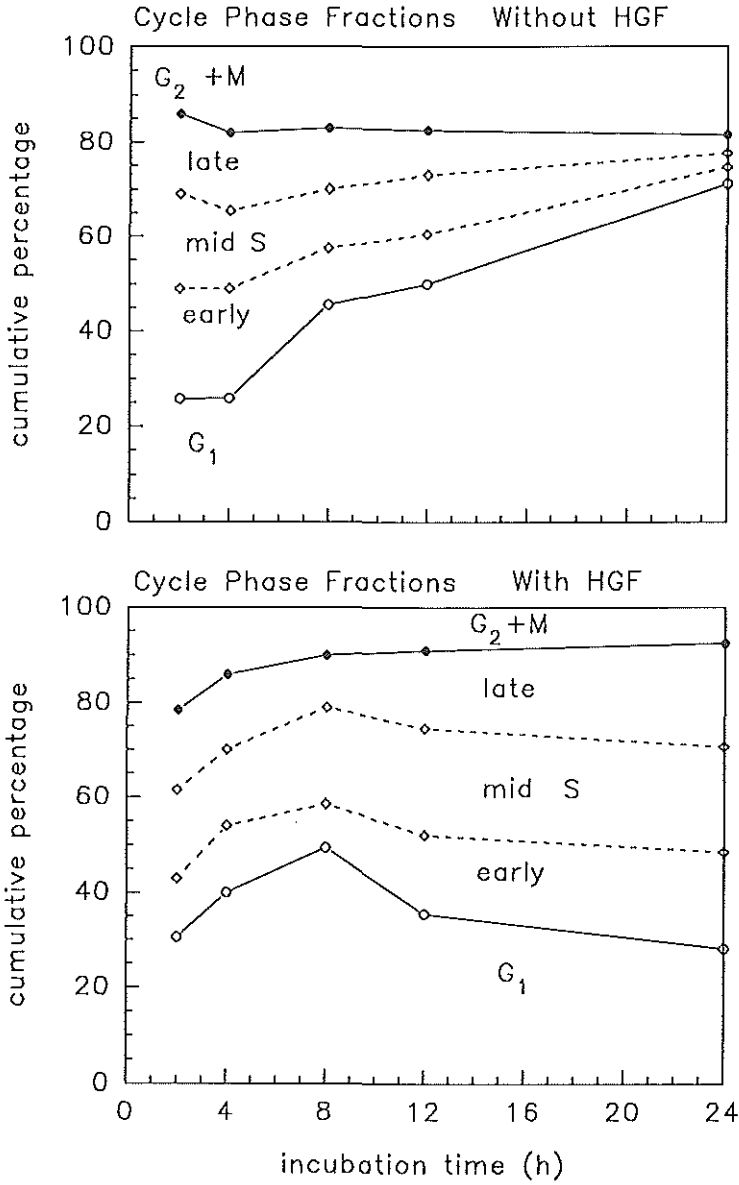


TABLE 5-20 RESULTS OF THE MULTIHARMONIC ANALYSIS TECHNIQUE; GRAY HISTOGRAMS

| Histogram:  | Gray 01; 0 h, start block in |               |          | Gray 03; 6 h after block |         |               | Gray 06; 15 h (6 h after |           |        |               |  |  |
|---|------------------------------|---------------|----------|--------------------------|---------|---------------|--------------------------|-----------|--------|---------------|--|--|
| Population:   | 1605817 cells; early S       |               |          | 2138720 cells            |         |               | 2442358; block release)  |           |        |               |  |  |
| nr. of channels, M=   | 128                          |               |          | 128                      |         |               | 128                      |           |        |               |  |  |
| G <sub>1</sub> fraction in channel, m <sub>1</sub> =            | 35                           |               |          | 35                       |         |               | 35                       |           |        |               |  |  |
| G <sub>2</sub> M fraction in channel, m <sub>2</sub> =          | 70                           |               |          | 70                       |         |               | 70                       |           |        |               |  |  |
| <u>Parameter values:</u>  | initial                      | final         | ± sd     | initial                  | final   | ± sd          | initial                  | final     | ± sd   |               |  |  |
|   |                              | (iteration 6) |          |                          |         | (iteration 6) |                          |           |        | (iteration 6) |  |  |
| nr. of G <sub>1</sub> cells, Z <sub>T</sub> [m <sub>1</sub> ]=  | 800000                       | 729197        | 9987     | 750000                   | 1687257 | 39292         | 226000                   | 232222    | 547    |               |  |  |
| nr. of G <sub>2</sub> M cells Z <sub>T</sub> [m <sub>2</sub> ]= | 280000                       | 284811        | 13770    | 425000                   | 142863  | 42404         | 185000                   | 131824    | 740    |               |  |  |
| Gaussian PC=  | 0.0400                       | 0.0606        | 0.0067   | 0.0600                   | 0.0119  | 0.0301        | 0.0400                   | 0.0407    | 0.0005 |               |  |  |
| spreading CF=   | 0.0000                       | -0.0220       | 0.2395   | 0.0000                   | 2.1280  | 1.0475        | 0.0000                   | 0.6554    | 0.0205 |               |  |  |
| Fourier coefficients  | A[1]=                        | 0.0           | -2316.16 | 512.81                   | 0.0     | 0.0           | 0.0                      | -11658.80 | 25.56  |               |  |  |
|   | A[2]=                        | 0.0           | -4796.78 | 709.22                   | 0.0     | -             | 0.0                      | -5956.94  | 34.53  |               |  |  |
|   | A[3]=                        | 0.0           | -        | -                        | 0.0     | -             | 0.0                      | -3780.46  | 58.07  |               |  |  |
|   | A[4]=                        | 0.0           | -        | -                        | 0.0     | -             | 0.0                      | -4467.90  | 116.52 |               |  |  |
|   | A[5]=                        | 0.0           | -        | -                        | 0.0     | -             | 0.0                      | -         | -      |               |  |  |
|   | B[1]=                        | 0.0           | -175.99  | 371.62                   | 0.0     | 0.0           | 0.0                      | 7656.52   | 16.55  |               |  |  |
|   | B[2]=                        | 0.0           | -1101.53 | 642.36                   | 0.0     | -             | 0.0                      | 1040.73   | 26.13  |               |  |  |
|   | B[3]=                        | 0.0           | -        | -                        | 0.0     | -             | 0.0                      | -274.71   | 55.22  |               |  |  |
|   | B[4]=                        | 0.0           | -        | -                        | 0.0     | -             | 0.0                      | -558.75   | 106.18 |               |  |  |
|   | B[5]=                        | 0.0           | -        | -                        | 0.0     | -             | 0.0                      | -         | -      |               |  |  |
| <u>yield:</u>   |                              |               |          |                          |         |               |                          |           |        |               |  |  |
| Goodness of fit   |                              |               |          |                          |         |               |                          |           |        |               |  |  |
| Z <sub>p</sub> to Z <sub>M</sub> , TCC=                         | 0.89891                      | 0.99618       |          | 0.84508                  | 0.97832 |               | 0.95666                  | 0.99999   |        |               |  |  |
| SD <sub>residuals</sub> :                                       | 12083.9                      | 2406.9        |          | 28437.0                  | 11016.4 |               | 1018.7                   | 110.3     |        |               |  |  |
| Fourier coefficient, A[0]=                                      | 15465.20                     | 17406.13      | 500.30   | 28344.70                 | 9076.46 |               | 59745.82                 | 61038.56  |        |               |  |  |
| Composition of population:                                      |                              |               |          |                          |         |               |                          |           |        |               |  |  |
| fraction G <sub>1</sub> (%), F <sub>1</sub> =                   | 49.8                         | 45.4          | 0.6      | 35.1                     | 78.9    | 1.8           | 9.2                      | 9.6       | 0.02   |               |  |  |
| fraction S (%), F <sub>S</sub> =                                | 32.8                         | 36.9          | 1.1      | 45.0                     | 14.4    | 2.9           | 83.2                     | 85.0      | 0.04   |               |  |  |
| fraction G <sub>2</sub> M (%), F <sub>2</sub> =                 | 17.4                         | 17.7          | 0.9      | 19.9                     | 6.7     | 2.3           | 7.6                      | 5.4       | 0.03   |               |  |  |
| Dispersion:   |                              |               |          |                          |         |               |                          |           |        |               |  |  |
| SD in channel m <sub>1</sub>                                    | 1.40                         | 2.10          |          | 2.10                     | 2.54    |               | 1.40                     | 2.08      |        |               |  |  |
| SD in channel m <sub>2</sub>                                    | 2.80                         | 4.22          |          | 4.20                     | 2.96    |               | 2.80                     | 3.50      |        |               |  |  |



TABLE 5-20 CONTINUED

Histogram: Gray 07; 18 h (9 h after  
 Population: 2612820 cells block release)  
 nr. of channels, M= 128  
 G<sub>1</sub> fraction in channel, m<sub>1</sub>= 35  
 G<sub>2</sub>M fraction in channel, m<sub>2</sub>= 70

Parameter values:

|   | initial       | final     | ±      | sd |
|---|---------------|-----------|--------|----|
|   | (iteration 6) |           |        |    |
| nr. of G <sub>1</sub> cells, Z <sub>T</sub> [m <sub>1</sub> ]=460000  | 376615        | 11513     |        |    |
| nr. of G <sub>2</sub> M cells Z <sub>T</sub> [m <sub>2</sub> ]=800000 | 793260        | 15019     |        |    |
| Gaussian PC=  | 0.0400        | 0.0548    | 0.0031 |    |
| spreading CF=   | 0.0000        | -0.1142   | 0.1512 |    |
| Fourier A[1]=   | 0.0           | -1628.50  | 586.60 |    |
| coefficients A[2]=  | 0.0           | -5078.59  | 792.27 |    |
| A[3]=   | 0.0           | -         | -      |    |
| A[4]=   | 0.0           | -         | -      |    |
| A[5]=   | 0.0           | -         | -      |    |
| B[1]=   | 0.0           | -14616.65 | 424.31 |    |
| B[2]=   | 0.0           | -7959.60  | 712.26 |    |
| B[3]=   | 0.0           | -         | -      |    |
| B[4]=   | 0.0           | -         | -      |    |
| B[5]=   | 0.0           | -         | -      |    |

yield:  
 Goodness of fit  
 Z<sub>p</sub> to Z<sub>M</sub>, TCC= 0.92460 0.99726  
 SD<sub>residuals</sub>: 14534.20 2819.6  
 Fourier coefficient, A[0]= 10000.00 42439.53 556.59  
 Composition of population:  
 fraction G<sub>1</sub> (%), F<sub>1</sub>= 17.6 14.4 0.4  
 fraction S (%), F<sub>S</sub>= 51.8 55.2 0.7  
 fraction G<sub>2</sub>M (%), F<sub>2</sub>= 30.6 30.4 0.6  
 Dispersion:  
 SD in channel m<sub>1</sub> 1.40 2.03  
 SD in channel m<sub>2</sub> 2.80 3.95

Fig.s 5.22 through 5.25  
 (on following pages)

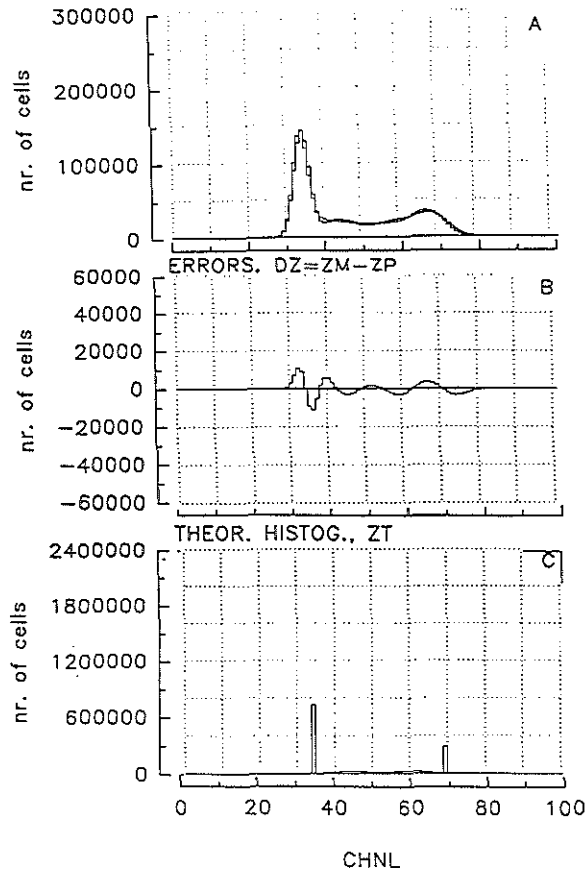
Sequential histograms (Gray)

- a) measured histogram (system output) and predicted histogram (optimum model output)
- b) deviations between measured and predicted histograms (measurement errors)
- c) corresponding theoretical histogram, assumed to be identical to the true histogram

respectively,  
 at start of block in early S-phase (0 h),  
 6 h after start of block,  
 15 h after start of block = 6 h after block release,  
 and 9 h after block release

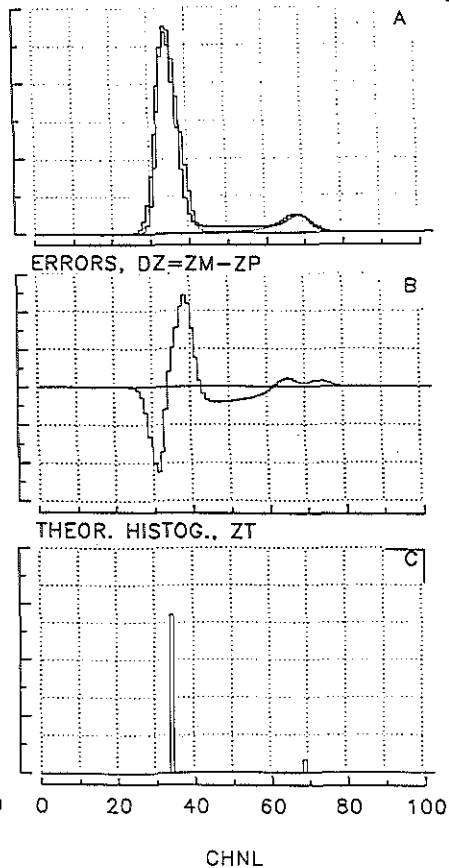
01

OBS. &amp; PREDICTED, ZM &amp; ZP



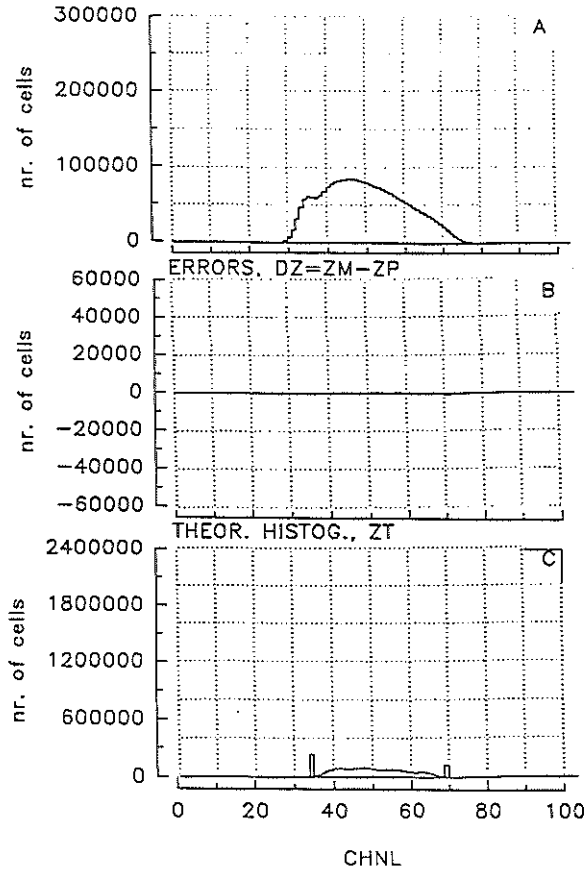
OBS. &amp; PREDICTED, ZM &amp; ZP

03



06

OBS. &amp; PREDICTED, ZM &amp; ZP



OBS. &amp; PREDICTED, ZM &amp; ZP

07

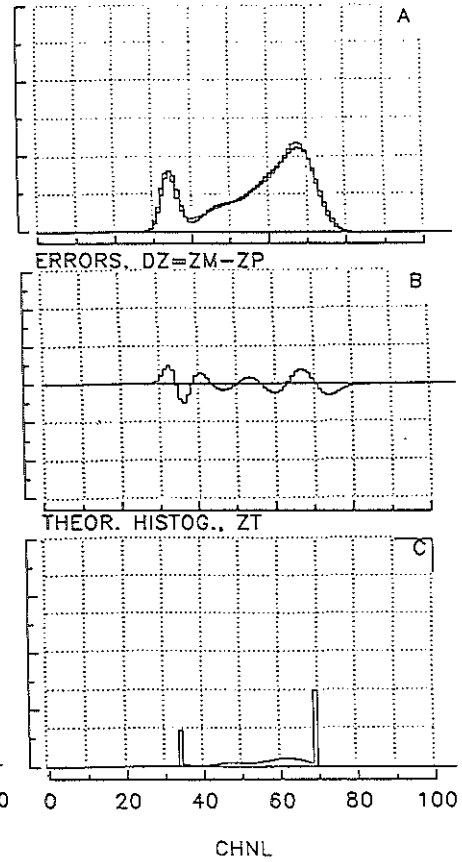
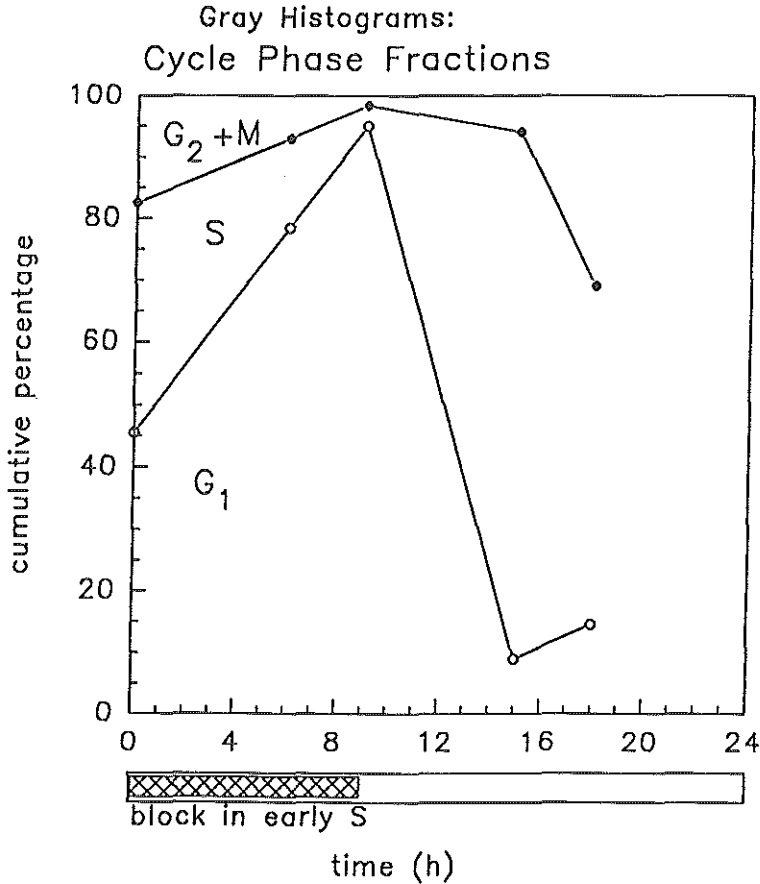


Fig. 5.26 Influence of a 9 h block in early S-phase on the kinetics of a cell population that originally grows in an exponential way



### 5.3.3 Histograms from BN Rat Tissues, With and Without infiltrated BNML Cells

Flow cytometric measurements were performed with cell suspensions derived from BN rat liver, spleen and (femoral) bone marrow on day 10 after the *i.v.* inoculation of  $10^7$  BNML cells. DNA histograms were made for all cells from a particular tissue (total nucleated cells, TNC) and for the malignant cells among them. These cells could be separated in the cell sorter device—a part of the FCM (FACS, fluorescence activated cell sorter)—because their surfaces had been marked with a fluorescent dye (FITC), attached to a monoclonal antibody (RM124), which binds to a receptor that is present in large numbers on the BNML cell surface.

From the bone marrow, samples were also taken at days 2 and 17 after  $10^7$  BNML cells. Before analysis these cell suspensions were prepared in two different ways, i.e., with or without adding a substance called nycodenz. This medium enables separation of non-proliferating cells. Finally, various subpopulations were sorted from the normal bone marrow cells.

Results are shown in Tables 5-21 through 5-24. Fig. 5.27 is an illustration of analysis results for the bone marrow on day 10. Looking at the day 10 tissues overall, the bone marrow and spleen show a similar behavior: about 80% of the cells are in  $G_1$ -, 15-20% in S-phase. In the liver only 50% of the cells are in  $G_1$ -, and almost no cells are in S-phase. This last result, however, may be an artifact because much fewer were analyzed and the fitting results were rather poor.

Looking at the malignant cells, again the liver distinguishes itself from the other two organs although in a less pronounced way. The liver has fewer malignant cells in  $G_1$ -, but—in contrast to the total cell population—many more cells in S-phase.

When looking at the BNML growth curves observed in these organs (Fig. 2.15), on day 10 after inoculation of  $10^7$  BNML cells the malignant population in the liver and in the spleen are still in the phase of exponential growth, whereas in the bone marrow conversion to Gompertzian growth has already taken place then. Therefore, one would expect fewer S-phase cells in the bone marrow as compared to liver and spleen (more proliferation in these last organs). This is not reflected by the DNA-histogram analysis. An explanation might be that conversion to Gompertzian growth sooner is a result of increased cell loss than a result of less proliferation.

Without nycodenz the bone marrow results tend to be a little higher in  $G_1$ - and  $G_2$ M-phase and lower in S. This could be expected, as the nycodenz treatment will have removed non-proliferating cells. Differences, however, are not large.

On day 2 insufficient BNML cells could be separated from the total population, the fraction RM124 positive cells still being too small.

Looking at the total population, with increase in time the S-phase fraction seems to increase and the  $G_1$  fraction seems to decrease. This suggests that at later stages there is more proliferation.

Apparently, the increased proliferation is not due to the BNML cells. The fraction of BNML cells in S-phase remains the same or decreases with time in the 10-17 d interval. Experiments with  $^3\text{H}$ -thymidine showed decreased S fraction as well [Martens and Hagenbeek, 1977]. From these data it follows that decreased proliferation also must make some contribution to the observed Gompertzian growth of BNML in bone marrow at this time interval (next to increased cell loss).

It should be noted here that, as at present it is difficult to obtain repro-

TABLE 5-21 CYCLE PHASE FRACTIONS OF TOTAL NUCLEATED CELLS (TNC) AND MALIGNANT CELLS (RM124<sup>+</sup>) IN VARIOUS TISSUES, AT DAY 10 AFTER INOCULATION OF 10<sup>7</sup> BNML CELLS

| organ       |                      | TNC        | RM124 <sup>+</sup> |
|-------------|----------------------|------------|--------------------|
| bone marrow | %G <sub>1</sub>      | 77.3 ± 0.3 | 77.0 ± 0.8         |
|             | %S (± sd)            | 18.0 ± 0.5 | 22.0 ± 1.0         |
|             | %G <sub>2</sub> M    | 4.7 ± 0.4  | 1.0 ± 0.5          |
|             | nr.of cells analyzed | 44880      | 7812               |
|             | goodness of fit      | 0.9992     | 0.9939             |
| liver       | %G <sub>1</sub>      | 49.0 ± 2.0 | 70.5 ± 0.8         |
|             | %S (± sd)            | 1.4 ± 3.6  | 29.1 ± 1.3         |
|             | %G <sub>2</sub> M    | 49.6 ± 3.0 | 0.4 ± 1.0          |
|             | nr.of cells analyzed | 24640      | 5530               |
|             | goodness of fit      | 0.9778     | 0.9948             |
| spleen      | %G <sub>1</sub>      | 81.1 ± 0.4 | 83.4 ± 1.1         |
|             | %S (± sd)            | 14.8 ± 0.8 | 14.6 ± 1.8         |
|             | %G <sub>2</sub> M    | 4.1 ± 0.7  | 2.0 ± 1.4          |
|             | nr.of cells analyzed | 43680      | 5412               |
|             | goodness of fit      | 0.9982     | 0.9854             |

TABLE 5-22 CYCLE PHASE FRACTIONS OF TOTAL NUCLEATED CELLS (TNC) AND MALIGNANT CELLS (RM124<sup>+</sup>) IN BONE MARROW AS FUNCTION OF TIME AND PRESENCE OF NYCODENZ

|                    |                   | time (d) after 10 <sup>7</sup> BNML cells i.v. |            |            |
|--------------------|-------------------|--|------------|------------|
|                    |                   | 2  | 10         | 17         |
| TNC                | %G <sub>1</sub>   | 74.9 ± 0.6                                     | -          | 62.5 ± 0.8 |
|                    | %S (± sd)         | 14.6 ± 1.2                                     | -          | 33.1 ± 1.6 |
|                    | %G <sub>2</sub> M | 10.5 ± 1.0                                     | -          | 4.4 ± 1.4  |
| RM124 <sup>+</sup> | %G <sub>1</sub>   | -  | 54.2 ± 0.6 | 62.4 ± 0.8 |
|                    | %S (± sd)         | -  | 40.0 ± 1.2 | 27.0 ± 1.4 |
|                    | %G <sub>2</sub> M | -  | 5.8 ± 1.0  | 10.6 ± 1.1 |
| TNC                | %G <sub>1</sub>   | -  | 81.8 ± 0.7 | 66.9 ± 0.5 |
|                    | %S (± sd)         | -  | 13.9 ± 1.3 | 29.2 ± 0.8 |
|                    | %G <sub>2</sub> M | -  | 4.3 ± 1.1  | 3.9 ± 0.6  |
| RM124 <sup>+</sup> | %G <sub>1</sub>   | -  | 60.2 ± 0.7 | 62.9 ± 0.6 |
|                    | %S (± sd)         | -  | 32.4 ± 1.1 | 30.1 ± 1.1 |
|                    | %G <sub>2</sub> M | -  | 7.4 ± 0.9  | 7.0 ± 1.0  |

nyco<sup>+</sup>

nyco<sup>-</sup>

TABLE 5-23 CYCLE PHASE COMPOSITION OF VARIOUS SUBPOPULATIONS OF NORMAL BONE MARROW CELLS IN THE BN RAT

| cell type:         | total BM   | lymphocytes | granulocytes | blasts     |
|--------------------|------------|-------------|--------------|------------|
| %G <sub>1</sub>    | 78.5 ± 0.5 | 98.0 ± 0.6  | 82.6 ± 0.5   | 52.0 ± 0.5 |
| %S (± sd)          | 16.9 ± 0.9 | 2.3 ± 1.0   | 16.0 ± 0.8   | 44.9 ± 0.8 |
| %G <sub>2</sub> M  | 4.5 ± 0.7  | -0.3 ± 0.7  | 1.4 ± 0.7    | 3.1 ± 0.6  |
| CV <sub>peak</sub> | 0.039      | 0.030       | 0.036        | 0.045      |
| nr.of cells anal.  | 7161       | 3600        | 6365         | 10692      |
| goodness of fit    | 0.9965     | 0.9966      | 0.9967       | 0.9965     |

*Lymphocytes and granulocytes populations have few S cells. Proliferation in bone marrow is largely due to blast cells (large S-phase fraction)*

TABLE 5-24 INFLUENCE OF SAMPLE SIZE ON ANALYSIS RESULTS

| case | nr. of cells | %G <sub>1</sub> ± sd | %S ± sd    | %G <sub>2</sub> M ± sd | peak channel pos., CV | TCC   |
|------|--------------|----------------------|------------|------------------------|-----------------------|-------|
| 1a   | 9354         | 68.0 ± 1.2           | 27.4 ± 2.0 | 4.6 ± 1.6              | 45 0.07               |       |
| 1b   | 4560         | 67.7 ± 0.4           | 27.7 ± 0.7 | 4.6 ± 0.6              | 88 0.03               |       |
| 2a   | 2160         | 66.8 ± 1.3           | 24.7 ± 2.2 | 8.5 ± 1.8              | 102 0.07              |       |
| 2b   | 1341         | 70.5 ± 1.0           | 24.8 ± 1.5 | 4.7 ± 1.2              | 103 0.07              |       |
| 2c   | 803          | 79.6 ± 1.0           | 18.3 ± 1.8 | 2.1 ± 1.4              | 103 0.07              |       |
| 2d   | 446          | 68.5 ± 1.6           | 27.7 ± 2.6 | 3.8 ± 2.0              | 102 0.07              |       |
| 3a   | 10692        | 52.0 ± 0.5           | 44.9 ± 0.8 | 3.1 ± 0.6              | 0.04                  | .9965 |
| 3b   | 3800         | 46.1 ± 0.6           | 50.6 ± 1.0 | 3.3 ± 0.8              | 0.03                  | .9937 |

The above cases with a same number are always taken from the same sample (prepared cell suspension); a-d denote different analysis runs.

- 1: RM124<sup>+</sup> cells (BNML) sorted from a sample of 10<sup>7</sup> bone marrow cells
- 2: RM124<sup>+</sup> cells (BNML) sorted from a sample of 10<sup>6</sup> bone marrow cells
- 3: blast cells from normal rat bone marrow

ducible results, firm conclusions cannot be drawn. Comparing the day 10 data in Tables 5-21 and 5-22 reveals that both for BNML and TNC in bone marrow large interexperimental fluctuations in phase fractions occur. Of course, this may be largely due to inevitable biological variation. But also the production of DNA histograms needs further development and standardization. For example, while the goodness of fit (total correlation coefficient, TCC) remains fairly constant, reducing the number of cells analyzed clearly influences the analysis

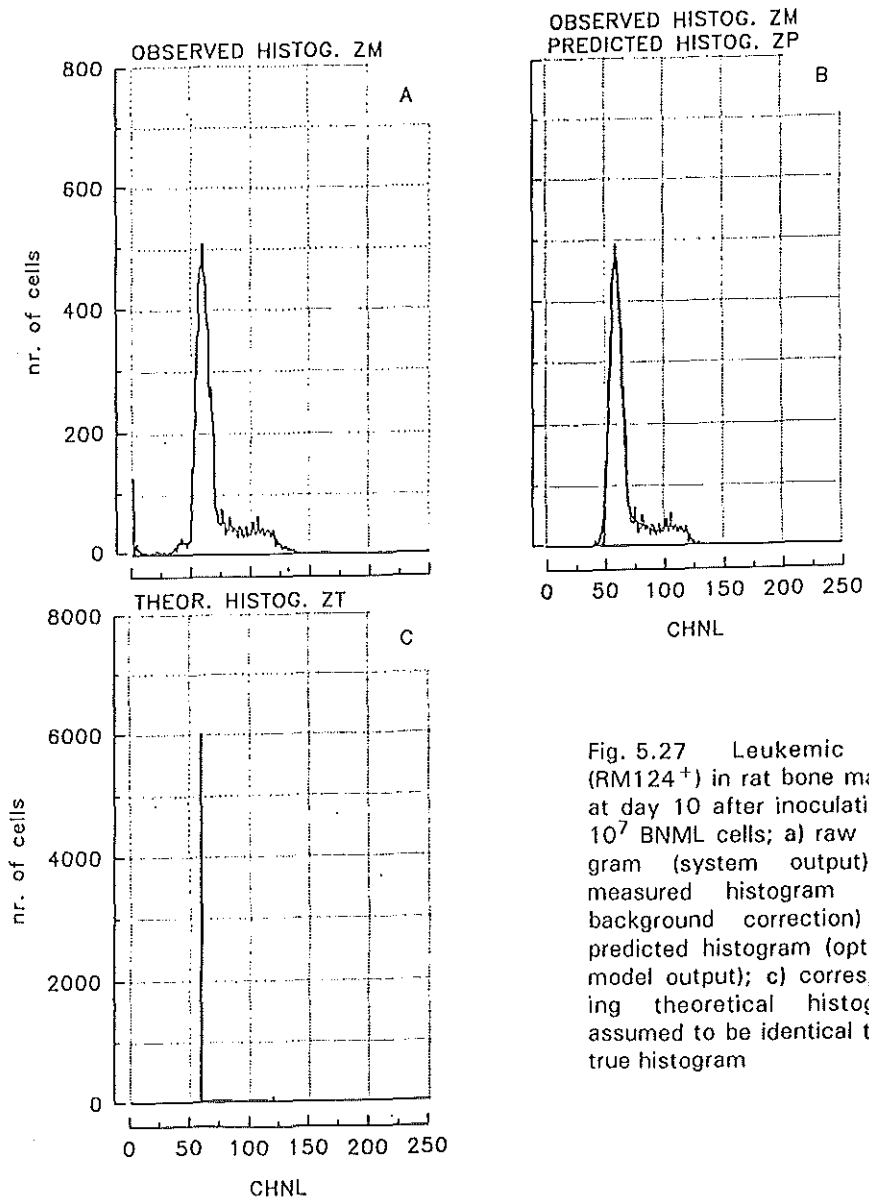


Fig. 5.27 Leukemic cells (RM124<sup>+</sup>) in rat bone marrow at day 10 after inoculation of 10<sup>7</sup> BNML cells; a) raw histogram (system output); b) measured histogram (after background correction) and predicted histogram (optimum model output); c) corresponding theoretical histogram, assumed to be identical to the true histogram



results (Table 5-24). Especially when the cell numbers get very low (less than a few thousands), fluctuations of 10% in phase fractions may be seen. As the model, as used in the analysis, assumes Gaussian redistribution of cells—an assumption that holds better when the numbers of cells are large—the number of cells to be analyzed should be chosen as large as possible.

#### 5.4 CONCLUDING REMARKS

A method has been presented that models the production of a single DNA histogram by FCM. By proper adjustment of the parameter values in the model equations, which is taken care of in an iterative numerical procedure, a given observed histogram can be analyzed, i.e., the cycle phase composition of the cell population under investigation can be determined. The method is to a large extent automatic and offers objective analysis of raw data. It includes the possibility of estimating the accuracies of the resulting parameter values and model response. Flexibility is a major feature. A kernel function describes the actual S-phase cell distribution with a set of harmonic functions. This set is characterized by a few parameters only that, however, by choice of proper values, allow a great variety of shapes. Therefore, it can be expected that this analysis method will be applicable to many different classes of DNA histograms. The constraints imposed by employment of a polynomial in the kernel function are thus relaxed, while the risk of introducing too many free parameters and encountering numerical problems (divergence, as may happen in the multiGaussian method if many S-phase spikes are required in the kernel function), is avoided.

The S-phase in the multiharmonic method has a distribution that is continuous through the DNA channels. Although uncertainty in the estimated Fourier coefficients is considerable, this allows a better representation of the physical reality than an approximation with a kernel function that consists of a number of spikes with empty spaces in between (for example, compare the theoretical histogram obtained by the multiharmonic method to the multiGaussian solution for FRAMA in Fig. 5.9c/d).

This also leaves the possibility to estimate not only the total fraction of S-phase cells but, in principle, also to give subfraction sizes explicitly. This may yield information on the rate of DNA synthesis, especially if sequential DNA histograms can be simultaneously analyzed. Adding a third dimension to the model, i.e., a mathematical expression that states how the contents of each specific channel will change with time (in other words, modeling the proliferation kinetics), is a logical next step. Predicted histograms at subsequent times can be calculated. By matching this model response to the actually observed series of histograms, information on cell proliferation characteristics may be obtained (e.g., [Yanagisawa et al., 1985]). This approach probably may yield

more accurate results for each separate DNA histogram than can be achieved by analyzing them one at a time [Lampariello et al., 1989]. Also, it may prove a good alternative for the double labeling experiments that are recently being developed (e.g., [Lacombe et al., 1988]) to assess cell kinetic information (treating cells with bromodeoxyuridine (BrdUrd) next to Propidium Iodide for DNA staining, and using a fluorescent monoclonal antibody against BrdUrd, the  $G_1$ , S and  $G_2M$  fractions can be distinguished on two-dimensional FCM plots of the fluorescence intensities).

Testing the multiharmonic method has thus far yielded promising results, suggesting the adequateness of this modeling approach, although only conventional, rather "clean" histograms have been analyzed. *Its better performance with respect to the multiGaussian and polynomial techniques has been shown above for several histograms.* Of course, histograms originating from various types of cell populations perturbed with cycle-specific agents should be analyzed with the multiharmonic method to prove its practical, general usefulness. A problem may arise when a channel of a theoretical histogram contains only very few cells. The subsequent redistribution among the neighbor channels then will be according to binomial statistics, whose tendency to the Gaussian function disappears with decreasing cell numbers. The assumption of normally redistributed quantities (Eq.(5.17)) thus may lose its validity. The consequences for the error model (variance of the residuals) should be looked into.

With respect to the present calculation of the variance of the residuals use of Eq.(5.29) may lead to some underestimation, since in general a histogram will not fully occupy the total range of M channels. To disregard the influence of irrelevant empty channels it might be better to truncate the range in Eq. (5.29), respectively, below  $m_1 - 3 \cdot SD_{m_1}$  and above  $m_2 + 3 \cdot SD_{m_2}$ .

The present analysis method relies upon a properly adjusted fluorescence amplification, i.e.,  $m_1$  must be sufficiently large so as to retain all "in front of the  $G_1$  peak" cells of the observed histogram in the channels 1 through  $m_1$ . But it also must be small enough not to spill redistributed cells in the tail of the histogram beyond the channel M boundary.

More operational experience in the application of the method must be accumulated, for instance, to enable a reasonable *à priori* choice of the minimum number of Fourier coefficients (N) that need to be considered for the best results. An improved routine that automatically searches for the best choice of  $G_1$  channel position  $m_1$  should be designed to reduce user interaction further. A presently available routine that optimizes  $m_1$  as a free parameter is too time consuming to be of practical use (as  $m_1$  occurs very frequently in many terms of the model equations, letting its value float means that many function evaluations must be performed, stretching required computation time to 15 minutes per iteration. At present it is more economic to run the program several times with

different fixed  $m_1$  values, and to choose the best model output by visual comparison). Finally, when and how to correct for the presence of cellular debris, background noise and occurrence of clumping also are problems that have only partially been solved.

## 5.5 REFERENCES

- Bagwell CB (1979) Theory and application of DNA histogram analysis. Thesis, University of Miami
- Baisch H, Göhde W and Linden WA (1975) Analysis of PCP-data to determine the fraction of cells in the various phases of the cell's cycle. *Rad Environm Biophys* 12:31-39
- Baisch H, Beck HP, Christensen IJ, Hartman NR, Fried J, Dean PN, Gray JW, Jett JH, Johnston DA, White RA, Nicolini C, Zietz S and Watson JV (1982) A comparison of mathematical methods for the analysis of DNA histograms obtained by flow cytometry. *Cell Tissue Kinet* 15:235-249
- Barlogie B, Drewinko B, Johnston DA, Büchner Th, Hauss WH and Freireich EJ (1976) Pulse cytophotometric analysis of synchronized cells in vitro. *Cancer Res* 36:1176-1181
- Beck HP (1980) Evaluation of flow cytometric data of human tumours; correction procedures for background and cell aggregations. *Cell Tissue Kinet* 13:173-181
- Brunsting A, Collins JM, Kane FR, Bagwell CB (1979) An examination of some basic assumptions of DNA distribution analysis using biological data. *Cell Tissue Kinet* 12:123-134
- Dean PN and Jett JH (1974) Mathematical analysis of DNA distributions derived from flow microfluorometry. *J Cell Biol* 60:523-527
- Dean PN, Dolbear F, Gratzner H, Rice GC and Gray JW (1984) Cell-cycle analysis using a monoclonal antibody to BrdUrd. *Cell Tissue Kinet* 17:427-436
- Edwards AWF (1972) *Likelihood*. Cambridge, New York
- Eijkhoff P (1974) *System identification, state and parameter estimation*. Wiley, New York
- Fried J (1976) Method for the quantitative evaluation of data from flow microfluorometry. *Comp Biomed Res* 9:263-276
- Fried J and Mandel M (1979) Multi-user system for analysis of data from flow cytometry. *Comp Prog Biomed* 10:218-230
- Fried J, Perez AG and Clarkson BD (1978) Rapid hypotonic method for flowcytometry of monolayer cell cultures: some pitfalls in staining and data analysis. *J. Histochem. & Cytochem.* 26:921-933
- Fulwyler MJ (1980) Flow cytometry and cell sorting. *Blood Cells* 6:173-184
- Gray JW (1974) Cell cycle analysis from computer synthesis of deoxyribonucleic acid histograms. *J Histochem Cytochem* 22:642-650
- Gupta NK and Mehra RK (1974) Computational aspect of maximum likelihood estimation and reduction in sensitivity function calculations. *IEEE Trans Automat Contr* AC19:774-783
- Karon M and Shirakawa S (1970) Effects of 1-beta-D arabinofuranosyl cytosine on cell passage time. *J Natl Cancer Inst* 45:861-867
- Kerker M, Van Dilla MA, Brunsting A, Krahtovil JP, Hsu P, Wang DS, Gray JW and Langlois RG (1982) Is the central dogma of flow cytometry true that fluorescence intensity is proportional to cellular dye content? *Cytometry* 3:71-78
- Klein KD, Adler D, Overing M, Klein PJ and Lennartz KJ (1976) Investigation on pharmacologic induction of partial synchronization of tumor cell proliferation; its relevance for cytostatic therapy. *Cancer Treatm Rep* 60:1959-1979
- Lacombe F, Belloc F, Bernard P, Boisseau MR (1988) Evaluation of four methods of DNA

- distribution data analysis based on bromodeoxyuridine/ DNA bivariate data. *Cytometry* 9:245-253
- Lampariello F, Mauro F, Uccelli R, Spanò M (1989) Automatic analysis of flow cytometric DNA histograms from irradiated mouse male germ cells. *Cytometry* 10:62-69
- Lansdorp PM, Aarden LA, Calafat J and Zeiljemaker WP (1986) A growth-factor dependent B-cell hybridoma. *Curr Top Microbiol Immunol* 132:105-113
- Martens ACM, Hagenbeek A (1977) Pulse cytophotometry of the BN myelocytic leukaemia during development and during treatment with cytostatic drugs. *Leukemia Res* 1:103-107
- Pinkel D and Steen HB (1982) Simple methods to determine and compare the sensitivity of flow cytometers. *Cytometry* 3:220-223
- Schultz FW, Sonneveld P, Mulder JA (1985) Identification of the flow cytometer system for DNA histogram analysis purposes. In: Barker HA, Young PC (eds) *Identification and system parameter estimation 1985, 1*, Proc. 7th IFAC/IFORS Symposium; Pergamon Press, Oxford, 1131-1136
- Sheck LE, Muirhead KA and Horan PK (1980) Evaluation of S phase distribution of flow cytometric DNA histograms by autoradiobiography and computer algorithms. *Cytometry* 1:109-117
- Sonneveld P and Mulder JA (1981) Development and identification of a multicompartment model for the distribution of adriamycin in the rat. *J Pharm Biophar* 9:577-601
- Sonneveld P, Schultz FW, Mulder JA (1988) New method for the analysis of flow cytometric data. *Anal Quant Cytol Histol* 10(4):261-268
- Tobey RA, Crissman HA and Oka MS (1976) Arrested and cycling CHO cells as a kinetic model; studies with adriamycin. *Cancer Treatm Rep* 60:1829-1837
- Trask BJ (1985) *Studies of chromosomes and nuclei using flow cytometry*. Thesis, Leiden University, The Netherlands
- Tuma JJ (1979) *Engineering Mathematics Handbook*. McGraw-Hill Book Company, New York
- Yanagisawa M, Dolbeare F, Todoroki T, Gray JW (1985) Cell cycle analysis using numerical simulation of bivariate DNA/bromodeoxyuridine distributions. *Cytometry* 6: 550-562
- Åström KJ (1979) Maximum likelihood and prediction error methods. In: *Tutorials on system identification* (Isermann R, ed) Institut für Regelungstechnik, Technische Hochschule Darmstadt, p53-67

## Chapter 6

# Bone Marrow Transplantation and Risk of Leukemia Relapse

In this chapter it is shown that the application of simple statistical models to (pre)clinical data like relapse rates may be helpful for developing guidelines or rationales for treatment protocols. At least the models may help in establishing new directions for further investigations so as to make bone marrow transplantation a better treatment modality for leukemia.

## 6.1 Introduction to Bone Marrow Transplantation

### 6.1.1 TRANSPLANTATION

For several diseases, involving an organ that no longer functions properly, transplantation may be the solution; e.g., certain conditions being satisfied, a defective kidney can be taken out of a patient in order to be replaced by a normally working donor kidney, restoring the patient's health without too serious complications. For leukemia, a disease of the hemopoietic system, a solution was sought in an analogous way. In principle, a patient's hemopoietic system can be renewed by means of bone marrow transplantation (BMT).

### 6.1.2 BONE MARROW, BLOOD CELLS AND HEMOPOIESIS

Bone marrow is present inside every bone of the body, in particular in the pelvis, the vertebrae, the ribs and the long bones of arms and legs. It is the main source of fresh blood cells, which are continuously being produced to replace old and worn ones (rate of production = rate of loss  $\approx 10^{11}$ - $10^{12}$  cells per day).

The various types of functional blood cells, such as erythrocytes, leukocytes and platelets are all descendants of a single cell type, i.e., the pluripotent hemopoietic stem cell (HSC). The HSC incessantly produces offspring that either is again an HSC or a committed cell (CC) that stands at the origin of a specific blood cell line. I.e., in a process of proliferation and stepwise differentiation—each new differentiation step reduces the proliferative capacity—CCs mature into the (non-proliferating) end cells of their (myelocytic, lymphocytic, erythrocytic, thrombocytic) line. Toward the end of the maturation the blood cells leave the bone marrow to start circulating and doing their specific functions throughout the body.

### 6.1.3 BONE MARROW TRANSPLANTATION PROCEDURE

A malfunctioning hemopoietic system, such as happens in leukemia, must be destroyed first before it can be replaced. This is usually done by total body irradiation and/or high-dose chemotherapy. The so-called conditioning therapy (CT) not only stops the patient's hemopoiesis by killing all HSCs, it also creates space for the incoming graft. Furthermore, the patient's immunological defense system is destroyed, which means that there will be little risk that the graft is rejected in an immunological reaction to the donor marrow that is foreign to the patient's body. It also means that during this period the immunodeficient patient is very vulnerable to all sorts of infections, so must be nursed in a protective environment.

Donor bone marrow cells administered intravenously have been shown to home to and occupy the vacated sites in the patient's bones. Relatively few bone marrow cells—in fact, even less HSCs among these—appear to be needed to restore the complete process of hemopoiesis eventually [Visser et al., 1984].

### 6.1.4 TYPES OF BONE MARROW TRANSPLANTATION

Four forms of BMT are distinguished, depending on the relation between donor and recipient.

With *autologous* BMT the patient is rescued with his own bone marrow. In leukemia the bone marrow is harvested before the marrow ablative CT is applied, but after the patient has attained a state of complete remission (CR). The thus obtained graft is believed to contain a negligible amount of leukemic cells, if any, but sufficient normal HSCs. The patient is hoped to be cured from leukemia by the CT.

If donor and recipient are genetically identical the graft is called *syngeneic* or *isologous*. This is the case with identical twins (or when animals of the same inbred strain are concerned).

BMT between a genetically different donor and host of the same species is called *allogeneic*. If BMT is applied interspecies, then it is called *xenogeneic*.

### 6.1.5 PROBLEMS IN BONE MARROW TRANSPLANTATION FOR LEUKEMIA; ALLOGENEIC; GRAFT-VERSUS-HOST DISEASE, AUTOLOGOUS; REINOCULATION OF LEUKEMIC CELLS

Obviously, with syngeneic or allogeneic BMT a healthy donor is selected. Therefore the graft certainly contains no leukemic cells, whereas it may contain malignant cells in case of autologous BMT. However, often a suitable (genetically matched) healthy donor is not readily available. Then, autologous BMT remains as one of the treatment modalities.

After allogeneic BMT there is a risk of facing another complication, namely, graft-versus-host disease (GvHD). This means that the graft cells evoke an immunological reaction against the foreign cells of the recipient host, which can be classified as mild or severe, but may end fatally. GvHD is known in two forms, acute and chronic, occurring early and late, respectively, after transplantation.

Certain measures to prevent GvHD can be taken, however. For instance, before BMT the graft may be manipulated so as to remove the T-lymphocytes that, as was experimentally established in a mouse model, are the cells responsible for causing the immune reaction [Van Bekkum and Löwenberg, 1985]. Alternatively, the patient may receive additional immunosuppressive treatment, e.g., cyclosporin A (CyA) and/or methotrexate (MTX), as a measure of GvHD-prophylaxis.

It has been observed that leukemia patients who develop GvHD after (allogeneic) BMT have better chances to remain in remission than those who do not. The induced immune reaction by the graft apparently attacks residual leukemic cells left over in the patient as well. Therefore, more and more clinicians tend to tolerate the occurrence of GvHD, even if it takes a toll in the form of GvHD related morbidity. Still, curing a patient's leukemia, but losing him because of GvHD, is an embarrassment to the potential successfulness of the therapy.

### 6.1.6 MATHEMATICAL MODELING

Through mathematical modeling it can be explored to what extent the graft-versus-leukemia reaction (GvLR), being part of the GvHD, contributes to the probability of curing leukemia after *allogeneic* BMT. It will be shown below that, although leukemia relapse rates decrease dramatically, a GvLR probably causes an equivalent of about 1 log cell kill (LCK) only. If this is true, then other treatment strategies—that induce approximately 1 LCK leukemic cell eradication while preventing the development of GvHD (e.g., through T-cell depletion of the graft)—should be preferred.

In *autologous* BMT for the treatment of leukemia the cause of relapse may be either the regrowth of leukemic cells that have survived the CT—in which case the treatment of minimal residual disease (MRD) should be improved—or the outgrowth of leukemic cells injected with the autograft. Then, the graft should be pretreated in some way or other. Of course, both sources may contribute (the possibility that a new leukemia develops in a patient is ignored).

Once again through mathematical modeling an estimate can be obtained on the likelihood of a leukemia relapse originating from MRD or from the graft. It will be shown that the contribution of the latter can be assumed to be much smaller than that of the former.

## 6.2 On the Quantitative Role of Graft-versus-host Disease in Decreasing the Leukemia Relapse Rate after Allogeneic Bone Marrow Transplantation<sup>1</sup>

Clinical data accumulated during the past decade show that patients undergoing allogeneic bone marrow transplantation (BMT) for leukemia have lower leukemia relapse rates if they develop graft-versus-host disease (GvHD) [Dicke et al., 1985; Bacigalupo et al., 1985; Masaoka, 1988; Weiden et al., 1979 and 1981; Butturini et al., 1987; Goldman et al., 1988]. Vice versa: measures to reduce the occurrence of GvHD such as T-cell depletion of the grafted allogeneic marrow increase the frequency of leukemia relapses after allogeneic BMT [Butturini et al., 1987; Goldman et al., 1988; Apperley et al., 1986]. This is highly significant in case of chronic myelocytic leukemia (CML) with an increase in relapse rate from 10% to 30-40% [Butturini et al., 1987; Goldman et al., 1988]. It is borderline significant in acute lymphocytic leukemia (ALL; from 35% to 55% [Butturini et al., 1987]). Additional evidence for a graft-versus-leukemia reaction (GvLR) is derived from the observations in patients with acute myelocytic leukemia (AML) that the leukemia relapse rate after allogeneic BMT in first remission is lower than after autologous or isologous BMT, i.e., 20-30% versus 50-60%, respectively [Butturini et al., 1987; Horowitz, 1988; Löwenberg et al., 1987; Gale and Champlin, 1984].

Experimental data suggest that the graft-versus-host reaction and the GvLR can be separated, although the evidence is not convincing [Butturini et al., 1987; Meredith and O'Kunewick, 1983; Denham et al. 1983]. While GvHD has a beneficial effect in terms of leukemia cure probability it also causes serious morbidity and mortality [Dicke et al., 1985; Goldman et al., 1988; Santos et al., 1988], thereby reducing the net success rate of allogeneic BMT. In this context it appears to be useful 1) to quantify what extra decrease in survival of clonogenic leukemic cells can be attained with GvHD and 2) to investigate whether the same amount of cell kill also might be achieved by other means that can be combined with GvHD prevention and thus are not causing extra GvHD-related deaths.

The first question is looked into by comparing theoretical and clinically observed leukemia cure probabilities in the presence or absence of GvHD. The extra "log cell kill" (LCK) value corresponding with the observed decreased leukemic cell survival will be derived in a mathematical analysis. Subsequently,

---

<sup>1</sup>Chapter 6.2 has been presented at the *Third International Symposium on Minimal Residual Disease in Acute Leukemia, 28-30 March, 1990, Rotterdam, The Netherlands* as: Schultz FW, Vriesendorp HM and Hagenbeek A. Allogeneic bone marrow transplantation, graft-vs-host disease and leukemia relapse



with respect to the second question, the feasibility of achieving an equally increased probability of cure with further pre- or post-BMT treatment will be reviewed.

## 6.2.1 METHODS

### 6.2.1.1 Treatment Models

First, the leukemia patient is brought into remission by remission induction (chemo/radio)therapy (RIT). RIT being successful, next, a much more severe conditioning therapy (CT) is applied so as to kill as many residual leukemic cells as possible. CT fatally damages the patient's normal hemopoietic system, so (allogeneic) BMT is applied to restore it.

To decrease the chances of leukemia relapse three different procedures may be followed: 1) No measures are taken to prevent GvHD; the associated GvLR is expected to work as an extra agent against residual leukemic cells; 2) CT is intensified and GvHD prophylaxis is given as well; 3) Standard CT is given, as well as GvHD prophylaxis and extra post-transplant antileukemia therapy.

### 6.2.1.2 Theoretical Probability of Leukemia Cure

For each treatment model the derivation of the theoretical probability of cure (TPC) is based on the old dogma that only one clonogenic leukemic cell surviving treatment is necessary to cause a relapse of the malignancy. Thus, the TPC is equal to the probability of all clonogenic leukemic cells being eradicated. This probability can be calculated from statistical distributions of leukemic cell numbers surviving treatment, assuming that each cell has the same chance,  $p$ , to survive and, consequently, chance  $1-p$  to be killed by the treatment. Let the tumor burden of a leukemia patient in remission be  $M$  clonogenic leukemic cells and let  $X$  denote the number of surviving clonogenic leukemic cells after conditioning treatment (CT).

Thus, if  $X \geq 1$  there will be no cure; the TPC equals the probability of  $X$  being zero:

$$Pr\{cure\} = Pr\{X=0\}. \quad (6.1)$$

If  $M$  is large ( $> 100$ ) and  $p$  is small ( $< 0.1$ ), then the conditions for the application of a Poisson distribution are satisfied, i.e., the probability that  $K$  cells

survive ( $0 \leq K \leq M$ ) is given by:

$$Pr\{X=K\} = \frac{(p \cdot M)^K}{K!} \cdot e^{-p \cdot M}, \quad (6.2)$$

where  $K!$  ( $K$  factorial) by definition means  $1 \cdot 2 \cdot 3 \cdot \dots \cdot K$  ( $0! = 1$ ). The expected value of  $X$  is  $\mu_X = p \cdot M$  with a standard deviation  $\sigma_X = \sqrt{p \cdot M}$ .

Thus, by substituting zero for  $K$ , the TPC can be written as:

$$Pr\{cure\} = Pr\{X=0\} = e^{-p \cdot M}. \quad (6.3)$$

If  $M$  is small and/or  $p > 0.1$  then instead of the Poisson distribution the binomial probability distribution function applies to obtain the probability of  $K$  surviving cells:

$$Pr\{X=K\} = p^K \cdot (1-p)^{M-K} \cdot \frac{M!}{K! \cdot (M-K)!}. \quad (6.4)$$

The expected value of  $X$  now is  $\mu_X = p \cdot M$  with standard deviation  $\sigma_X = \sqrt{(1-p) \cdot p \cdot M}$ . The TPC then becomes:

$$Pr\{cure\} = Pr\{X=0\} = (1-p)^M. \quad (6.5)$$

To derive the TPC for the treatment models considered above the two probability distributions must be combined. Before conditioning the tumor load will be high ( $M$  large). The cytoreductive potential of the CT also is large, i.e., each cell has only a small chance ( $p_{CT}$ ) to survive this treatment so relatively few leukemia cells will remain. Letting  $X'$  denote the number of such residual leukemic cells, the expected distribution of  $X'$  therefore can be calculated from Eq.(6.2) (Poisson). If, subsequently, an additional (small) cytoreductive effect is exerted on the small number of residual cells—either by a GvLR or by a post-transplant therapy—each of the residual cells has a (comparatively large) chance ( $p_S$ ) to survive once more. Thus, the resulting probability of cure is composed of 1) the chance,  $Pr\{X'=K\}$ , that  $K$  out of  $M$  cells survive CT (Poisson); and 2) the conditional chance,  $Pr\{X=0|X'=K\}$ , that zero out of  $K$  cells survive the second effect (binomial), given the chance that (the small number of)  $K$  cells have survived the CT. In formula:

$$Pr\{cure\} = \sum_K Pr\{X'=K\} \cdot Pr\{X=0 \mid X'=K\} = \sum_K \frac{(p_{CT} \cdot M)^K}{K!} \cdot \exp\{-P_{CT} \cdot M\} \cdot (1 - p_S)^K, \quad K = 0, 1, 2, \dots \quad (6.6)$$

For practical reasons it is assumed that the conditional probability of cure will be zero when the residual population is larger than a certain number of cells, L:

$$Pr\{X=0 \mid X' > L\} = 0. \quad (6.7)$$

Thus, the summation in Eq.(6.6) need not extend beyond  $K = L$ . If L is set low this assumption tends to be rather pessimistic.

### 6.2.1.3 Choice of Parameter Values

A survey of clinical data suggests that when a patient is diagnosed for leukemia the tumor load will often be some  $10^{12}$  cells. A first course of successful remission-induction therapy (RIT) will generally reduce the leukemic cell burden by at least two logs, making it disappear below the clinical detection level (e.g., 80% of AML patients younger than 50 y obtains complete remission [Smith et al., 1986]). Thus, fewer than  $10^{10}$  leukemic cells remain.

Patient data on the malignant cell population size as function of time and treatment are virtually non-existent. Estimates can be obtained in a systematic way from in vivo laboratory experiments only. In this respect the availability of a good rat model—the BNML model—proved very valuable. In this model various treatment protocols have been evaluated, including RIT, CT and (allogeneic) BMT as well as post-transplant chemotherapy, using total body irradiation and/or chemotherapy with cytostatic agents of various classes. A comprehensive overview of experiments and results was recently presented by Martens et al. [1990a,b].

From the BNML experiments, mimicking clinical treatment protocols, it is deduced that two courses of RIT probably will reduce the leukemic cell load by about two times two logs, leaving on average  $10^8$  out of  $10^{12}$  cells. Intensive CT will cause some 8 LCK [Hagenbeek and Martens, 1981, 1983, 1987b], so only a few residual leukemic cells may survive.

Based on these findings it was decided that the base point values of interest would be  $M = 10^8$  and  $SF = (SF_1 =) 10^{-8}$ . SF is the surviving fraction, equivalent to each cell's chance to survive.

However, as probably not all patients will carry the same tumor burden after successful RIT [Rohatiner et al., 1988], the influence of distributed M on the probability of cure must be evaluated. A few approximating simulations

TABLE 6-1 THEORETICAL PROBABILITY OF CURE ACCORDING TO POISSON DISTRIBUTION, AS FUNCTION OF LEUKEMIC CELL LOAD AND "LOG CELL KILL" EFFECT

| surviving<br>fraction,<br>SF = | leukemia burden before conditioning, M = |                     |                     |                     |                     |                     |                     |
|--------------------------------|--|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
|                                | $10^{12}$                                | $10^{10}$           | $10^9$              | $10^8$              | $10^7$              | $10^6$              | $10^4$              |
| $10^{-1}$                      | 0  | 0                   | 0                   | 0                   | 0                   | 0                   | 0                   |
| $10^{-2}$                      | 0  | 0                   | 0                   | 0                   | 0                   | 0                   | $4 \times 10^{-44}$ |
| $10^{-3}$                      | 0  | 0                   | 0                   | 0                   | 0                   | 0                   | $5 \times 10^{-5}$  |
| $10^{-4}$                      | 0  | 0                   | 0                   | 0                   | 0                   | $4 \times 10^{-44}$ | 0.37                |
| $10^{-5}$                      | 0  | 0                   | 0                   | 0                   | $4 \times 10^{-44}$ | $5 \times 10^{-5}$  | 0.90                |
| $10^{-6}$                      | 0  | 0                   | 0                   | $4 \times 10^{-44}$ | $5 \times 10^{-5}$  | 0.37                | 0.99                |
| $10^{-7}$                      | 0  | 0                   | $4 \times 10^{-44}$ | $5 \times 10^{-5}$  | 0.37                | 0.90                | 1                   |
| $10^{-8}$                      | 0  | $4 \times 10^{-44}$ | $5 \times 10^{-5}$  | 0.37                | 0.90                | 0.99                | 1                   |
| $10^{-9}$                      | 0  | $5 \times 10^{-5}$  | 0.37                | 0.90                | 0.99                | 1                   | 1                   |
| $10^{-10}$                     | $4 \times 10^{-44}$                      | 0.37                | 0.90                | 0.99                | 1                   | 1                   | 1                   |
| $10^{-11}$                     | $5 \times 10^{-5}$                       | 0.90                | 0.99                | 1                   | 1                   | 1                   | 1                   |

were performed. By assigning fractions,  $f_i$ , of patients to tumor burdens,  $M = 10^i$ —the largest fraction at  $M = 10^8$ , the other fractions symmetrically decreasing around it—the total increase in probability of cure can be calculated for various assumed values of increase in LCK. For each fraction—the sum of fractions  $f_i$  must obviously be one—the increase in probability of cure due to increase in LCK (or decrease in SF) can be found using Eq.(6.3). Their sum should, of course, match the observed value. This procedure then yields an estimate of the actual increase in LCK.

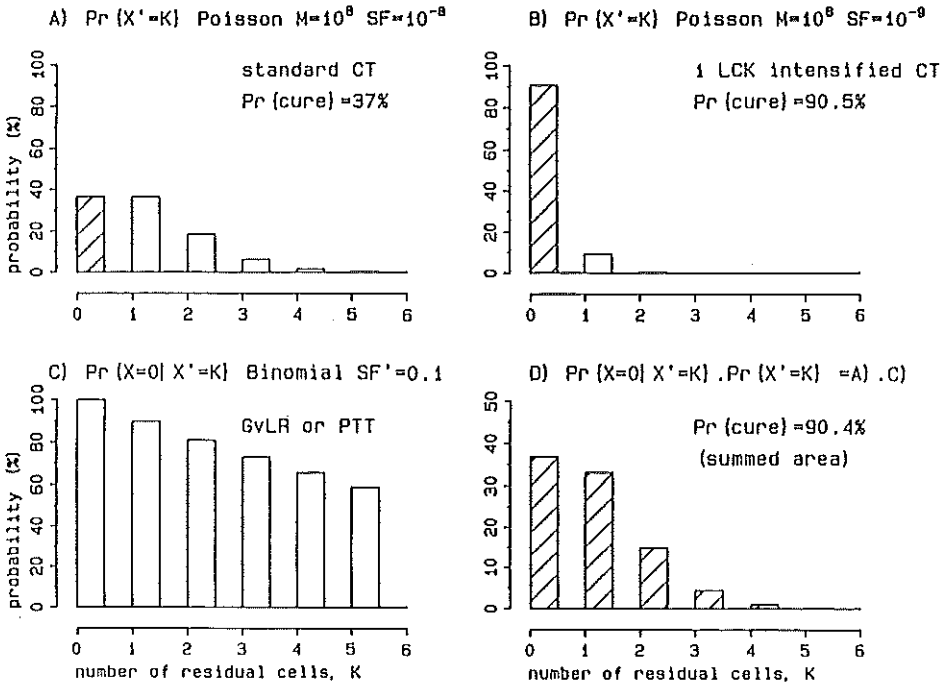
### 6.2.2 RESULTS

Several initial tumor loads,  $M$ , and surviving fractions,  $p$ , as listed in Table 6-1, were chosen such that the TPC could be computed according to the Poisson distribution, Eq.(6.3). It decreases with either increasing original leukemic cell load or increasing surviving fraction. Although not tabulated here, the binomial distribution too (Eq.(6.5)) yields a decrease in TPC with increasing  $M$  and/or SF.

The largest difference between leukemia relapse rates for non-GvHD and

Fig. 6.1 Theoretical Probability of Cure for Several Therapy Procedures

A,B) probability that K cells survive conditioning if therapy is started at  $M = 10^8$  cells (probability of cure,  $K = 0$ , is 40% and 90% for  $SF = 10^{-8}$  and  $10^{-9}$ , respectively); C) probability that the K residual cells are all killed, by GvLR or post-transplantation therapy (PTT) causing 1 LCK; D) probability of cure ( $= 90\%$ ) then is obtained by multiplying A and C for each K and adding the results



(both acute and chronic) GvHD developing patients was found from clinical data concerning acute lymphoblastic or non-lymphoblastic leukemia. The probability of being in remission at day 150 after transplantation was reported to be about 40% and 90%, respectively [Dicke et al., 1985].

In Table 6-1 it is shown that a rise in percent cure from about 40% to 90%—or a decrease in leukemia relapse rate from 60% to 10%—corresponds with a difference of 1 LCK for several combinations of M and two subsequent p values. In particular this is the case when assuming—as suggested—that the leukemic cell population just after an average quality remission-induction therapy has a size of  $M = 10^8$  cells and that the LCK induced by the conditioning regimen has the value eight or nine, respectively.

Figure 6.1 shows again that the probability of cure ( $K = 0$ ) will shift from 40% to 90% when the CT intensity is increased by 1 LCK (Fig. 6.1A;

TABLE 6-2 INCREASE IN THEORETICAL PROBABILITY OF CURE AS FUNCTION OF LEUKEMIC CELL LOAD DISTRIBUTION, IF SF DECREASES BY EITHER 1 OR 2 LOGS

Before conditioning therapy an (assumed) fraction  $f_i$  of the patients carries a leukemic cell load of  $M = 10^i$  cells. Increase in cure probability ( $\Delta\text{Pr}\{\text{cure}\}$ ) for each  $i$  from Table 6-1. Total increase in cure probability should be compared to the observed 50% due to the presence of GvHD

|  | for 1 LCK (SF: $10^{-8} \rightarrow 10^{-9}$ ) |        |                          |        |        | for 2 LCK (SF: $10^{-8} \rightarrow 10^{-10}$ ) |        |                          |        |        |
|--|--|--------|--------------------------|--------|--------|---|--------|--------------------------|--------|--------|
| $i$  | 10   | 9      | <u>8</u>                 | 7      | 6      | 10  | 9      | <u>8</u>                 | 7      | 6      |
| $M$  | $10^{10}$                                      | $10^9$ | <u><math>10^8</math></u> | $10^7$ | $10^6$ | $10^{10}$                                       | $10^9$ | <u><math>10^8</math></u> | $10^7$ | $10^6$ |
| $\Delta\text{Pr}\{\text{cure}\}$           | 0  | 37     | <u>53</u>                | 9      | 1      | 37  | 90     | <u>62</u>                | 10     | 10     |
| <b>A</b>                                   |  |        |                          |        |        |   |        |                          |        |        |
| $f_i$                                      | 0.0  | 0.0    | <u>1.0</u>               | 0.0    | 0.0    | 0.0   | 0.0    | <u>1.0</u>               | 0.0    | 0.0    |
| $f_i \cdot \Delta\text{Pr}\{\text{cure}\}$ | 0.0  | 0.0    | <u>53.0</u>              | 0.0    | 0.0    | 0.0   | 0.0    | <u>62.0</u>              | 0.0    | 0.0    |
|  | sum = 53.0 ( <i>all right</i> )                |        |                          |        |        | sum = 62.0 ( <i>too high</i> )                  |        |                          |        |        |
| <b>B</b>                                   |  |        |                          |        |        |   |        |                          |        |        |
| $f_i$                                      | 0.01   | 0.04   | <u>0.90</u>              | 0.04   | 0.01   | 0.01  | 0.04   | <u>0.90</u>              | 0.04   | 0.01   |
| $f_i \cdot \Delta\text{Pr}\{\text{cure}\}$ | 0.01   | 1.48   | <u>47.7</u>              | 0.36   | 0.01   | 0.37  | 3.6    | <u>55.8</u>              | 0.4    | 0.01   |
|  | sum = 49.55 ( <i>all right</i> )               |        |                          |        |        | sum = 60.18 ( <i>too high</i> )                 |        |                          |        |        |
| <b>C</b>                                   |  |        |                          |        |        |   |        |                          |        |        |
| $f_i$                                      | 0.01   | 0.09   | <u>0.80</u>              | 0.09   | 0.01   | 0.01  | 0.09   | <u>0.80</u>              | 0.09   | 0.01   |
| $f_i \cdot \Delta\text{Pr}\{\text{cure}\}$ | 0.01   | 3.33   | <u>42.4</u>              | 0.81   | 0.0    | 0.37  | 8.1    | <u>49.6</u>              | 0.9    | 0.01   |
|  | sum = 46.55 ( <i>all right</i> )               |        |                          |        |        | sum = 58.98 ( <i>too high</i> )                 |        |                          |        |        |
| <b>D</b>                                   |  |        |                          |        |        |   |        |                          |        |        |
| $f_i$                                      | 0.03   | 0.12   | <u>0.70</u>              | 0.12   | 0.03   | 0.03  | 0.12   | <u>0.70</u>              | 0.12   | 0.03   |
| $f_i \cdot \Delta\text{Pr}\{\text{cure}\}$ | 0.03   | 4.34   | <u>37.1</u>              | 1.08   | 0.0    | 1.11  | 10.8   | <u>43.4</u>              | 1.2    | 0.03   |
|  | sum = 42.55 ( <i>too low</i> )                 |        |                          |        |        | sum = 56.54 ( <i>too high</i> )                 |        |                          |        |        |
| <b>E</b>                                   |  |        |                          |        |        |   |        |                          |        |        |
| $f_i$                                      | 0.05   | 0.24   | <u>0.51</u>              | 0.24   | 0.05   | 0.05  | 0.24   | <u>0.51</u>              | 0.24   | 0.05   |
| $f_i \cdot \Delta\text{Pr}\{\text{cure}\}$ | 0.05   | 8.68   | <u>25.5</u>              | 2.16   | 0.0    | 1.65  | 21.6   | <u>31.6</u>              | 2.4    | 0.05   |
|  | sum = 36.42 ( <i>too low</i> )                 |        |                          |        |        | sum = 57.32 ( <i>too high</i> )                 |        |                          |        |        |

$M = 10^8$ ;  $SF = 10^{-8}$  or  $10^{-9}$ , respectively). Furthermore, it shows that in this case the same improvement of cure probability can be obtained by maintaining the 8 LCK conditioning treatment and adding the extra 1 LCK afterwards (Fig. 6.1A-C).

Table 6-2 shows roughly that broadening the distribution of the leukemic cell population size after RIT has little influence on the expected cell kill effect as established above. The observed improvement of the probability of cure of 50% still is explained best by a cell kill effect of not much more than 1 LCK. Note that the same figures in the table are valid for other ranges of  $M$ , as long as the corresponding values of  $p$  (Table 6-1) are taken account of as well.

At present GvHD is held responsible for the improved cure rate; the conclusion then must be that GvHD causes about 1 LCK. If, instead of through GvHD the same 1 LCK is induced through additional post-transplantation therapy, the same improvement can be expected.

### 6.2.3 DISCUSSION

The relapse rate of leukemia after allogeneic BMT is strongly correlated with either the occurrence of GvHD or T-cell depletion of the graft. E.g., Goldman et al. [1988] have reported for chronic myelocytic leukemia that the probability of relapse within 4 yr after BMT in the absence of GvHD is a factor of 2 higher than in its presence. In the same study the chance of relapse in the case of T-cell depleted grafts appeared to be a factor of 2-3 higher than for non T-cell depleted cases, independent on GvHD development. The same tendency although not quite that significant is seen in both acute myelocytic and acute lymphocytic leukemia [Horowitz, 1988].

In the present theoretical study simple statistical models have been applied in an effort to analyze and explain the clinical observations on differences in leukemia relapse rates. Little can be said about the actual processes of cell kill as function of time in both categories. At the most a more elaborate simulation study as suggested by Birkhead [1985], using disease free survival time observations as input for a mathematical model of remission duration, might shed some light on the most likely depth of a remission, the cure probability and the rate of leukemia regrowth. However, it is clear from the present simple calculations, which are based on the average patient and on preclinical data on therapy efficacy, that GvHD contributes no more than the equivalence of about 1 LCK to the eradication of the malignant cells. Yet, this extra 1 LCK on top of the 8 LCK by conditioning treatment does explain the relatively large increase in probability of cure, i.e., from 40 to 90 percent.

Letting GvHD develop thus may result in more leukemia cures. The disadvantage, unfortunately, is that GvHD itself evokes morbidity and mortality. Consequently, one cause of death might in this situation be replaced by another.

Several bone marrow transplantation centers have advocated to cancel T-cell depletion, in particular in chronic myelocytic leukemia, to make use of the GvLR to prevent high leukemia relapse rates. This causes a strong feeling that another treatment strategy should be developed, maintaining the significant benefits brought about by introducing T-cell depleted allogeneic marrow grafting in terms of preventing GvHD related morbidity and mortality. A strategy must be found that yields the same improvement in leukemia cure probability, yet will not tolerate GvHD development and therefore will show a better overall therapeutic ratio. As is shown above, such additional treatment would need to induce no more than a little over one extra log leukemic cell kill to raise the probability of cure to over twice its original value. Apart from turning to strategies that are still poorly explored, like monoclonal antibody therapy [Ferrara et al., 1989] or employment of lymphokine activated killer cells [Prentice and Brenner, 1989], this might be achieved in different ways by application of pre- or post-transplant additional chemotherapy.

As shown in Fig. 6.1A the TPC can be raised by intensifying the antileukemia conditioning treatment. Achievement of an additional 1 LCK in conditioning therapy will allow later suppression of GvHD and makes up for losing the benefit of the GvLR. This is presently being explored in the Dr Daniel den Hoed Cancer Center in Rotterdam by increasing the dose of total body irradiation from 2 x 5.0 Gy to 2 x 6.0 Gy together with high-dose chemotherapy before T-cell depleted allogeneic BMT. The addition of 2 x 1.0 Gy to the conditioning regimen is thought to induce an extra 1-2 log leukemic cell kill, while the risk of non-leukemic deaths remains at an acceptable level.

Another approach would be to apply additional low-dose antileukemia treatment shortly after bone marrow transplantation, for instance with alkylating agents in combination with GvHD suppressing drugs. Of course there is the danger of interfering with the recovering hemopoietic system. However, in a rat model for human acute myelocytic leukemia (BNML) it has been shown that the method is feasible, without jeopardizing the graft, employing low dose cyclophosphamide [Hagenbeek and Martens, 1981], or biological response modifiers [Hagenbeek and Martens, 1983 and 1987b]. In this way a significant increase in cure rates could be achieved. A similar conclusion could be drawn from a clinical study [Horowitz et al., 1989] concerning the effect of methotrexate (MTX) as a method of GvHD prophylaxis.

In summary, three options have been discussed for better leukemia control after allogeneic BMT: 1) tolerate the development of GvHD; a profit is gained in the form of an equivalent of a 1 LCK, presumably by a GvLR acting on residual cells after BMT; 2) suppress the development of GvHD, e.g., by T-cell depletion, and intensify the CT such as to achieve an extra 1 LCK; 3) suppress the development of GvHD and give supplementary low dose (1 LCK) antileukemia



treatment as soon as possible after BMT.

All three options will predictively result in the same raised level of leukemia cure probability. Option 1 has the disadvantage that serious GvHD-related life-threatening complications must be faced. Furthermore, it implies that GvHD should be encouraged in those patients who do not develop the disease spontaneously. Option 2 implies a further intensification of an already heavy therapy that, however, may be feasible within narrow limits. Option 3 opens another pathway to eradicate the one or two logs of residual leukemic cells and deserves to be tested in clinical protocols for patients with a high risk of leukemic recurrence.

### 6.3 Contributions to Leukemia Relapse of Residual Leukemic Cells in the Patient and in the Graft after Autologous Bone Marrow Transplantation<sup>1</sup>

High-dose chemo/radiotherapy followed by autologous bone marrow transplantation (ABMT) has become a regular and accepted form of treatment of leukemia [Bast and Ritz, 1984; Burnett et al., 1984; Hervé et al., 1983; Kaizer et al., 1983]. Leukemia often is diagnosed when the patient is already in an advanced stage of the disease. The bone marrow then has been fully infiltrated and the patient carries on average a burden of some  $10^{12}$  leukemic cells in this organ. Usually chemotherapy will induce a clinical complete remission [Smith et al., 1986], i.e., less than 5 per cent leukemic cells survive in the bone marrow (the clinical detection level). If the patient is selected for ABMT then some  $1.5 \times 10^{10}$  bone marrow cells will be harvested at this time point (on average,  $2 \times 10^8$  cells/kg and a body weight of 75 kg), because A) the leukemic cell frequency is low, and B) the exposure of the normal hemopoietic stem cells to cytostatic agents has still been relatively mild. The bone marrow cells are frozen according to standard procedures and properly stored in liquid nitrogen for future use. In general, the hemopoietic stem cells thus will retain their blood cell forming capacity. Subsequently, to destroy all remaining leukemic cells a high intensity chemo/radiotherapy regimen (conditioning treatment) is applied. To restore normal hemopoiesis the bone marrow graft is thawed and reinfused immediately afterwards.

Clinical results show that 50-60 percent of the ABMT patients, transplanted in first complete remission of acute myelocytic leukemia (AML), will relapse [Löwenberg et al., 1987; Burnett et al., 1987; Gorin et al., 1987]. This occurs at a median time of 240 days after transplantation. At that moment the leukemic cell population again has reached a size of about  $5 \times 10^{10}$  ( $> 5\%$  blast cells in the marrow).

Beside a remote possibility of a new leukemia having emerged, two different sources may be responsible for a relapse, A) minimal residual disease (MRD), i.e., any leukemic cell(s) in the patient having survived the intensive conditioning treatment; and B) any leukemic cells that have survived the remission-induction treatment, were taken out and, inevitably, were reinfused again with the graft.

An indication that the contribution of "source B", the graft, is much smaller than that of "source A", MRD, is found in the fact that patients who receive the

---

<sup>1</sup>Chapter 6.3 was published with minor modifications as: Schultz FW, Martens ACM and Hagenbeek A (1989) The contribution of residual leukemic cells in the graft to leukemia relapse after autologous bone marrow transplantation: mathematical considerations. *Leukemia* 3:530-534

same treatment, but a "clean" graft originating from an identical twin brother or sister (isologous bone marrow transplantation, IBMT), show a relapse rate similar to that after ABMT (about 60 percent; [Gale and Champlin, 1984]). Below, another argument in favor of the relatively minor contribution of the graft to the relapse of leukemia is presented. It is based on a hypothetical model for the development of leukemia from engrafted malignant cells. Its validity will be elucidated by the use of data derived from laboratory experiments with the Brown Norway (BN) rat acute myelocytic leukemia (BNML) that represents one of the best models for human AML.

### **6.3.1 METHODS**

#### **6.3.1.1 Experimental Animals (See Chapter 1.9.1)**

The experiments were performed in the barrier derived inbred Brown Norway rat strain BNBi/Rij from the Rijswijk colony. Male rats were used of ages between 13 and 16 weeks (220 g mean body weight).

#### **6.3.1.2 The Brown Norway Rat Acute Myelocytic Leukemia (BNML) (See Chapter 1.9.2)**

The BNML was chemically induced in a female BN rat and appeared to be transplantable within the BN rat strain, yielding reproducible growth patterns. The major characteristics of the disease, i.e., a slow growth rate and severe suppression of normal hemopoiesis, show striking similarities with human AML, making it a suitable model to perform therapy studies with [Martens et al., 1990a,b; Hagenbeek and Martens, 1991].

The principle target organs to be occupied by transplanted BNML cells are the bone marrow and the liver, and to a lesser extent the spleen. This has been established by an experiment [Hagenbeek and Martens, 1979] in which chromium labeled BNML cells were inoculated. Thus the homing and lodging could be monitored quantitatively by measuring the amounts of the radioactive isotope in the various organs. Quantities in lungs, blood and other organs are smaller by at least a factor of ten. It made no difference whether or not the test animals had been subjected to marrow ablative chemotherapy [Hagenbeek and Martens, 1985] before.

#### **6.3.1.3 BNML Growth *in vivo* after Cellular Transfer (See Chapter 2.2 and [Schultz et al., 1987])**

The development of the leukemic cell population after *i.v.* inoculation of  $10^7$  BNML cells has been investigated in the major target organs (bone marrow,

liver, spleen) of the recipient rats. The number of BNML cells present at various time points could be experimentally determined in several ways, ranging from counting cells microscopically on a morphological basis or by flow cytometry after labeling leukemic cells with monoclonal antibodies and fluorescent dyes to more complex bio-assay methods, such as survival experiments and the spleen colony assay for clonogenic leukemic cells.

Mathematical analysis of the datapoints consisted of computerized nonlinear least squares curve fitting. Classes of hypothetical growth curves describing the size of the leukemic cell population as function of time, e.g., an exponential size—time relation or a Gompertz function, were postulated. For each organ the best fitting growth curve was determined, i.e., the class of the growth curve and the parameter values that characterize the specific growth curve in this class.

#### 6.3.1.4 *In vivo* Growth of a Drug-Treated BNML Cell Line (See Chapter 2.4.2)

As part of a larger experiment whose purpose was the *in vivo* development of a drug resistant BNML cell subpopulation, the following test was conducted to investigate whether the growth kinetic properties of drug-treated BNML cells differed from the parent line's. Rats were inoculated with  $10^7$  untreated BNML cells. Subsequently, every 13 or 14 days they were treated with a single *i.p.* dose (100 mg/kg) of cyclophosphamide (CFA). After having been thus *in vivo* exposed seven times to the drug the BNML cells were harvested and, in quantities of  $10^7$ , were transplanted into fresh recipient rats. With most rats the same CFA-treatment procedure was then repeated. Some ( $n = 6$ ) were not further treated but time until death from leukemia was observed. For comparison, assuming that finding a difference in survival times means that there must be a difference in the respective growth patterns, survival times of rats inoculated with  $10^7$  BNML cells of the parent line were observed as well. The significance of differences in mean survival time (MST) was tested with Student's T-test at a 95% confidence level.

#### 6.3.1.5 Determination of the ED<sub>50</sub> value for BNML (See Chapter 2.1.1.2)

When rodents are inoculated with a certain number of leukemic cells, then the chance that leukemia will develop is dependent on this dose. This observation is based on data from laboratory experiments [Hagenbeek and Martens, 1985 and 1987a; Hewitt, 1958] in which percent mortality was associated with the mean cell dose injected in animals. Below a certain dose no leukemia will arise, above another critical dose all animals will develop the disease. In between, the chances of leukemia development increase with increasing cell dose. The ED<sub>50</sub> value is defined as the number of malignant cells that is required to induce leu-

TABLE 6-3 THE PROBABILITY OF LEUKEMIA DEVELOPMENT BASED ON INOCULATION EXPERIMENTS AGREES WELL WITH THE THEORETICAL PROBABILITY BASED ON THE ED<sub>50</sub> CONCEPT (compare columns 4 and 6)

| nr of BNML cells injected | observed nr of leukemias / nr of recipient rats |     |                      | nr of ED <sub>50</sub> units injected | theoretical probability of leukemia development % |
|---------------------------|---|-----|----------------------|---------------------------------------|---|
|                           | ratio   | %   | % after probit anal. |                                       |   |
| 1                         | 0/6   | 0   | 0                    | 0.04                                  | 3   |
| 5                         | 1/6   | 17  | 10                   | 0.2                                   | 13  |
| 25                        | 1/6   | 17  | 50                   | 1.0                                   | 50  |
| 50                        | 5/6   | 83  | 76                   | 2.0                                   | 75  |
| 100                       | 6/6   | 100 | 92                   | 4.0                                   | 94  |
| 1000                      | 6/6   | 100 | 100                  | 40.0                                  | 100   |

The theoretical probability of leukemia development, i.e., the chance that at least one unit of ED<sub>50</sub> cells yields a cell to grow out, is given by  $(1-0.5^X) \cdot 100\%$ , where X denotes the number of ED<sub>50</sub> units injected (ED<sub>50</sub> for BNML: 24.7 cells).

kemia in fifty percent of the individuals of an inoculated group.

Six groups of BNML rats (n = 6 rats per group) were inoculated *i.v.* with different (low) numbers of BNML cells (1-1000). The occurrence of leukemia in each group was registered, with death from leukemia as the endpoint in an observation period of 500 days totally. A probit analysis was performed to determine the ED<sub>50</sub> value from the thus obtained datapoints of percentage-leukemic-animals as a function of cell dose.

### 6.3.2 RESULTS

The left part of Table 6-3 shows the observations of the experiment conducted to determine the ED<sub>50</sub> value for the BNML. The result of the probit analysis is shown in Fig. 2.1. The derived ED<sub>50</sub> value, the dose corresponding with 50 percent of the rats developing leukemia, appears to be 24.7 cells [Hagenbeek and Martens, 1985 and 1987a].

The results of the experiments conducted to analyze the growth of the BNML cell populations in the rat's organs after transplantation of the leukemia (10<sup>7</sup> BNML cells *i.v.*) are illustrated in Fig. 2.15. Figure 2.15A shows the bone marrow datapoints, as well as the growth curve found by nonlinear least squares fitting. This curve consists of an exponential part (constant population doubling time of 0.8 day) that is contiguously followed by a Gompertz curve (from the conversion point at 2x10<sup>8</sup> cells the doubling time increases exponentially with rate 0.4 per day) making the population size level off toward a steady state plateau phase. For liver and spleen similar curves were found, with slightly different parameter values. The size—time course of the total population is reflect-

ed by the sum of the three individual curves (Fig. 2.15B). The animals die on day 22 (median value) with a leukemia burden of some  $2 \times 10^{10}$  BNML cells.

The median survival time of rats inoculated with a dose of  $10^7$  cells of the parent BNML cell line was 22 d (range 19-24). The mean survival time and standard deviation were  $22.0 \pm 1.9$  d. When the inoculum consisted of CFA-treated BNML cells ( $7 \times 100$  mg/kg) the corresponding values were 26 d (range 20-26) and  $23.6 \pm 3.3$  d. The T-test revealed that the null-hypothesis of no significant difference in MST could not be rejected at the 95% confidence level.

### 6.3.3 DISCUSSION

#### 6.3.3.1 The $ED_{50}$ Model Concept

The  $ED_{50}$  model concept, postulated here, states that a quantity of inoculated cells can be regarded as a number of  $ED_{50}$  units. Upon inoculation each unit will either disappear completely—i.e., no proliferative activity results—or produce only one cell to grow out. Both possibilities have equal probability. If at least one cell remains to grow out (when all units but one disappear), then leukemia will develop.

Evidence for the validity of the  $ED_{50}$  concept will be given below—using the BNML rat model—by A) comparison of experimentally established data on BNML induction and expected leukemia incidence derived theoretically on the basis of the  $ED_{50}$  concept; and B) analysis of experimental growth data on the development of the leukemic population after the inoculation of  $10^7$  BNML cells *i.v.*

A) Based on the  $ED_{50}$  model concept the theoretical chance can be calculated that leukemia will develop. This chance is a function of the number of injected cells and of the  $ED_{50}$  value.

Let the number of injected cells be  $N$  and the  $ED_{50}$  value be  $M$ . Then, the inoculum consists of  $X = N/M$   $ED_{50}$  units. One  $ED_{50}$  unit will yield either 0 or 1 cell to grow out; the chance that 0 cells remain is 0.5 (50 percent). If two  $ED_{50}$  units are considered, then four situations are possible: both yield 1 cell each, both yield 0 cells each, or either one yields 1 cell while the other yields 0. As the  $ED_{50}$  units are independent, the chance that both will yield 0 cells equals  $0.5 \times 0.5 = 0.25$  (the only case corresponding with no leukemia; if either or both would yield 1 cell to remain, then leukemia does develop). In general, considering  $X$  independent  $ED_{50}$  units, the chance for zero cells to remain (no leukemia) becomes  $0.5^X$ . Complementary, the chance that leukemia will develop can be written as  $(1 - 0.5^X)$ .

Calculating this chance for various values of  $X$  and comparing the results with the outcome of the actual experiment (also used for the  $ED_{50}$  determina-

tion) reveals that good agreement exists (Table 6-3; compare columns 4 and 6).

B) By backwards extrapolation of the overall BNML growth curve to day zero (Fig. 2.15B) it follows that of the  $10^7$  leukemic cells inoculated on day zero only a small fraction ( $1.5 \times 10^5$ ) appears to grow out. In this region of low cell numbers the available detection methods fail in mapping the development of the population, but indeed, the extrapolation agrees well with the number of cells theoretically available for growing out according to the  $ED_{50}$  concept, indicating its validity. The  $ED_{50}$  value for BNML equals 24.7 cells, thus, the  $10^7$  inoculated BNML cells can be considered as  $10^7/24.7 = 4 \times 10^5$   $ED_{50}$  units. As this is a large number and as each unit has a fifty percent chance to disappear completely, on average half the number of  $ED_{50}$  units will be able to produce 1 proliferating cell (compare tossing a coin many times and counting the number of heads). Thus, the expected initial proliferating population size will be  $2 \times 10^5$ , a value quite similar to the  $1.5 \times 10^5$  cells on day zero extrapolated from Fig. 2.15B.

The  $ED_{50}$  concept implies that when leukemic cells are (re)inoculated with a bone marrow graft, a relatively large number of these cells just seem to disappear, i.e., they are not proliferatively active. Only a few—at the most one out of every  $ED_{50}$  cells—may grow out to produce a relapse of leukemia, besides any leukemic cells that might have been left in the patient (MRD). The chances of leukemia development due to the contribution of the graft increase with increasing cell dose or/and decreasing  $ED_{50}$  value.

Although at present not verified by the most proper experiment, i.e., the direct determination of the  $ED_{50}$  value for CFA-treated BNML cells, indirect evidence has revealed that there is no reason to suspect that drug treatment will influence the  $ED_{50}$  value. Equal doses of leukemic cells of either the parent BNML cell line or the CFA-treated cell line both resulted in the same MST of the recipients. Therefore, probably the leukemia growth pattern also will be the same, after starting with the same initial leukemic cell load.

### 6.3.3.2 Application to ABMT

Having demonstrated—or at least having made plausible—the validity of the  $ED_{50}$  concept both for parent and drug-treated cells, its application in ABMT will be given next. The probability of the development of leukemia due to reinfusing leukemic cells with the graft could be calculated exactly, if the number of transplanted malignant cells and their  $ED_{50}$  value were known. Unfortunately, for human patients treated with ABMT this is not so. However, in a simulation study certain likely values can be assumed. In Table 6-4 certain combinations of assumed  $ED_{50}$  value and assumed number of clonogenic leukemia cells in the

TABLE 6-4 THEORETICAL CHANCE (%) OF LEUKEMIA DEVELOPMENT AS FUNCTION OF ASSUMED ED<sub>50</sub> VALUE AND THE NUMBER OF LEUKEMIC CELLS IN THE HUMAN AUTOLOGOUS BONE MARROW GRAFT

| ED <sub>50</sub> value | number of leukemic cells in graft |      |      |       |       |       |       |       |       |  |
|------------------------|-----------------------------------|------|------|-------|-------|-------|-------|-------|-------|--|
|                        | 10                                | 15   | 100  | 150   | 500   | 1000  | 1500  | 5000  | 10000 |  |
| 10                     | 50.0                              | 64.6 | 99.9 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |  |
| 100                    | 6.7                               | 9.9  | 50.0 | 64.6  | 96.9  | 99.9  | 100.0 | 100.0 | 100.0 |  |
| 1000                   | 0.7                               | 1.0  | 6.7  | 9.9   | 29.3  | 50.0  | 64.6  | 96.9  | 99.9  |  |
| 5000                   | 0.1                               | 0.2  | 1.4  | 2.1   | 6.7   | 12.9  | 18.8  | 50.0  | 75.0  |  |
| 10000                  | <.1                               | 0.1  | 0.7  | 1.0   | 3.4   | 6.7   | 9.9   | 29.3  | 50.0  |  |

see note beneath Table 6-3

graft will yield 100 percent, other combinations will yield less than one percent probability of leukemia development (at least one cell grows out) from engrafted cells.

### 6.3.3.3 The Probability of the Graft's Contribution to Leukemia Relapse

The number of leukemic cells present in the human graft can be estimated in the following way. If the total leukemic population in the bone marrow of a human patient at diagnosis has a size of  $10^{12}$  cells and if the effect of the remission-induction therapy is estimated to correspond with some 4 log cell kill (LCK; i.e., the population size is divided by a factor of  $10^4$ ), then, at the moment of taking out the marrow graft, the frequency of occurrence will be 1 leukemic cell in  $10^4$  normal bone marrow cells ( $10^8$  in  $10^{12}$ ). Assuming the mixture to be homogeneous and considering a 75 kg patient with an autologous marrow graft containing  $2 \times 10^8$  marrow cells per kg body weight, the corresponding graft of size  $1.5 \times 10^{10}$  will thus contain  $1.5 \times 10^6$  leukemic cells.

This is a conservative estimate, assuming that one course of remission-induction therapy will just result in a remission (2 LCK, meaning that one percent of the leukemic cells is left). The second course of remission-induction therapy, having the same influence, reduces the population by again a factor of 100. If in reality remission-induction therapy would be more effective, then the bone marrow graft would contain even less than  $1.5 \times 10^6$  leukemic cells.

Cryopreservation of the graft, i.e., freezing and storage in liquid nitrogen and subsequent thawing before reuse, will kill some 99% of the *in vivo* clonogenic leukemic cells [Hagenbeek and Martens, 1989]. Furthermore, of the remainder only about 0.1-1% will still be clonogenic as judged from *in vitro* culture experiments on clonogenicity of human AML cells [Löwenberg et al., 1984; Swart et al., 1982]. Therefore, an estimated number of only 15-150 viable clonogenic leukemic cells will be reinfused with the marrow graft. Should



remission-induction therapy prove more effective than causing 4 LCK, even fewer malignant cells are transplanted.

If  $ED_{50}$  for human AML cells is 1000 and there are between 15 and 150 leukemic cells in the graft, then (Table 6-4) there is a chance of 1-10% that they will cause leukemia. If  $ED_{50}=10000$ , which might be the case if an extrapolation from rat to man were allowed (for BNML:  $ED_{50}=25$ ; rat/man body weight ratio = 0.2/75; therefore,  $ED_{50}$  for AML:  $25 \times 75 / 0.2 = 10000$ ), then this chance is reduced to only 0.1-1% (Table 6-4), i.e., in 99 or more out of 100 cases the graft would not at all contribute to the relapse of leukemia. But even if  $ED_{50}$  would be equal to 100 and there would be 150 clonogenic malignant cells in the graft, then there is still a 35% chance that no leukemia would evolve. The same is true if  $ED_{50}$  were 1000 and the number of clonogenic viable cells in the graft amounts to 1500, thus would be a factor of at least 10 higher than expected.

The residual leukemia ( $10^8$  cells) in the body of a patient in remission may just be reduced to either one or zero cells by the intensive conditioning radio-chemotherapy (8 LCK [Hagenbeek and Martens, 1983 and 1987b]). The latter case would mean the achievement of a cure, the former one would lead to a relapse. Clinical observations with isologous bone marrow transplantation [Gale and Champlin, 1984] reveal a relapse frequency of about 60%. After a median time interval of 240 days the leukemic population then has reached the size of  $5 \times 10^{10}$  cells again, i.e., 5% or more of all marrow cells are recognized as leukemic blasts. Assuming that they originate from a single cell and that growth is exponential, the population doubling time is calculated to be 6.75 days. This agrees well with another study [Schultz et al., 1989], concerning relapsing childhood acute lymphocytic leukemia, in which a population doubling time of 6.52 days was found.

It may be reasonable to expect a similar distribution of leukemic cell burdens in AML patients considered for isologous and autologous bone marrow transplantation, both before and after remission-induction treatment (but before transplantation). Then, to be able to explain the almost equal relapse patterns, it must be assumed that the influence of the only major difference between the groups, i.e., the reinfusion of leukemic cells with the graft, is almost negligible. That this influence can be considered as negligible indeed, can be explained with the  $ED_{50}$  concept, the validity of which has been shown in the BNML rat model. Application to the human situation would mean that the  $ED_{50}$  for human leukemia should be larger than at least 1000 cells. With a maximum amount of 150 clonogenic malignant cells in the graft, the chance that these cells contribute to the regrowth of the disease is less than 10 percent.

In a recent clinical investigation [Brenner et al., 1993] two children with AML developed a relapse after ABMT. Of the blast cells detected in the marrow and peripheral blood a low percentage (2-3%) was shown to originate from the autologous marrow graft. These cells carried the neomycin-resistance gene

(as detected by their survival in *in vitro* culture, in the presence of otherwise toxic neomycin concentrations) that had been introduced as a marker into the normal and malignant bone marrow progenitors before reinfusion of the graft. The thus derived percentage is a lower limit, because the efficiency of gene marking is only modest. Therefore, more than 3% of the leukemic cells may originate from the graft. On the other hand, as the transferred gene is present in equal concentration in all progeny of a marked cell for every generation, it is not possible to determine whether a neomycin gene carrying cell is a descendent of the originally marked cell, or this cell itself. In other words, the proliferative state of this cell, ergo its potential contribution to a relapse, is not certain at this moment.

#### 6.4 REFERENCES

- Apperley JF, Jones L, Hale G, et al. (1986) Bone marrow transplantation for patients with chronic myeloid leukemia: T-cell depletion with campath-1 reduces the incidence of GvHD but may increase the risk of leukemia relapse. *Bone Marrow Transplantation* 1:53-66
- Bacigalupo A, Van Lint MT, Frassoni F, Marmont A (1985) Graft-versus-leukemia effect following allogeneic bone marrow transplantation. *Br J Haematol* 61:749-751 (letter)
- Bast RC, Ritz J (1984) Application of monoclonal antibodies to autologous bone marrow transplantation. In: *Biological Responses in Cancer 2*. Plenum Press, New York and London, pp:185-193
- Birkhead BG (1985) Mathematical model of remission duration in acute myelogenous leukemia. *Cancer Treatm Rep* 69:595-601
- Bortin MM, Rimm AA, Saltzstein EC (1973) Graft-versus-leukemia effect III, Apparent independent antitox and antileukemic activity of transplanted immunocompetent cells. *Transplantation* 16:182-188
- Brenner MK, Rill DR, Moen RC, Krance RA, Mirro J Jr, Anderson WF, Ihle JN (1993) Gene-marking to trace origin of relapse after autologous bone-marrow transplantation. *Lancet* 341:85-86
- Burnett AK, Watkins R, Maharaj D, McKinnon S, Tansey P, Alcorn M, Singer CRJ, McDonald GA, Robertson AG (1984) Transplantation of unpurged autologous bone marrow in acute myeloid leukemia in first remission. *Lancet* ii:1068-1071
- Burnett AK, Mackinnon S, Morrison A (1987) Autologous transplantation of unpurged bone marrow during first remission of acute myeloid leukemia. In: Dicke KA, Spitzer G, Jagannath S, Favrot M, Peters W, eds. *Autologous Bone Marrow Transplantation, Proceedings of the third international symposium*. The University of Texas, MD Anderson Hospital and Tumor Institute at Houston, pp:23-29
- Butturini A, Bortin MM, Gale RP (1987) Graft-versus-leukemia following bone marrow transplantation. *Bone Marrow Transplantation* 2:233-242 (review)
- Denham S, Attridge S, Barfoot RK, Alexander P (1983) Effect of cyclosporin A on the anti-leukaemia action associated with graft-versus-host disease. *Br J Cancer* 47:791-795
- Dicke KA, Zander AR, Vellekoop L, Spitzer G, Verma DS (1985) The treatment of leukemia. In: *Bone marrow transplantation: biological mechanisms and clinical practice*. Van Bekkum DW, Löwenberg B (eds.). Marcel Dekker: New York, pp:435-474

- Ferrara JLM, Wall DA, Van Dijken PJ (1989) Graft versus Host Disease: Mechanisms of immunodeficiency and monoclonal antibody therapy. In: Bone marrow transplantation: current controversies. Gale RP, Champlin RE (eds.). Alan R Liss, Inc: New York, pp:179-193
- Gale RP, Champlin RE (1984) How does bone marrow transplantation cure leukaemia? *Lancet* ii:28-29
- Goldman JM, Gale RP, Horowitz MM, et al. (1988) Bone marrow transplantation for chronic myelogenous leukemia in chronic phase. *Ann Internal Med* 108:806-814
- Gorin NC, Laporte JP, Douay L, Lopez M, Stachowiak J, Aegerter P, Deloux J, Marin JL, Duhamel G, Laugier A, Salmon Ch, Najman A (1987) Use of bone marrow incubated with mafosfamide in adult acute leukemia patients in remission. The experience of the Paris - Saint Antoine transplantation team. In: Dicke KA, Spitzer G, Jagannath S, Favrot M, Peters W, eds. *Autologous Bone Marrow Transplantation, Proceedings of the third international symposium*. The University of Texas, MD Anderson Hospital and Tumor Institute at Houston, pp:15-22
- Hagenbeek A, Martens ACM (1979) Functional cell compartments in a rat model for human acute myelocytic leukemia. *Cell Tissue Kinet* 12:361-370
- Hagenbeek A, Martens ACM (1981) High-dose cyclophosphamide treatment of acute myelocytic leukemia; studies in the BNML rat model. *Eur J Cancer Clin Oncol* 18:763-769
- Hagenbeek A, Martens ACM (1983) The efficacy of high-dose cyclophosphamide in combination with total body irradiation in the treatment of acute myelocytic leukemia; studies in a relevant rat model (BNML). *Cancer Res* 43:408-417.
- Hagenbeek A, Martens ACM (1985) Reinfusion of leukemic cells with the autologous marrow graft: preclinical studies on lodging and regrowth of leukemia. *Leukemia Res* 9:1389-1395
- Hagenbeek A, Martens ACM (1987a) On the fate of leukemic cells infused with the autologous marrow graft. In: Büchner T, Schellong G, Hiddemann W, Urbanitz D, Ritter J, eds. *Acute Leukemias. Prognostic factors and treatment strategies*. Berlin, Heidelberg, New York, London, Paris, Tokyo, Springer Verlag, pp:553-558
- Hagenbeek A, Martens ACM (1987b) Conditioning regimens before bone marrow transplantation in acute myeloid leukemia. In: Dicke KA, Spitzer G, Jagannath S, Favrot M, Peters W (eds.) *Autologous Bone Marrow Transplantation, Proceedings of the third international symposium*. The University of Texas, MD Anderson Hospital and Tumor Institute at Houston, pp:99-103
- Hagenbeek A, Martens ACM (1989) Cryopreservation of autologous marrow grafts in acute leukemia: survival of in vivo clonogenic leukemic cells and normal hemopoietic stem cells. *Leukemia* 3:535-537
- Hagenbeek A, Martens ACM (1991) Minimal residual disease in leukemia: Preclinical studies in a relevant rat model (BNML). In: Jasmin C, Proctor SJ (eds) Ballières "Clinical Haematology" 4 (3) Chapter 3:609-635. Ballière Tindall, WB Saunders Company, London
- Hervé P, Philip T, Flesh M, Rozenbaum A, Plouvier E, Cahn JY, Noir A, Peters A, Leconte des Floris R (1983) Intensive cytoreductive regimen and autologous bone marrow transplantation in leukemia. Present status and the future. A review. *Eur J Cancer Clin Oncol* 19:1043-1051
- Hewitt HB (1958) Studies of the dissemination and quantitative transplantation of a lymphocytic leukemia of CBA mice. *Br J Cancer* 12:378-401
- Horowitz MM (1988; for the Advisory Committee of the International Bone Marrow Transplant Registry) Evidence for a graft versus leukemia effect in clinical bone marrow transplantation. *Exp Hemat* 16:547 (abstract)
- Horowitz MM, Gale RP, Barrett AJ, et al. (1989) Effect of methotrexate on relapse after bone marrow transplantation for acute lymphoblastic leukaemia. *Lancet* i:535-536
- Kaizer H, Tutschka P, Stuart R, Körbling M, Braine H, Saral R, Colvin M, Santos G (1983) Autologous bone marrow transplantation in acute leukemia and non-Hodgkin's lymphoma; a

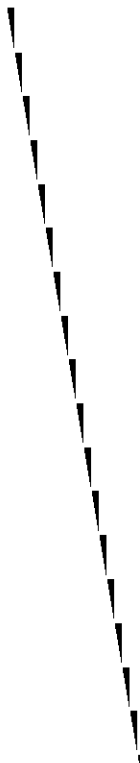
- phase I study of 4-hydroperoxycyclophosphamide (4HC) incubation of marrow prior to cryopreservation. In: *Hematology and Blood Transfusion* 28. Springer, Berlin and Heidelberg, pp:90-97
- Löwenberg B, Abels J, Van Bekkum DW, Dzoljic G, Hagenbeek A, Hendriks WDH, Van de Poel J, Sizoo W, Sintnicolaas K, Wagemaker G (1984) Transplantation of non-purified autologous bone marrow in patients with AML in first remission. *Cancer* 54:2840-2846
- Löwenberg B, Van der Lelie J, Goudsmit R, Willemze R, Zwaan FE, Hagenbeek A, Van Putten WJL, Verdonck LF, De Gast GC (1987) Autologous bone marrow transplantation in patients with acute myeloid leukemia in first remission. In: Dicke KA, Spitzer G, Jagannath S, Favrot M, Peters W, eds. *Autologous Bone Marrow Transplantation, Proceedings of the third international symposium*. The University of Texas, MD Anderson Hospital and Tumor Institute at Houston, pp:3-7
- Martens ACM, Van Bekkum DW, Hagenbeek A (1990a) The BN acute myelocytic leukemia (BNML). A rat model for studying human acute myelocytic leukemia. *Leukemia* 4:241-257
- Martens ACM, Van Bekkum DW, Hagenbeek A (1990b) Minimal residual disease in leukemia. Studies in an animal model for acute myelocytic leukemia (BNML). *Int J Cell Cloning* 8:27-38
- Masaoka T (1988) Allogeneic bone marrow transplantation for leukemia - Japanese experience. In: *Recent advances and future directions in bone marrow transplantation*. Baum SJ, Santos GW, Takaku F (eds.). Springer Verlag: New York, pp:114-118
- Meredith RF, O'Kunewick JP (1983) Possibility of graft-versus-leukemia determinants independent of the major histocompatibility complex in allogeneic marrow transplantation. *Transplantation* 35:378-385
- Prentice HG, Brenner MK (1989) Donor marrow T-cell depletion for prevention of GvHD without loss of the GvL effect. Current results in a myeloblastic leukemia and further directions. In: *Bone marrow transplantation: current controversies*. Gale RP, Champlin RE (eds.). Alan R Liss, Inc: New York, pp:117-128
- Rohatiner AZ, Gregory WM, Bassan R, et al. (1988) Short-term therapy for acute myelogenous leukemia. *J Clin Oncol* 6:218-226
- Santos GW, Yeager AM, Saral R, et al. (1988) Allogeneic and autologous marrow transplants in acute nonlymphocytic leukemia (ANLL). In: *Recent advances and future directions in bone marrow transplantation*. Baum SJ, Santos GW, Takaku F (eds.). Springer Verlag: New York, pp:74-81
- Schultz FW, Martens ACM, Hagenbeek A (1987) Computer simulation of the progression of an acute myelocytic leukemia in the Brown Norway rat. *Comput Math Applic* 14:751-761
- Schultz FW, Van Dongen JJM, Hählen K, Hagenbeek A (1989) Time-history of the malignant population in the peripheral blood of a T-cell acute lymphoblastic leukemia patient: a pilot study. *Comput Math Applic* 18:929-936
- Smith RG, Hetherington ML, Huntsman PR and Buchanan GR (1986) Surveillance of terminal deoxynucleotyl transferase-positive cells in peripheral blood of patients with acute lymphoblastic leukemia. In: Hagenbeek A and Löwenberg B (eds.), *Minimal residual disease in acute leukemia 1986*. Martinus Nijhoff Publishers, Dordrecht, 134-140
- Swart K, Hagemeijer A, Löwenberg B (1982) Acute myeloid leukemia colony growth in vitro: differences of colony forming cells in PHA supplemented and standard leukocyte feeder cultures. *Blood* 59:816-822
- Van Bekkum DW, Löwenberg B (Eds.; 1985) *Bone marrow transplantation: biological mechanisms and clinical practice*. Marcel Dekker, New York
- Visser JWM, Bauman JGJ, Mulder AH, Eliason J, De Leeuw AM (1984) Isolation of murine pluripotent hemopoietic cells. *J Exp Med* 59:1576-1590
- Weiden PL, Flournoy N, Thomas ED, et al. (1979) Antileukemic effect of graft-versus-host

disease in human recipients of allogeneic marrow grafts. N Engl J Med 300:1068-1073

Weiden PL, Sullivan KM, Flournoy N, Storb R, Thomas ED and the Seattle Marrow Transplant Team (1981) Antileukemic effect of graft-versus-host disease; contribution to improved survival after allogeneic marrow transplantation. N Engl J Med 304:1529-1533

100

100



## Chapter 7

# General Discussion, New Developments and Concluding Remarks

### 7.1 MATHEMATICAL MODELING FOR LEUKEMIA TREATMENT

The previous chapters dealt with several aspects of mathematical modeling for treatment of leukemia. When possible the "gray box" system identification approach (Appendix F) was followed, integrating known facts, (accepted) theories and experimental observations. A few differences with previous work by other investigators are summarized in Table 7-1.

Development of the leukemic cell population, unperturbed or challenged with a cytostatic drug; drug distribution; development of drug resistance; probability of relapse after bone marrow transplantation: only a few parts of the vast problem area at a time can be, and have been, isolated to be studied in some detail. For clinical application, eventually, a synthesis of the separate components is required. An idealized scheme of on-line chemotherapy control is shown in Fig. 7.1. For optimal delivery of the proper cytostatic drug(s) it is required to a) monitor an easily observable but relevant variable that is indicative for the patient's state with respect to the leukemia burden; and b) to use the mathematical model(s), to calculate a prognosis of the disease development. The models must be fed with further important external data, e.g., what type and stage of leukemia and what toxicity constraints.

A similar set-up has been realized for automatic insulin administration in diabetic patients. An artificial pancreas contains an insulin pump whose continuous output is controlled based on plasma glucose concentration measurements and a pharmacokinetic model of glucose metabolism, including production and uptake in the liver, renal excretion, glucose utilization dependent (muscles, adipose tissue) or independent (central nervous system, red blood cells) on insulin level, with insulin and glucagon as hormonal controllers [Carson et al., 1983].

A flowchart for arriving at optimal leukemia chemotherapy, from experimental data to clinical application, is shown in Fig. 7.2. It illustrates the connections among the various research fields with their mathematical models. A few developments can be elaborated upon.

Many new developments concern improved techniques for biomedical experimentation. Progress in theoretical population dynamics, system dynamics and signal processing have not so much resulted in new principles and revolutionary models that are readily applicable to cancer chemotherapy. An exception, perhaps, may be the concept of *fuzzy logic*, first formulated by Zadeh [1965]. The concept has been further developed [Takagi and Sugeno, 1985; Ped-

Table 7-1 MAIN DIFFERENCES WITH RESPECT TO PREVIOUS WORK BY OTHER INVESTIGATORS

The present work is more than an integration and extension of the models and working hypotheses of Skipper, Goldie/Coldman, Norton/Simon and Birkhead and the optimization principles by, for instance, Acharaya. It involves:

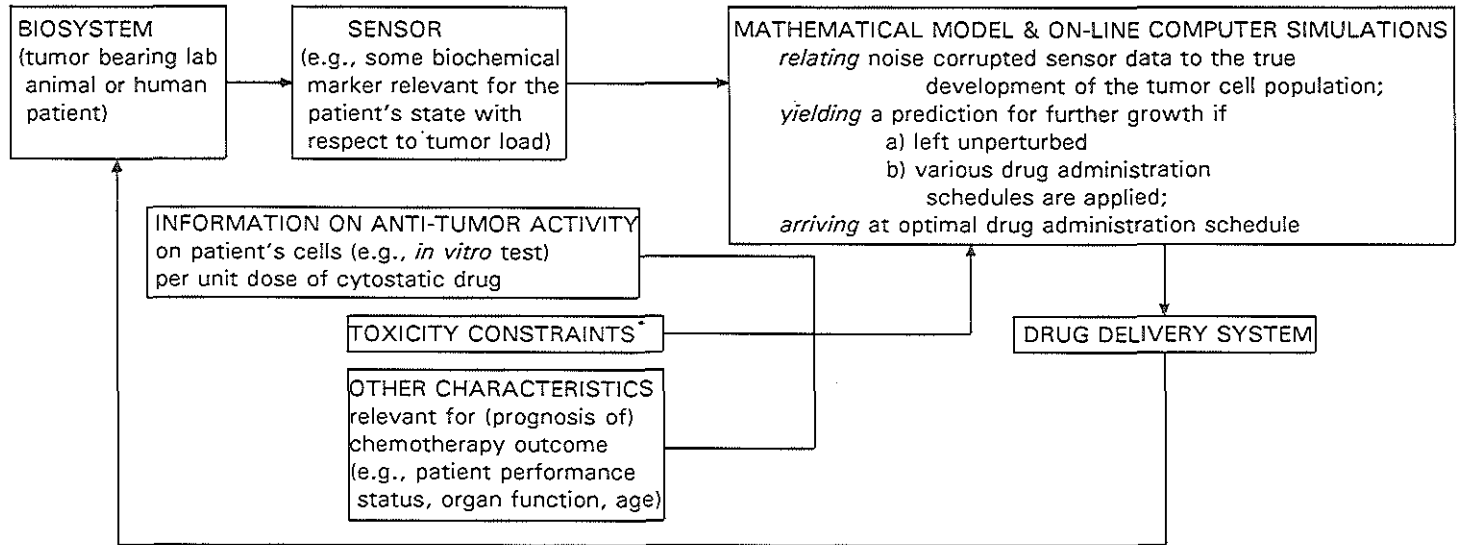
- a) more elaborate *in vivo* data acquisition methods that enable the tuning of experiments to the needs of mathematical modeling; in turn, modeling is based on the hypotheses that arise from the examination of the experimental data
- b) system identification techniques (combining efficiently theoretical considerations and actual experimental observations) applied to the problem of exposing the relationships between the growth of a malignant cell population, drug distribution, cell kill by chemotherapy, and the development of drug resistant sub-lines
- c) use of sophisticated computer software to estimate the parameters of these processes and to optimize controllable factors (therapy regimen) in order to maximize chemotherapy effectiveness. Administration of different doses of one or two drugs in any sequence at non-constant intervals should eventually be allowed
- d) consideration of both phases of leukemia growth, i.e., the early phase when growth is exponential, and the subsequent steady state plateau phase where Gompertzian growth applies
- e) the integral handling of information gained on one complete and coherent system

rysz, 1985, 1989] and, in recent years, was applied to modeling and control of complex and highly nonlinear systems like socioeconomic and industrial processes [Wang and Zou, 1988; Fu and Cai, 1988]. In summary, this methodology offers a possibility to 'calculate', or 'reason', with variables qualified by vague descriptive terms—like 'sometimes, often' or 'good, reasonable, bad'—rather than quantified by exact numerical values [Sugeno, 1985; Van Nauta Lemke, 1991]. It is being developed as a tool to support decision making when circumstances are not all well-defined or data is incomplete, and yet a reasonable choice among several alternatives is required. *Models based on fuzzy logic might, eventually, be useful to estimate and compare the probable success rates of various chemotherapy strategies.* They allow processing and combining vague (clinical) observations like 'the patient shows decreased metabolism' and 'most people of his age respond well to that drug within two days'.

Closely related is the use of knowledge based systems in modeling, simulation and identification [Bekey, 1988]. A verbal description to ascertain effects of medication, toxic substances and physical stresses should be used for modeling systems (e.g., environmental, sociological or biological processes) of which the complexity and many unmeasurable state variables defy crisp and detailed mathematical models.

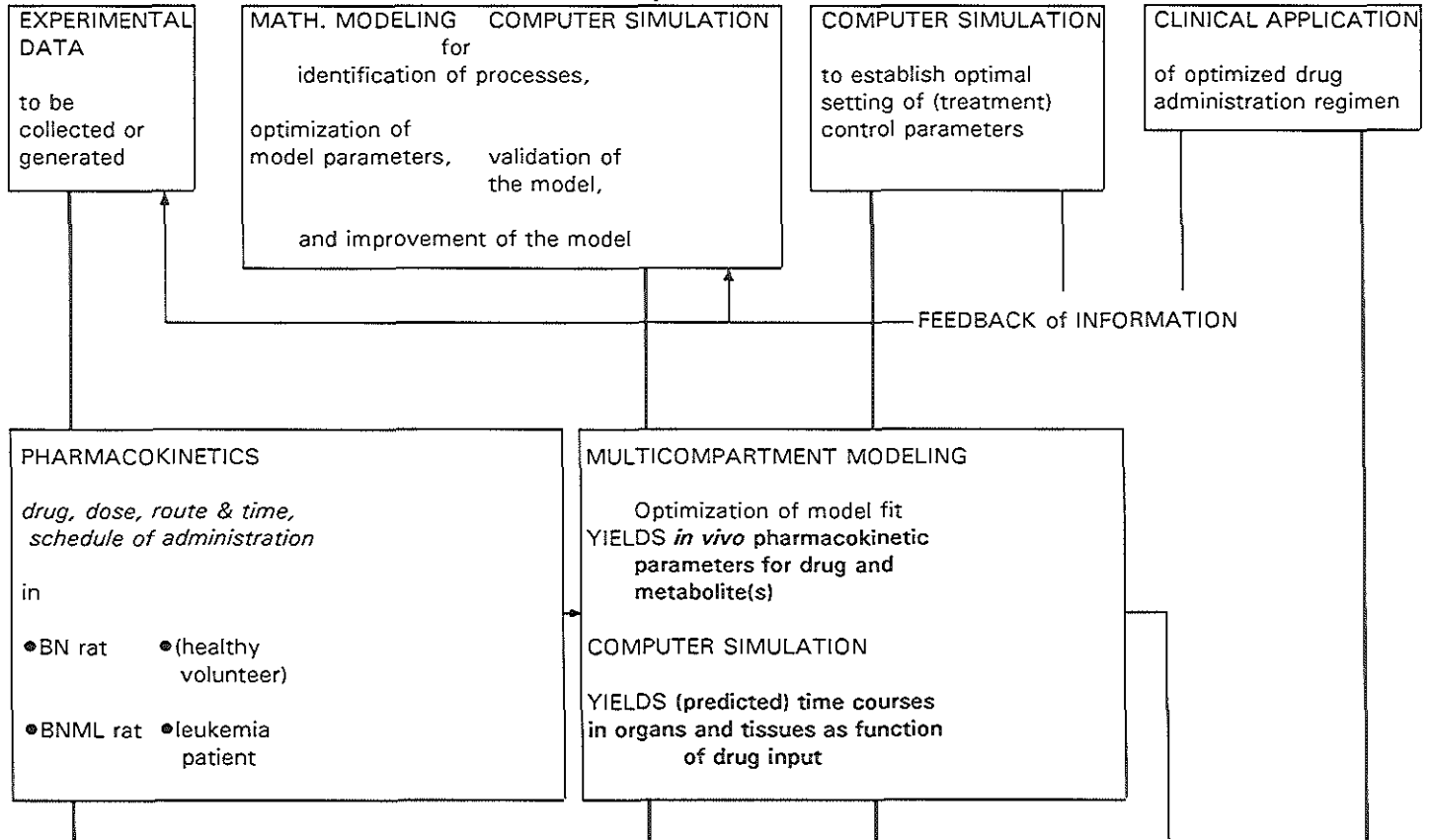


Fig. 7.1 Idealized scheme of on-line chemotherapy control (parts from [Swan, 1987])



\* toxicity limits may be exceeded if rescue is possible afterwards, e.g., through bone marrow transplantation

Fig. 7.2 Flowchart for Leukemia Chemotherapy Optimization



PHARMACODYNAMIC DRUG EFFECTS (*in vivo*)  
*drug concentration, exposure time, environment, tumor load*

for

- normal tissue cells  
(drug sensitive)
- BNML cells  
(drug resistant)
- human leukemia cells

Modeling of intracellular processes relevant for anti-tumor effect and toxicity of cytostatic drugs;

Identification of important parameters involved, e.g., DNA or protein binding, saturation, active membrane drug efflux pump

YIELDS (predicted) cell kill as function of tissue drug concentration and proliferative state

LEUKEMIA GROWTH (*in vivo*) before and REGROWTH after chemotherapy  
*stage of disease, site and size of initial tumor load, cell population heterogeneity*

for

- BNML rat
- human leukemia

Modeling of cell kinetics and population dynamics

Identification of growth rate (growth fraction, doubling time, cell loss factor), heterogeneity (spectral drug sensitivity), mutation rate to drug resistance, cell cycle phase parameters (cell maturity)

YIELDS time courses of cell (sub)population sizes as function of chemotherapy history

COMBINATION YIELDS  
the predicted time course of the size of a malignant cell population before, during and after a specific regimen of cytostatic drug administration

COMPARISON of results of various regimens ENABLES the choice of the optimal regimen

## 7.2 DETECTION AND REGROWTH OF MINIMAL RESIDUAL DISEASE

In the current stage of mathematical modeling of leukemia treatment *the greatest emphasis must still be put on system identification*, i.e., hypotheses on the mechanisms of leukemic cell population growth and drug action are being tested and verified through modeling (see Appendix F). It was shown for one initial condition, how contiguous exponential and Gompertz growth curves describe the development of BNML cell populations in rat organs (Chapter 2.2). Also, that the effect of a single *i.v.* drug dose can best be described as an instantaneous reduction in the number of malignant cells. Later, the validity of the thus acquired knowledge must be confirmed for other circumstances. Then, it is to be used in models for optimization of administration regimens for man. For collecting data that can serve as input for the system identification models, *the availability of sensitive and accurate methods of detection and quantification of (residual) tumor cells is a necessity.*

### 7.2.1 Development of Experimental Detection Methods

**7.2.1.1 In the BNML Rat Leukemia Model.** Up to now the number of BNML cells present in a given tissue at some time point—both before and after application of therapy—has been determined by one of five different methods (Fig. 2.4). In order of ascending sensitivity (i.e., the last method is to be used preferably when very few BNML cells are present relative to the number of normal cells): a) *organ weights*. Compared to normal values, excess weight is attributed to leukemic cells. One leukemic cell in 1 to 10 normal cells can be detected; b) *morphology* (direct counting) detects at best 1 in 20; c) *monoclonal antibody* (RM-124) labeling and flow cytometry detects down to 1 in 10,000. The antibody discriminates leukemic cells from normal marrow cells by a higher antigen density on the former, which is visualized by a fluorescent dye conjugated to the antibody; d) *in vivo clonogenic leukemic stem cell assay* detects 1 in  $10^5$  to  $10^6$ . The number of colonies counted on the surface of spleens of recipient rats, after inoculation of BNML cell contaminated tissue of the donor rat, determines the number of BNML cells in that tissue; and e) *survival time bioassay* detects 1 in  $10^8$ . The median survival time of recipient rats being inoculated with the BNML cell contaminated tissue of the donor rat, determines the number of BNML cells in the inoculum.

The advantage of method c is that it does not require recipient animals, in contrast to methods d and e. Profit of the reduction in time and costs, however, can be harvested only after considerable improvement of the method's sensitivity.

A new detection method under development concerns *genetic markers*. In principle it is meant to recover leukemic cells, and their progeny, after transfer of leukemia. After introducing marker genes into leukemic cells by using retroviral infection as gene transfer technique, these cells and their descendants betray their presence by releasing the specific products the marker genes code for. Various assays are developed to retrieve these products [Yan et al., 1992; Hendrikx et al., 1993].

Integration of the bacterial (*E.coli*) *lacZ* gene into the BNML cell causes production of the enzyme  $\beta$ -galactosidase. Its activity can be measured through colorimetric blue discoloration (down to 1 malignant "blue" cell in  $10^6$  or  $10^7$  normal "white" marrow cells), or fluorometric or flow cytometric fluorescence.

Introduction of the *neo<sup>R</sup>* gene makes the BNML cell resistant to neomycin and its analogues, a property that is expressed in the survival of colonies when a cell population is exposed to this drug. The lower detection limit here is 1 leukemic colony forming cell per  $10^6$  normal cells plated.

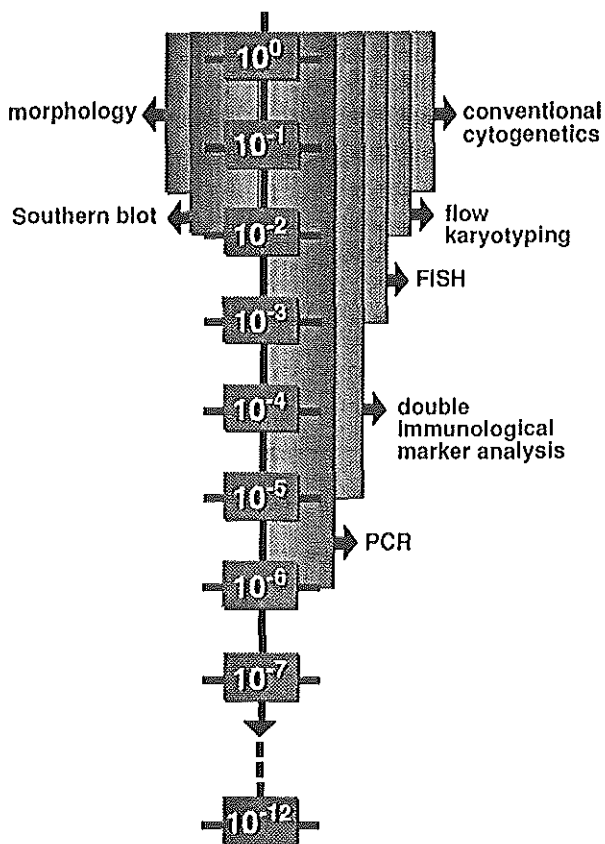
A problem to be conquered is the assessment of the efficiency of the retroviral infection (how to make sure that all relevant cells are genetically marked), as well as gene transfer in the *in vivo* situation.

**7.2.1.2 In Human Leukemias.** Early detection of (residual) leukemic cell proliferation is important in relation to the patient's cure probability. In human leukemia several detection methods are used (Fig. 7.3) that operate at three different levels, i.e., a) cellular (cell surface, cytoplasm and/or nucleus markers), b) chromosomal, and c) DNA/RNA level.

*a) Double Immunological Marker Analysis (cellular level).* There are no single antigens that are specific for leukemic cells. The detection method is based on unique combinations of two immunological markers that are present in/on malignant cells and absent in/on normal marrow cells. The method was successfully used to detect and quantify malignant cells in childhood T-ALL (See Chapter 3), with a sensitivity of about  $10^{-5}$ . Unique combinations of immunological markers, characteristic for several types of leukemia (T-ALL, B-ALL, AML) are being identified and catalogued, e.g., [Adriaansen, 1992; Van Dongen et al., 1992, 1993].

*b) Cytogenetics (chromosome aberrations).* Leukemia can often be associated with characteristic chromosome aberrations. When such genetic abnormalities are found, e.g. through conventional karyotyping [Rowley, 1973; Yunis, 1983; Croce, 1987], this not only indicates the presence of malignant disease, but also helps its classification [Sandberg and Turc-Carel, 1987]. Prognostic value can be derived from retrospective studies correlating chromosomal aberrations and disease development [e.g., Secker-Walker, 1990]. Well-known is the so-called

Fig. 7.3 Levels of Detection of Leukemic Cells in Human Leukemia [Hagenbeek, 1992]



Philadelphia chromosome [Rowley, 1973], a translocation of chromosomes 9 and 22, which is present in the majority of chronic myelocytic leukemia cases. Other structural aberrations include deletions. Furthermore, numerical aberrations have been described in a variety of human leukemias.

Preliminary studies on the applicability of *flow karyotyping* for the detection and quantification of chromosome aberrations have been performed, both in BNML and human CML [Arkesteijn et al., 1987, 1988, 1990]. The minimum detection level has not yet been established, but is estimated to be as good as that of conventional karyotyping, i.e., in the order of  $10^{-2}$ . Flow cytometry enables fast processing of large numbers of chromosomes in suspension. The A(denosine)-T(hymidine) and C(ytidine)-G(uanosine) base pairs are stained with two different fluorescent dyes (e.g., Hoechst 33258 and chromomycin A3). Flow cytometry results in a bivariate plot of fluorescence intensities. Abnor-

malities can be detected as the chromosomes present, each with a specific ratio of A-T and C-G contents, appear at characteristic locations in the plot.

A disadvantage of karyotyping is that cells must be in mitosis to show the separate chromosomes. To enlarge the usually small fraction of metaphases, *premature chromosome condensation* (PCC) techniques have been attempted [Johnson and Rao, 1970; Hittelman, 1986]. To avoid this disadvantage, and be able to use interphase cells as well, a new approach uses *fluorescent in situ hybridization* (FISH) with chromosome specific DNA probes [Pinkel et al., 1986; Cremer et al., 1988; Van Dekken et al., 1989]. Such a probe is a certain small sequence of bases that can bind to a complementary DNA sequence, pinpointing a corresponding unique location on the 'marker' chromosome, if present. As the probe is labeled with a fluorescent conjugate, the malignant cell thus is recognisably marked. The detection limit is expected to be at least  $10^{-3}$ . A useful probe—probe development is a research area in itself—implies knowledge about the sequence of bases at the spot to be searched for.

c) *DNA/RNA amplification*. The last remark holds when using the *polymerase chain reaction* (PCR) technique to find molecular lesions characteristic for leukemia or lymphoma [Potter et al., 1993]. Lesions occur at breakpoints of chromosomal translocations, as rearranged genes coding for antigen receptors on B- and T-lymphocytes, or as point mutations in cell regulation genes. The PCR technique is based on splicing the double stranded DNA by heating and, after cooling down, letting each single strand synthesize its complementary part in one direction, starting from one end of the aberrant DNA sequence. This end is marked—beforehand, at yet another temperature—with a so-called primer, a short piece of the known DNA sequence. Repeating this three-phased cycle some 30 times,  $2^{30}$  pieces of aberrant DNA are produced. Of course, this multiplication only works if the searched aberrant DNA sequence actually is present. Then, one abnormal cell in about  $10^5$ - $10^6$  normal cells can be detected.

### 7.2.2 Regrowth of the Residual Leukemic Cell Population

It should be stressed that the detection of residual cells at one time point in itself is not sufficient to make predictions about the development of disease or the imminence of relapse. Not only their presence must be recognized, but *information must be obtained as well about the proliferative state of the cells and the cell population dynamics*. To this purpose it might, for instance, be measured to what extent BrdUrd, a compound incorporated into the DNA of actively proliferating cells only (See Chapter 7.4.1), is taken up. For genetically marked leukemic cells a functional assay may be used [Hendrikx et al., 1993]. Otherwise, the cell population must be monitored at several time points to determine its dynamic behavior. The multiharmonic DNA histogram analysis method

presented in Chapter 5 is a useful tool for determining cell kinetics.

BNML growth curves in the principle target organs have been constructed for leukemia induction with  $10^7$  malignant cells and regrowth after a dose of cyclophosphamide has been studied (Chapter 2.2). *Drug effect could be modeled as instantaneous log cell kill, with identical growth and regrowth curve parameters.* However, a discrepancy was noticed in Chapter 2.4.3.1, concerning the relation between median survival time and inoculated BNML cells on the one hand, and log cell kill and increase in lifespan on the other hand. To shed more light on this problem and verify present extrapolations, among which the ED<sub>50</sub> hypothesis (see Chapter 6.2.4.1), *BNML growth after at least one other inoculum size should be experimentally observed (datapoints of time versus number of BNML cells in the various organs).*

A recent modeling development for human leukemia seeks to determine the mean and spread parameters of the amount of minimal residual disease after treatment, as well as regrowth doubling time, from examination of remission durations [Gregory et al., 1991]. From a recent AML trial it was inferred that patients with long first remission durations had less residual disease and a slower rate of regrowth. Also, their second remission durations are longer. Making correlations explicit may help in finding therapy strategies for the second remission induction. Of importance is that the model *seems to confirm the necessity to eradicate all leukemic cells in order to cure.*

### 7.3 PHARMACOKINETICS AND PHARMACODYNAMIC DRUG EFFECTS

Further development of the pharmacokinetic models is desirable, to increase knowledge on the mechanisms of drug distribution and metabolic processes. At present only a single dose of a single drug in the normal BN rat has been studied. Chapter 4 shows the applicability of multicompartiment modeling and maximum likelihood parameter estimation for system identification in pharmacokinetics. *The influence of dose variation and of repeated doses—intermittent and continuous (infusion)—should be investigated with respect to (nonlinear) effects like drug accumulation, saturation and attainment of steady state levels.* Thus, eventually, the bioavailability of a drug at certain relevant sites and during certain time intervals, as function of administration regimen, can be accurately predicted.

*Pharmacokinetics in the leukemic rat is the next step.* For example, the influence of the stage of the disease (leukemic cell burden) still is ill-known although the prognostic importance of tumor load has long been recognized [Langermann et al., 1982]. The malignant cell mass will change the distribution pattern and possibly influences the rate of metabolism.

New measurement techniques for drug uptake in organs and tumors are



appearing, e.g., the use of radiolabeled drugs allows registration of cumulated activity in organs as function of time [Strand et al., 1993]. On the cellular level flow cytometry can be applied for cytostatic agents with fluorescent properties. The *in vitro* drug uptake can be measured in neoplastic cells in patient-derived cell cultures. In this way changes—due to exposure to cyclosporin A—in accumulation of daunomycin in various resistant tumor cells have been measured [Nooter et al., 1989, 1990] in a study of the potential of *in vivo* treatment with calcium channel blockers or other cell membrane transport modifying compounds to reverse existing drug resistance.

*Finally, combined administration of multiple drugs should be studied for additive, synergistic or antagonistic effects.*

### 7.3.1 Analysis of Existing Pharmacokinetic Data (BN or BNML and Adriamycin or Daunomycin)

Data is available on the pharmacokinetics of the anthracyclines adriamycin (ADR) and daunomycin (DAU) in the normal BN rat and the BNML rat (drug concentration in various tissues and body fluids, detected by HPLC or fluorescence methods at time points during 48 h after *i.v.* administration of a single dose). These data can be, and partially have been, used as input in the multicompartiment models for further analysis.

ADR in the BN rat has been analyzed in a ten-compartment model, yielding good fits in the tissues, but a too high plasma response. The latter can be improved, at the expense of the tissue fits, by enlarging the plasma volume (to include extracellular fluid) or by changing the kidney-to-urine pathway to kidney-to-plasma transport. Putting extra weight to the plasma data is a mathematical manipulation that also forces the concentration—time history to a better plasma fit, but in contrast to the other two mentioned modifications this does not add any physiological meaning.

Application of the ten-compartment model to BNML data yielded poorer fits. The tumor load apparently influences the mechanism of the pharmacokinetics [Nooter et al., 1984]. Plasma concentration is higher, while lower ADR concentrations are observed in liver, spleen and bone marrow, which are the organs in which the malignant cell population grows in greatest abundance.

DAU (and its principal metabolite daunomycinol, DOL) data in the BN rat yielded good fits in models with few compartments, which are clinically irrelevant, but poorer fits in the larger models that should offer more detail information. Analyzing three subsequent periods, i.e., 0-24 h, 24-48 h and 48-72 h, each starting with an equal *i.v.* bolus dose, showed a non-linear response, i.e., the concentration—time histories were different during each period, especially between periods one and two.

These (preliminary) analyses prove conceptual shortcomings of the present models, which are based on passive diffusion processes. *More sophisticated*

*models should be developed, the effort being justified by the fact that anthracyclines are likely to be continuously used in the clinic for quite some time to come.* This requires more investigations on, e.g., the enterohepatic cycle (reabsorption with a time-delay); inactivation of the drugs (mass balance okay?); recirculation (drug being released from cells that are killed); cell type dependent passive influx/active efflux (cellular drug pump) and saturation phenomena.

Protein binding or covalent binding to DNA might explain to some extent the observations of the three-doses administration. If the drug binds when the first dose is given it will go undetected (due to the measurement method), resulting in too low observed values. The model interprets this "lost mass" as excretion to tissue. Because of the occupied binding sites the second dose yields higher free drug concentrations picked up by the detection method. Concentrations after the third dose are not very different from the previous ones, suggesting prolonged saturation of binding sites. Two-dimensional electrophoresis might check this hypothesis.

Other modes of drug administration also should be investigated. Data collected on a three hour *i.v.* infusion of DAU in BN rats should be processed in a compartmental analysis, to establish any differences of the transfer rate constants with respect to the *i.v.* injection mode. Differences in observed concentration—time points seem to suggest an advantage of the infusion as less cardiotoxicity can be expected for the same degree of myelosuppression [Nooter et al., 1986]. Observations on DAU pharmacokinetics in human patients (bolus dose and infusion) are available as well [DeGregorio et al., 1984; Kokenberg, 1991].

### **7.3.2 Drug Sensitivity Testing *in vitro***

*One information category necessary for modeling leukemia chemotherapy is the antitumor activity of certain fixed doses of various cytostatic agents.* The sensitivity of a patient's neoplastic cells to different drugs can now be measured with the *in vitro* MTT (3-(4,5-dimethylthiazol-2yl)-2, 5-diphenyltetrazolium-bromide) assay [Pieters et al., 1993]. Essentially, this assay explores the survival of cultured malignant cells that are exposed to graded concentrations of a drug. Several drugs can be tested simultaneously, using an array of cultures growing on one tray.

During the past few years the MTT assay has been used to examine the variation in sensitivity of childhood acute lymphoblastic leukemia cells to vincristine, cytosine arabinoside, daunorubicin, methotrexate, maphosphamide, thioguanine, prednisolone and other cytostatic drugs [Kasper et al., 1991; Pieters et al., 1991, 1992a, 1992b, 1993]. The *in vitro* tests bear relevance to the *in vivo* situation, as the degree of cellular drug resistance to thioguanine, daunorubicin or prednisolone appears to correlate with the duration of complete remission. Such a relationship was not found for vincristine. Malignant cells of

relapsed children appear to be more resistant to several drugs than during first remission induction. Again vincristine is an exception, indicating that increased resistance is acquired rather than being caused by selective kill of more sensitive (intrinsic property) subclones. Some correlation was found between *in vitro* drug resistance and the age class of the patient (< 18 mo, in between, > 10 y) or immunophenotype. In general, normal peripheral blood cells are more resistant than leukemic cells. The source of the leukemic cell sample makes no difference, i.e., malignant cells from the peripheral blood or the bone marrow of the same patient are equally drug resistant/sensitive.

*Sensitivity testing in vitro may be a step toward personalizing chemotherapy, providing information about cell kill per unit of dose for different drugs and individual patients.*

### 7.3.3 Drug Resistance

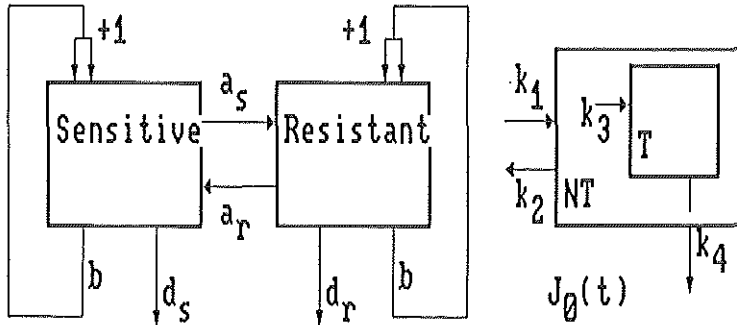
Refinement of the simple model for resistance development (Chapter 2.4.3) requires more data. BNML subcell lines resistant to various drugs are now available for *in vitro* and *in vivo* experiments. Leukemia induction with a constant inoculum size, but different compositions of sensitive and resistant BNML cells should reveal any differences in growth kinetics. Treatment with a constant drug dose at the same day should prove that the treatment results closely correlate with the expected potential log cell kill of that drug dose, corrected for the composition of the BNML cell mixture. If not so, indications for additional drug resistance development would be obtained.

In other experiments flow cytometry with BrdUrd (See Chapter 7.4.1) could be used to measure the rate of resistance development. This rate can be deduced by treating samples of cells at various times with a lethal dose of a cytostatic agent, followed by exposure to BrdUrd. Each time the fraction of cells that have incorporated BrdUrd (i.e., the surviving, ergo drug resistant, fraction) is compared with the result of the previous time point.

*To study drug resistance phenomena a generalized mathematical model, proposed by Michelson and Slate [1989], could be useful.* This model allows many pathways leading to the emergence of drug resistance, like decreased drug uptake, increased efflux, increased metabolism, change in drug target properties, and combinations of these mechanisms. It is a hybrid model, consisting of a stochastic part that describes cell population growth, and a deterministic part that describes the micropharmacological distribution of the drug (Fig. 7.4). The population growth model is very similar to the one shown in Fig. 2.26. The micropharmacology of the cell is modeled with a drug target compartment in a nontarget compartment, surrounded by environment. In the environment the drug concentration may vary, as denoted by the function  $J_0(t)$ . Different values may be assigned to the parameters for drug sensitive and resistant cells. Transition rates  $k_1$  and  $k_2$  represent drug uptake and the drug efflux pump, respec-

Fig. 7.4 Model for Emergence of Drug Resistance [Michelson and Slate, 1989]

Left: stochastic compartment representation of heterogeneous tumor, with sensitive and resistant subpopulations. Transition probabilities include drug-induced cytotoxicity ( $d_s$ ,  $d_r$ ), mitosis ( $b$ ) and phenotype mutation ( $a_s$ ,  $a_r$ ). Right: Target and Nontarget compartments to describe micropharmacology of the cell. Transition rates include drug uptake ( $k_1$ ), efflux ( $k_2$ ), transport into target compartment ( $k_3$ ) and inactivation ( $k_4$ ).  $J_0(t)$  denotes extracellular drug concentration



tively. The latter plays an important role in multidrug resistance. A clearance/inactivation rate,  $k_4$  is associated with drug metabolism by repair and clearance enzymes, like glutathione reacting with alkylating agents. Overexpression of target enzyme can be modeled by assigning different values to the transition probabilities for cytotoxicity,  $d_s$  and  $d_r$ , when drug concentration levels are similar for sensitive and resistant cells. The model was applied to study resistance reversal scenarios, e.g, by simulating doxorubicin plus calcium entry blocker treatment of multidrug resistant cells [Slate and Michelson, 1990]. The model could be used in a similar way, in combination with accumulated experimental data, to study emergence of cyclophosphamide resistance in the BNML. Incorporation of cell cycle phase specific drug action would make the model suitable for simulating e.g. ARA-C resistance. Other extensions could include growth rate heterogeneity of leukemic cells [Chan and Kuczek, 1991], spectral sensitivity (heterogeneity with respect to the degree of drug resistance [Carl, 1989] and treatment with two different drugs [Abundo et al., 1989].

#### 7.4 CELL (CYCLE) KINETICS AND CELL POPULATION DYNAMICS

Work on cell cycle models continues because accurate data on the fractions of cells of a certain biological age (maturity) at a given chronological time is needed for various applications among which cancer radiotherapy and chemotherapy (e.g., [Zaider and Minerbo, 1993; Jones et al., 1993]).

More than a decade ago it was recognized that the time sequence of drug

administration is an important factor influencing the cytostatic effect of cell cycle (phase) specific drugs (e.g., [Burke et al., 1982]). Understanding the cell kinetics of the malignant cell population allows explaining this phenomenon. Recruitment, i.e., triggering resting or quiescent cells into proliferation, enlarges the number of malignant cells that are sensitive for cycle specific drugs. Subsequent synchronization of the proliferating cells—all cells brought into the same state of maturity—allows maximal cell kill when a cell cycle phase specific drug is given at the right moment (e.g., [Colly et al., 1984a,b]). Detection of cell kinetic changes is possible through flow cytometry and the multiharmonic DNA histogram analysis method (Chapter 5).

Progress in molecular biology allows taking a closer look at phenomena at the cellular level (e.g., processes of drug transport across the cell membrane or cellular production of and response to growth factors) and even at the level of the cell nucleus (DNA, RNA; damage of the genetic code). Thus, new information is obtained on the regulation of the production of normal and malignant cells, and on the direct impact of cytostatic agents on those cells themselves.

When studying what drives their malignant counterparts, knowledge about the regulation of normal hemopoietic cell populations is more than welcome. This research area also offers many subjects to be still explored, and, for better understanding, use is made of mathematical models [e.g., Wichmann et al., 1988; Fokas et al., 1991; Grabosch and Heijmans, 1991]. Growth factors, also known as interleukins or colony stimulating factors, play an important role in the production and maturation of hemopoietic subpopulations as well as leukemic cells. Their mechanisms of control are being investigated, a.o. through mathematical modeling [e.g., Schmitz et al., 1993].

#### 7.4.1 Flow Cytometry

*Flow cytometry remains an important tool for observing proliferation and maturation of (sub)populations of cells*, thus generating data that, after proper analysis, can be used in the development and validation stages of mathematical modeling. For example, Fokas et al. [1991] presented a theoretical model for granulocytopoiesis and chronic myelogenous leukemia (CML). This model shows how CML cells can ultimately outnumber the normal cells, and how this process can be very slow, on the assumption—to be verified—that a greater fraction of CML cells is produced by division rather than by maturation.

In the present thesis a method for the analysis of DNA distributions is described (Chapter 5). Today, similar programs come standard with new commercial flow cytometers. Better programs are still being developed, e.g., [Del Bino et al., 1985; Baldetorp et al., 1989]. Some of them employ the maximum likelihood principle [Lampariello and Del Bino, 1988]. They can deal with the simultaneous presence of cell populations with different ploidy levels (abnormal DNA contents, i.e., aneuploidy, hypoploidy, near-diploidy or

tetraploidy), presence of debris (cellular fragments, and they use internal standards (chicken and/or trout red blood cells) of fluorescence intensity.

Modern flow cytometers can process cells very fast. They are often equipped with multiple lasers, enabling better distinction and sorting of cell subpopulations—and, of a subpopulation, the various cycle phase fractions—based on multiple fluorescent labels that are specific for certain membrane characteristics and/or nuclear DNA contents. For example, to determine the fractions of cells in the various phases of the cell cycle and the ratio of resting and proliferating cells, a method of *in vivo* labeling with bromodeoxyuridine (BrdUrd) has been developed [Dean et al., 1984] and applied (e.g., [Raza et al., 1985, 1987]). BrdUrd is only taken up by cells actively synthesizing DNA. (Therefore, the method may be less suitable for analyzing samples of human solid tumors that generally have small growth fractions). The uptake is detected by means of a monoclonal antibody, specific against BrdUrd, to which a fluorescent dye is coupled. Counter staining with e.g. propidium iodide (PI) reveals the amount of total DNA. From the (series of) two-dimensional flow cytometric fluorescence plots the desired variables can be derived (e.g., [Yanagisawa et al., 1985; Lacombe et al., 1988; Bertuzzi et al., 1993]) even where badly perturbed DNA histograms are concerned [Ormerod et al., 1987]. Cell cycle times and phase durations can be determined as well. Total cycle time and the period of sensitivity to a phase specific drug are required as input for the Z-method of improving therapy scheduling (see Chapter 7.4.2).

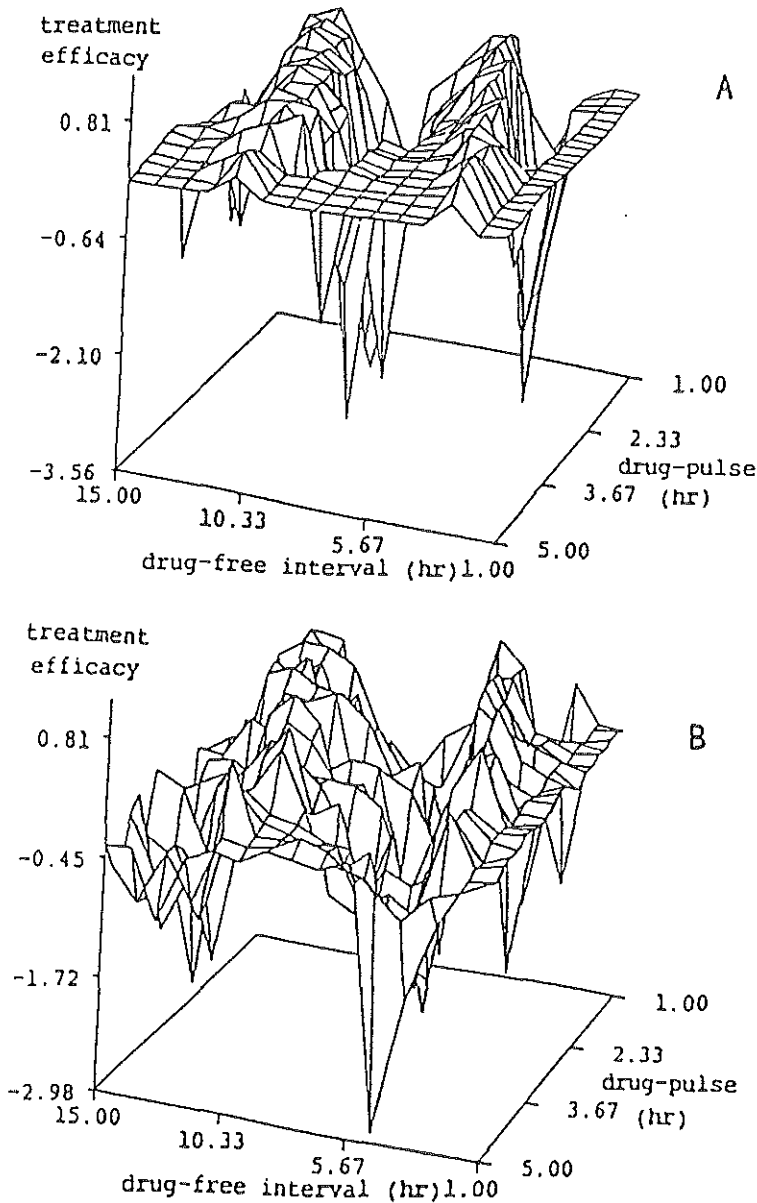
#### 7.4.2 Improving the Therapeutic Index of Cell Cycle Specific Cytostatic Agents; the Z-Method

Administration of cytostatic drugs in cancer chemotherapy is limited by toxicity to normal tissues. Certain administration regimens exhibit a better therapeutic index than do others, i.e., they yield more tumor cell kill for less adverse effects on normal tissues. Which are the better regimens and why?

To answer this question where cell cycle specific drugs are concerned, a systematic investigation has been performed by Agur et al. [1986, 1988].

Their Z-method has been put forward, based on a recent theory on population dynamics in harshly varying environments [Agur, 1985], as a new policy for minimizing the cytotoxicity to host tissues of antineoplastic cell cycle specific drugs. When the drug is administered as a number of short pulses at a certain frequency, with drug-free periods in between, using stochastic mathematical models (incorporating knowledge on log cell kill per dose and on neoplastic and normal cell proliferation, among which cycle time and duration of drug susceptible phase) the time till extinction of each of the two populations can be calculated,  $T(T)$  and  $T(N)$  for tumor and normal cells, respectively. With computer simulations the treatment efficacy,  $Z(\delta, \omega)$  as function of drug pulse

Fig. 7.5 Worst Case Examples of Simulation Results with Z-Method [Agur, 1988]  
 Treatment efficacy, Z as function of drug pulse duration and drug free interval. Z should tend to 1 for fast extinction of the malignant cells and slow extinction of normal cells. A) Erythroblasts (cycle time  $T_C=7.0\pm 0.5$  h, drug sensitive interval  $T_S=3.0$  h) and Leukemic Cells ( $T_C=6.0\pm 0.5$  h,  $T_S=2.5$  h), B) Erythroblasts ( $T_C=7.0\pm 0.5$  h,  $T_S=3.0$  h) and Leukemic Cells ( $T_C=7.0\pm 2.0$  h,  $T_S=3.0$  h) in WT & Balb C mice



duration,  $\delta$  and drug free interval,  $\omega$  can be evaluated.

$$Z(\delta, \omega) = 1 - T(T) / T(N). \quad (7.1)$$

The longer it takes to eradicate the normal cells ( $T(N)$  up) and the faster the malignant cells vanish ( $T(T)$  down), the more  $Z$  tends to one, corresponding with a better therapeutic index. The time till extinction can be evaluated by calculating the population size after every time interval  $T_c$  (mean cell cycle time), assuming that during that time  $A$  ( $2 \leq A \leq 1$ ) daughter cells per cell are produced, unless drug is present during the sensitive period,  $T_s$  of the cell cycle, in which case the cell is killed.

Subsequently, optimization of the treatment schedule (length of  $\delta$  and  $\omega$ ) is possible, e.g., by determining the minimum elimination time of the tumor for either maximum elimination time of the normal cell population or a maximum number of normal cells left after treatment.

The  $Z$ -method does not depend on an exact knowledge about the intrinsic population growth nor on the exact values of most parameters in the system. Crucial assumptions are: 1) a periodical change between drug sensitive and insensitive phase of the limiting normal cell population; 2) the mean cycle time of the cell population can be estimated; 3) normal and malignant cell populations differ in mean cycle time, or the normal cell population has a smaller variance in cell cycle time. Figure 7.5 illustrates two worst case examples. Favorable combinations of  $(\delta, \omega)$  yielding high values of  $Z$  can be distinguished from unfortunate ones (low  $Z$ ), even when cycle times and ARA-C sensitive phase durations are nearly equal for both cell populations (erythroblasts in WT & Balb C mice:  $T_c = 7.0 \pm 0.5$  h,  $T_s = 3.0$  h and leukemic cells:  $T_c = 6.0 \pm 0.5$  h,  $T_s = 2.5$  h or  $T_c = 7.0 \pm 2.0$  h,  $T_s = 3.0$  h). Short drug pulses ( $\delta = 1-2$  h) after a period that is a multiple of the normal cells' mean cycle time yield best results.

It would be very interesting to check the  $Z$ -method's theoretical results with respect to populations of normal hemopoietic stem cells and leukemic cells when treating BNML with the cell cycle specific drug ARA-C, as an addition to the studies of Colly et al. [1986]. This would involve: 1) Validation of the assumptions for ARA-C, BNML and normal hemopoiesis, which includes gathering log cell kill data and growth kinetics data, mainly from literature and previous research projects; 2) Specification of the optimization problem and the dynamic processes involved for implementation on a computer system; 3) Running computer simulations for a. evaluation of treatment efficacy as function of drug pulse duration and drug pulse frequency; b. evaluation of consequences of working with inaccurate estimates of system parameter values; 4) Verification of selected optimal strategies in experimental animals and, after that, possibly in human patients. This would provide a scientifically based theoretical background for the design of administration schemes with improved therapeutic index for



cell cycle specific cytostatic drugs in general.

### 7.4.3 Visualization of Tumor Growth

Computer hardware developments during the past decade have resulted in powerful yet affordable personal computers equipped with color monitors and suitable for real-time 2D and 3D computer graphics. It seems a logical step to employ this tool in modeling cell population dynamics, to visualize simulated tumor growth. Düchting et al. [1989, 1990] have developed software to show how the composition of tumor cross sections changes with time, with or without chemo-/radiotherapy. Although designed for solid tumors, with some modification the program may be made relevant also for leukemic cells, e.g., growing in a matrix of bone marrow cells. A random maturity is assigned to a cell and a random draw from the cell cycle phase duration distribution determines the transition to the next phase. This transition can be seen, as cells in different phases of the cell cycle are represented by different symbols. The effects of using various production rules can be evaluated, e.g., normal cells may only divide a limited number of times and only if there is space available (because a neighbor cell has died), while a cancer cell may divide without further restriction than that it must be sufficiently close to a capillary blood vessel (for nutrition). With similar rules for occurrence of natural cell death and effects of therapy, the development of the tumor can be studied—one sees what happens—and compared to experimental data. Thus, the production rules used can be analyzed with respect to validity, and knowledge is gained about the underlying processes. Also, the effects of alternative therapies, different treatment intervals or dose rates, etc., can be compared. At present the model itself is still too simple, e.g., the cell volume has a constant value; there is no possibility for metastasis or any immunological reactions. But, in the long run, the objective is to arrive at optimal treatment strategies and schedules by computer simulation prior to clinical therapy.

Notwithstanding the considerable increase in cure rate achieved during the past decades, leukemia treatment today is still very much a matter of working with empirical methods and rules of thumb. Little is known with great certainty about cell population dynamics, pharmacokinetics and drug action in the leukemic patient, let alone the combined *in vivo* effects of multiple drugs on malignant and normal cell populations. Such knowledge is necessary for appropriate disease control and design of an optimal treatment strategy. Building mathematical models contributes to the generation of this knowledge, as the examples given in this thesis may illustrate. These examples address a number of topics in leukemia treatment, from rather scattered parts of the research area. The

challenge is both in (stepwise) extending and refining existing models to a high level of sophistication (realism), and in interconnecting them. It is hoped, that this thesis encourages people with a (bio)medical education and those with a systems control background to cooperate in accumulating, processing and interpreting data. Such a multidisciplinary approach offers good prospects of reducing a chaos of observed data, systematically through evaluation of working-hypotheses, to firmly formulated laws of nature. Then, these laws can be used to optimize leukemia treatment in a truly rational way.

## 7.5 REFERENCES

- Abundo M, Rossi C, Benassi M, Gentile FP, Mauro F (1989) A stochastic model of the acquisition of drug resistance during antineoplastic chemotherapy. *Health Phys* 57:349-354
- Adriaansen HJ (1992) Heterogeneity in acute myeloid leukemia: basic and diagnostic studies. Thesis. Rotterdam: Erasmus University.
- Agur Z (1985) Randomness, synchrony and population persistence. *J Theor Biol* 112:677-693
- Agur Z (1986) The effect of drug schedule on responsiveness to chemotherapy. *Ann NY Acad Sci* 504:274-277
- Agur Z (1988) A new method for reducing cytotoxicity of the anti-AIDS drug AZT. In: Vichnevetsky R, Borne P, Vignes J (eds) *Proc 12th IMACS World Congress on Scientific Computation*, Paris, France, July 18-22, 1988, 4:602-604
- Agur Z, Arnon R, Schechter B (1988) Reduction of cytotoxicity to normal tissues by new regimens of cell-cycle specific drugs. *Math Biosci* 92:1-15
- Arkesteijn GJA, Martens ACM, Jonker RR, Hagemeyer A, Hagenbeek A (1987) Bivariate flow karyotyping of acute myelocytic leukemia in the BNML rat model. *Cytometry* 8:618-624
- Arkesteijn GJA, Martens ACM, Hagenbeek A (1988) Bivariate flow karyotyping in human Philadelphia positive chronic myelocytic leukemia. *Blood* 72:282-286
- Arkesteijn GJA, Van Dekken H, Martens ACM, Hagenbeek A (1990) Clinical applications of flow karyotyping in myelocytic leukemia by stimulation of different subpopulations of cells in blood or bone marrow samples. *Cytometry* 11:196-200
- Baldetorp B, Dalberg M, Holst U, Lindgren G (1989) Statistical evaluation of cell kinetic data from DNA flow cytometry (FCM) by the EM algorithm. *Cytometry* 10:695-705
- Bekey GA (1988) Knowledge based systems in modeling, simulation and identification. In: Han-Fu Chen (ed) *Identification and parameter estimation*, Preprints of the 8th IFAC/IFORS Symposium, Beijing, Peoples Republic of China, August 27-31 1988, Academia Sinica and Pergamon Press, 69-74
- Bertuzzi A, Gandolfi A, Sinisgali C, Starace G (1993) Mathematical methods for cell cycle analysis from flow cytometric DNA-BrdUrd distributions. Presentation at the International Seminar on Control and Modelling of Cancer Cell Population, Gliwice, Poland, 20-23 Oct 1993
- Burke PJ, Vaughan WP, Karp JE, Saylor PL (1982) The correlation of maximal drug dose, tumor recruitment, and sequence timing with therapeutic advantage: schedule dependent toxicity of cytosine arabinoside. *Med Pediat Oncol Suppl* 1:210-208
- Carl J (1989) Drug-resistance patterns assessed from tumor marker analysis. *J Natl Cancer Inst* 81:1631-1639
- Carson ER, Cobelli C, Finkelstein L (1983) *The mathematical modeling of metabolic and endocrine systems; model formulation, identification, and validation*. John Wiley & Sons, New York

- Chan TCK, Kuczek T (1991) The rate normal model of tumor drug resistance. *Proc of the Am Assoc Cancer Res* 32:358
- Colly LP, Van Bekkum DW, Hagenbeek A (1984a) Cell kinetic studies after high dose ARA-C and adriamycin treatment in a slowly growing rat leukemia model (BNML) for human acute myelocytic leukemia. *Leuk Res* 8:945-952
- Colly LP, Van Bekkum DW, Hagenbeek A (1984b) Enhanced tumor load reduction after chemotherapy induced recruitment and synchronization in a slowly growing rat leukemia model (BNML) for human acute myelocytic leukemia. *Leuk Res* 8:953-963
- Colly LP, Peters WG, Willemze R (1986) Effect of the interval between high dose 1- $\beta$ -D-arabinofuranosylcytosine injections on leukemic cell load, intestinal toxicity, and normal hematopoietic stem cells in a rat model for acute myelogenous leukemia. *Cancer Res* 46:3825-3827
- Cremer T, Lichter J, Borden DC, Ward DC, Manuelis L (1988) Detection of chromosome aberrations in metaphase and interphase tumor cells by in situ hybridization using chromosome-specific library probes. *Human Genet* 80:235-246
- Croce MC (1987) Role of chromosome translocations in human neoplasia. *Cell* 49:155-156
- Dean PN, Dolbear F, Gratzner H, Rice GC, Gray JW (1984) Cell cycle analysis using a monoclonal antibody to BrdUrd. *Cell Tiss Kinet* 17:427-436
- DeGregorio MW, Holleran WM, Macher BA, Linker CA, Wilbur, JR (1984) Kinetics and sensitivity of daunorubicin in patients with acute leukemia. *Cancer Chemother Pharmacol* 13:230-234
- Del Bino G, Bruni C, Koch G, Mazzini G, Costa A, Silvestrini R (1985) Validation of a mathematical procedure for computer analysis of flow cytometric DNA data in human tumors. *Cytometry* 6:31-36
- Düchting W, Lehrig R, Rademacher G, Ulmer W (1989) Computer simulation of clinical irradiation schemes applied to in vitro spheroids. *Strahlenther Onkol* 165:873-878
- Düchting W (1990) Tumor growth simulation. *Comput & Graphics* 14:505-508
- Fokas AS, Keller JB, Clarkson BD (1991) Mathematical model of granulocytopenia and chronic myelogenous leukemia. *Cancer Res* 51:2084-2091
- Fu JG, Cai JB (1988) Fuzzy identification of systems and its application to process control. In: Han-Fu Chen (ed) Identification and parameter estimation, Preprints of the 8th IFAC/IFORS Symposium, Beijing, Peoples Republic of China, Aug 27-31 1988, Academia Sinica and Pergamon Press, 754-756
- Grabosch A, Heijmans HJAM (1991) Production, development and maturation of red blood cells; a mathematical model. In: Arino O, Axelrod D, Kimmel M (eds) Mathematical population dynamics. Lecture notes in pure and applied mathematics 131. Marcel Dekker, New York, 189-210
- Gregory WM, Richards MA, Slevin ML, Souhami RL (1991) A mathematical model relating response durations to amount of subclinical resistant disease. *Cancer Res* 51:1210-1216
- Hendriks PJ, Martens ACM, Ophorst-Van Marrewijk C, Visser JWM, Hagenbeek A (1993) Comparison of functional and nonfunctional assays for detection of minimal residual disease of genetically marked leukemic cells. *Exp Hematol* 21:1099
- Hittelman WN (1986) The technique of premature chromosome condensation to study the leukemic process: review and speculations. *CRC Crit Rev Oncol Hematol* 6:147-221
- Johnson RT, Rao PN (1970) Mammalian cell fusion; induction of premature chromosome condensation in interphase nuclei. *Nature* 226:717-722
- Jones TD, Morris MD, Young RW, Kehlet RA (1993) A cell-kinetics model for radiation-induced myelopoiesis. *Exp Hematol* 21:816-822
- Kasper GJ, Pieters R, Van Zantwijk CH, De Laat PA, De Waal FC, Van Wering ER, Veerman AJ (1991) In vitro drug sensitivity of normal peripheral blood lymphocytes and childhood

- leukemic cells from bone marrow and peripheral blood. *Br J Cancer* 64:469-474
- Kokenberg EANM (1991) Pharmacokinetic studies of daunorubicin in patients with acute myeloid leukemia. Thesis, Rotterdam: Erasmus University
- Lacombe F, Belloc F, Bernard P, Boisseau MR (1988) Evaluation of four methods of DNA distribution data analysis based on bromodeoxyuridine/ DNA bivariate data. *Cytometry* 9:245-253
- Lampariello F, Del Bino G (1988) Automatic parameter estimation of flow cytometric DNA distributions in the study of tumor cell kinetics. In: Han-Fu Chen (ed) Identification and parameter estimation, Preprints of the 8th IFAC/IFORS Symposium, Beijing, Peoples Republic of China, Aug 27-31 1988, Academia Sinica and Pergamon Press, 1767-1772
- Langermann HJ, Henzè G, Wulf M, Riehm H (1982) Abschätzung der Tumorzellmasse bei der akuten lymphoblastischen Leukämie im Kindesalter: prognostische Bedeutung und praktische Anwendung. *Klin Pädiat* 194:209-213
- Michelson S, Slate DL (1989) Emergence of the drug-resistant phenotype in tumor subpopulations: a hybrid model. *J Natl Cancer Inst* 81:1392-1401
- Nooter K, Sonneveld P, Deurloo J, Oostrum R, Schultz FW, Martens ACM, Hagenbeek A (1984) Repeated daunomycin administration in rats; pharmacokinetics and bone marrow toxicity. *Cancer Chemother Pharmacol* 12:187-189
- Nooter K, Sonneveld P, Martens ACM, Hagenbeek A, Schultz FW (1986) Tissue distribution and myelotoxicity of daunomycin in the rat; rapid bolus injection versus continuous infusion. *Eur J Cancer Clin Oncol* 22:801-806
- Nooter K, Oostrum R, Jonker R, Van Dekken H, Stokdijk W, Van den Engh G (1989) Effect of cyclosporin A on daunorubicin accumulation in multidrug resistant P388 leukemia cells measured by realtime flow cytometry. *Cancer Chemother Pharmacol* 23:296-300
- Nooter K, Sonneveld P, Oostrum R, Herweijer H, Hagenbeek A, Valerio D (1990) Overexpression of the *mdr1*-gene in blast cells from patients with acute myelocytic leukemia is associated with decreased anthracycline accumulation that can be restored by cyclosporin-A. *Int J Cancer* 45:263-268
- Ormerod MG, Payne AWR, Watson JV (1987) Improved program for the analysis of DNA histograms. *Cytometry* 8:637-641
- Pedrysz W (1989) Fuzzy control and fuzzy systems. Research Studies Press Ltd, England
- Pedrysz W (1985) Application of fuzzy relational equations for methods of reasoning in presence of fuzzy data. *Fuzzy sets and systems* 16:163-175
- Pieters R, Huisman DR, Loonen AH, Hählen K, Van der Does van den Berg A, Van Wering ER, Veerman AJ (1991) Relation of cellular drug resistance to long term clinical outcome in childhood acute lymphoblastic leukemia. *Lancet* 338:399-403
- Pieters R, Huisman DR, Loonen AH, Hählen K, Van der Does van den Berg A, Van Wering ER, Veerman AJ (1992a) Adenosine deaminase and purine nucleoside phosphosylase in childhood lymphoblastic leukemia; relation with differentiation stage, in vitro drug resistance and clinical prognosis. *Leukemia* 6:375-380
- Pieters R, Hongo T, Loonen AH, Huisman DR, Broxterman HJ, Hählen K, Veerman AJ (1992b) Different types of non-P-glycoprotein mediated multiple drug resistance in children with relapsed acute lymphoblastic leukemia. *Br J Cancer* 65:691-697
- Pieters R, Kasper GJ, Van Wering ER, Huisman DR, Loonen AH, Hählen K, Veerman AJ (1993) Cellular drug resistance profiles that might explain the prognostic value of immunophenotype and age in childhood acute lymphoblastic leukemia. *Leukemia* 7:392-397
- Pinkel D, Straune T, Gray JW (1986) Cytogenetic analysis using quantitative, high sensitivity, fluorescence hybridization. *Proc Natl Acad Sci USA* 83:2934-2938
- Potter MN, Cross NCP, Van Dongen JJM, Saglio G, Oakhill A, Bartram CR, Goldman JM (1993) Molecular evidence of minimal residual disease after treatment for leukaemia and

- lymphoma: an updated meeting report and review. *Leukemia* 7:1302-1314
- Raza A, Preisler HD (1985) Double labeling of human leukemic cells using  $^3\text{H}$ AraC and monoclonal antibody against BrdU. *Cancer Treatm Rep* 69:195-198
- Raza A, Maheshwari Y, Ucar K, Mayers G, Preisler HD (1987) Proliferative characteristics of acute nonlymphocytic leukemia in vivo. *Acta Haemat* 77:140-145
- Rowley JD (1973) A new consistent chromosomal abnormality in chronic myelogenous leukemia by quinacrine fluorescence and Giemsa staining. *Nature* 243:290-293
- Sandberg AA, Turc-Carel C (1987) The cytogenetics of solid tumors; relation to diagnosis, classification and pathology. *Cancer* 59:387-395
- Schmitz S, Franke H, Brusis J, Wichmann HE (1993) Quantification of the cell kinetic effects of G-CSF using a model of human granulopoiesis. *Exp Hematol* 21:755-760
- Secker-Walker LM (1990) Prognostic and biological importance of chromosome findings in ALL. *Cytogenet Cell Genet* 49:1-13
- Slate DL, Michelson S (1990) A mathematical model for the development and reversal of tumor cell resistance to chemotherapeutic agents. *Proc of the Am Assoc Cancer Res* 31:357
- Strand SE, Zanzico P, Johnson TK (1993) Pharmacokinetic modeling. *Med Phys* 20 (pt 2):515-527
- Sugeno M (1985) An introductory of fuzzy control. *Int J Infor Sci* 36:59-83
- Swan GW (1987) Tumor growth models and cancer chemotherapy. In: Thompson JR, Brown BW (eds) *Cancer modeling*. Marcel Dekker Inc, New York/Basel. pp 91-179
- Takagi T, Sugeno M (1985) Fuzzy identification of systems and its application to modelling and control. *IEEE Trans Syst Man Cyber SMC*-15:116-185
- Van Dekken H, Hagenbeek A, Bauman JGJ (1989) Detection of host cells following sex-mismatched bone marrow transplantation by fluorescent in situ hybridization with a Y-chromosome specific probe. *Leukemia* 3:724-728
- Van Dongen JJM, Breit TM, Adriaansen HJ, Beishuizen A, Hooijkaas H (1992) Detection of minimal residual disease in acute leukemia by immunological marker analysis and polymerase chain reaction. *Leukemia* 6,Suppl 1:47-59
- Van Dongen JJM, Breit TM, Adriaansen HJ, Beishuizen A, Hooijkaas H (1993) Immunophenotypic and immunogenotypic detection of minimal residual disease in acute lymphoblastic leukemia. *Recent Results Cancer Res* 131:157-183
- Van Nauta Lemke HR (1991) Fuzzy control of systems, a birdseye view. In: Vichnevetsky R, Miller JJH (eds) *Proc 13th IMACS World Congress on Computation and Applied Mathematics*, Trinity College, Dublin, Ireland, July 22-26 1991, 3:1197-1199
- Wang DF, Zhou CJ (1988) On the fuzzy modelling and control of complex systems. In: Han-Fu Chen (ed) *Identification and parameter estimation*, Preprints of the 8th IFAC/IFORS Symposium, Beijing, Peoples Republic of China, August 27-31 1988, Academia Sinica and Pergamon Press, 780-785
- Wichmann HE, Loeffler M, Schmitz S (1988) A concept of hemopoietic regulation and its biomathematical realization. *Blood Cells* 14:411-429
- Yan Y, Martens ACM, De Groot CJ, Hendrikk PJ, Valerio D, Van Bekkum DW, Hagenbeek A (1992) Retrovirus mediated transfer and expression of marker genes in the BN rat acute myelocytic leukemia model for the study of minimal residual disease. *Leukemia* 7:131-139
- Yanagisawa M, Dolbeare F, Todoroki T, Gray JW (1985) Cell cycle analysis using numerical simulation of bivariate DNA/ Bromodeoxyuridine distributions. *Cytometry* 6:550-562
- Yunis JJ (1983) The chromosomal basis of human neoplasia. *Science* 221:227-236
- Zadeh LA (1965) Fuzzy sets. *Information and Control* 8:338-353
- Zaider M, Minerbo GN (1993) A mathematical model for cell cycle progression under continuous low-dose-rate radiation. *Rad Res* 113:20-26



## Summary

In science it is common to use models to gain better insight into the mechanisms of action of complex systems and processes. Not only do they help understanding, they also enable prediction of the (future) behavior of systems, in response to internal and/or external changes. Such changes may (control) or may not (environmental influences) be imposed on purpose. In mathematical modeling a system is described in terms of mathematical formulae. This simplifies analysis by doing calculations as well as computer simulations.

System identification techniques can be used for model construction. Based on *à priori* knowledge and accepted theory, model equations are drafted that should describe the system. Such equations often are a set of algebraic or differential equations. Numerical values of parameters in these equations are estimated by an iterative algorithm, which makes adaptations until the solution of the model equations (=model response) agrees as much as possible with observations obtained from input/output measurements. If good agreement cannot be attained, it is a sign that the model does not adequately describe the system. Ergo, the description should be improved.

In the present thesis it is shown how in a multidisciplinary approach mathematical modeling can be applied to support leukemia treatment, with emphasis on chemotherapy. Several subproblems can be distinguished. To solve them is finding the best compromise among contradictory requirements. In other words, they form optimization problems.

With (induction) chemotherapy most leukemia patients enter complete remission. The number of leukemic cells drops below the clinical level of detection. In this state of "minimal residual disease" remaining cells may continue proliferation that results, eventually, in a relapse. How do cells regrow? How long and how intensively should maintenance chemotherapy be given to prevent relapse, without putting unnecessary strain—in view of the many adverse effects of the usual cytostatic drugs—on the patient?

These questions are immediately followed by other ones, like what dose-effect relations apply to cytostatic drugs? What are their pharmacokinetic properties? I.e., how is a drug distributed in and eliminated from the body. How (dose, route of administration) and how often should a drug be given to assure a sufficiently high concentration during a sufficiently long period to kill leukemic cells, but to stay within the drug tolerance levels of normal tissues?

To answer such questions data have been accumulated from laboratory experiments with the Brown Norway rat. A preclinical leukemia (BNML), transplantable in this species, was shown to be a relevant model for human acute myelocytic leukemia.

After an introduction in *Chapter 1*, where a few illustrative examples of model-

ing are given already, *in vivo* unperturbed growth of leukemia and regrowth after chemotherapy are discussed in Chapters 2 and 3.

In *Chapter 2* the BNML is investigated. In *Chapter 2.1* a few basic experiments are discussed that are important for the remainder of the thesis, as well as detection methods for leukemia.

In *Chapter 2.2* it is shown that unperturbed growth in the main target organs, after leukemia induction with  $10^7$  BNML cells *i.v.*, is characterized by an exponential phase (constant population doubling time) that is contiguously followed by a phase of Gompertzian growth (exponentially increasing doubling time). To the different organs (liver, spleen, bone marrow) slightly different growth curve parameter values apply. The growth curve of the total BNML cell population is obtained by summation of the curves of the mentioned organs at each time point.

Chemotherapy was given as a single *i.p.* dose of cyclophosphamide (100 mg/kg). The effect on the BNML cell population was described with several models. A pulse effect, i.e., an instantaneous reduction in population size, followed by regrowth similar to unperturbed growth, appeared to agree better with observed datapoints than models that describe a more gradual course of drug effect in the time.

Growth of the BNML cell population and the influence of two different courses of AMSA—a one-time and a fractioned administration of an equal dose of this drug—is evaluated in *Chapter 2.3*. From observed BNML cell survival data it appeared that the effect of an identical dose of AMSA varies with time. By means of computer simulations with a growth model it was verified whether BNML cells may be more sensitive during a certain phase of their cell cycle than during the other phases. Instantaneous cell kill by the drug was assumed, as well as regrowth identical with unperturbed growth. Although time-dependency of the drug dose effect resulted, the experimentally observed pattern could not be reproduced. A phase of extra drug sensitivity therefore could not be identified with the model, although some evidence exists in the literature that it must be the late S-phase.

The appendix shows that a growth curve can be described with two sets of parameters, i.e., the cell cycle time in combination with either a growth fraction and a cell loss factor, or the probabilities of cell death and cell doubling after one cycle time.

Development of drug resistance is a great problem in leukemia chemotherapy. *Chapter 2.4* is started with a discussion on drug resistance. The development of a BNML subcell line resistant against cyclophosphamide is reported. A simple model of growth of sensitive and resistant cell populations—assuming identical growth kinetics and mutation from sensitive to resistant—is used together with observed increase in lifespan of leukemic rats, treated at different times with different doses of cyclophosphamide. Another variable is the ratio of



resistant and sensitive BNML cells in the  $10^7$  cells used for leukemia induction. Curves are derived that show the relation between potential drug effect (if all BNML cells were sensitive) and net drug effect (decreased due to resistance development) as function of mutation rate, treatment time and amount of initially present resistant cells. Early treatment results in longer survival. This agrees with clinical experience. Without development of drug resistance, based on the log cell kill principle (a same dose always kills a same fraction of cells) and on the Gompertzian growth model, the opposite would be expected.

The model for leukemia development is applied to available clinical data in *Chapter 3*. The kinetics of childhood T-cell ALL (acute lymphoblastic leukemia) are analyzed in peripheral blood. Leukemic cells were detected during remission-induction and maintenance therapy by a sensitive double immunofluorescence method. Remarks on uncertainty are given in the appendix. One patient was extensively analyzed. She was monitored for 500 d, during which she relapsed twice after two remissions. The third remission induction course failed. The doubling time of the exponentially growing leukemic cells (under maintenance therapy) remained approximately constant (6.5 d). During remission induction the leukemic cells disappeared with a halftime of 2.8 d. Chemotherapy thus did not seem to change cell kinetics. An increasing resistance against the chemotherapy courses could be numerically expressed as a decrease in therapy efficacy.

In *Chapter 4* another important aspect of chemotherapy is investigated, i.e., distribution and elimination of the administered cytostatic drug. With multicompartment models of variable structure the pharmacokinetic processes of daunomycin (and its metabolite daunomycinol) in the rat are identified. Transfer rates of diffusion (the model parameters) are estimated according to the so-called maximum likelihood principle, based on concentration—time datapoints observed in various organs after a rapid *i.v.* daunomycin inoculation.

The system identification technique used is extensively discussed (with elucidations on elementary matrix manipulations and Laplace transformation in appendices). Three algorithms for minimizing the log(likelihood)-function are tested. They are a modified Gauss-Newton (gradient) method, a finite differences approach and a direct determination of the second order derivatives of the log(likelihood)-function with respect to the parameters. The first method appears to converge fast and well in comparison to the other ones. The maximum accuracy of the final parameter values is in the range of 5-10%.

First order kinetics (passive diffusion) appears to describe distribution and elimination of the parent drug reasonably well. Metabolism occurs in all tissues, including plasma, at different rates. For the metabolite the calculated concentration—time curves do not agree as well with the observations as for the parent drug, especially in the plasma. This indicates that model refinement is necessa-

ry, which may include nonlinearities.

Should system identification be possible with a minimum number of observations, it means that fewer laboratory animals need to be sacrificed. The influence of presence/absence of observations in compartments, as well as the number of observations in a compartment and time intervals between them, was evaluated with simulated observations. Reducing the number of observations to 25% makes the maximum accuracy of the estimated parameters decrease from 1% to 5-10%. Especially early observations might be left out, as long as all compartments are observed. If this is not the case, the maximum accuracy decreases rapidly to 30% or worse. It is better to have fewer observations in all compartments, than many in only few.

Limited accuracy of estimated parameters causes deliberately introduced small model structure errors to go unnoticed.

Analysis of DNA histograms obtained from a flow cytometer is the subject of *Chapter 5*. System identification is applied, in a way similar to Chapter 4, to a laboratory instrument that reveals the kinetics (proliferation) of cell populations. At a given time point the composition of a population with respect to the fractions of cells in  $G_1$ -, S- and  $G_2+M$ -phases of the cell cycle, can be determined from a flow cytometric histogram, which shows the number of cells as function of measured fluorescence (=DNA content) per cell. The wanted information cannot be read directly, but can be derived through modeling the histogram production process. Characteristic for the model that is proposed here, is that the S-phase in the histogram is composed of a number of harmonic functions (sines and cosines). This yields great flexibility in the allowed shape of the resultant for only few parameters to be estimated.

The method of analysis was tested on a series of simulated standard histograms obtained from literature. It performed well in comparison to other published analysis methods. This was also the case when analyzing populations with unconventional histograms, e.g., with relatively many cells in  $G_2+M$ -phase. From these results it may be expected that the method is suitable for analyzing histograms of cell populations treated with cell cycle phase specific drugs.

As an application two series of DNA histograms were analyzed that reflect the development of a B-lymphocyte cell population, with or without addition of a growth factor. Without growth factor many cells remain in  $G_1$ -phase (no proliferation); with growth factor more DNA synthesizing cells (S-phase) are detected, indicating proliferation. Histograms of normal and malignant cells in bone marrow, liver and spleen of the BN rat at several time points after inoculation with  $10^7$  BNML cells, were analyzed as well. With increase in time, a slight decrease in S-phase BNML cells was noticed. Gompertzian growth is caused by both decreased cell proliferation and increased cell loss.

Bone marrow transplantation (*Chapter 6*) offers perspectives with respect to increased cure rate in leukemia. It allows intensified chemotherapy, with greater chance to eradicate the leukemia, because limited drug tolerance in bone marrow (producer of normal blood cells) no longer forms an obstruction. After chemotherapy, destroyed bone marrow is replaced with donor bone marrow. *Chapter 6.1* is a short introduction to bone marrow transplantation.

In *Chapter 6.2* allogeneic bone marrow transplantation is discussed. The patient receives marrow from a healthy donor. A problem is the foreign tissue causing an immunological reaction against the patient (Graft versus Host Disease; GvHD). Three options are open for additional kill of residual leukemic cells in the patient. A) No special measures are taken—like removal of T-cells from the graft—to prevent GvHD. The profit is in the Graft versus Leukemia Reaction (GvLR) that is manifest alongside the GvHD. Complications due to GvHD must be accepted. If measures are taken to prevent GvHD, the GvLR is lost. Then, B) conditioning therapy (before transplantation) may be intensified; or C) after transplantation additional low dose chemotherapy may be prescribed in an attempt to prevent leukemia relapse.

From analysis of survival curves it can be derived that the GvLR should correspond with about 1 log leukemic cell kill (i.e., a factor of  $10^1=10$  reduction in the population of leukemic cells). It is shown that this explains the observed increase in cure probability from 40% to 90%. It is argued that 1 log cell kill also can be achieved with either one of the other two options. The advantage is that GvHD prophylaxis can be simultaneously applied. The cure rate thus increases just as much, but without having to face GvHD related complications.

The subject of *Chapter 6.3* is autologous bone marrow transplantation. Bone marrow is taken from the patient in remission, before intensive conditioning therapy aimed against minimal residual disease, cryopreserved and returned afterwards. An analysis is presented to answer the question to what extent relapse is to be attributed, on the one hand, to regrowth of leukemic cells left in the patient, and on the other hand, to the leukemic cells that probably were reinfused with the graft. Based on information obtained from the BNML it is argued that, as only a small fraction of *i.v.* inoculated leukemic cells actually starts proliferation (1-2%), and because cryopreservation of the graft strongly reduces the number of viable clonogenic cells (down to 0.1-1%), the contribution of the graft must be low (less than 10%). Therefore, more attention should be given to eradicating minimal residual disease in the patient, than to purging the graft from leukemic cells before transplantation.

In *Chapter 7* an overview is given on how the mathematical models, described in the previous chapters, as components fit into the overall modeling approach aimed at optimization of treatment of leukemia. Before real attention can be

given to the optimization part—eventually meant to adapt therapy to the needs of the individual patient—considerable effort must be put into the identification of the mechanisms of action of several complex biological systems. Among these are *in vivo* dose—effect relationships for various cytostatic drugs and various administration regimens, development of drug resistance, and pharmacokinetics in various stages of leukemia.

Several new developments are mentioned, among which mathematical models based on fuzzy logic, a simulation model for the improvement of the therapeutic index by optimally choosing the dose frequency (Z-method), and simulation models to visualize tumor growth. Better experimental methods for detection of leukemia, and the use of flow cytometry for cell kinetics of malignant cell populations treated with cytostatic drugs, are also discussed.

## Samenvatting

Op velerlei terrein wordt in de wetenschap gebruik gemaakt van modellen om een beter inzicht te verkrijgen in de werkingsmechanismen van complexe systemen en processen. Hiermee wordt niet alleen het begrip gediend, men komt ook tot het voorspellen van (toekomstig) gedrag van systemen in respons op in- en/of externe veranderingen. Deze worden al (besturing) dan niet (omgevingsinvloeden) opzettelijk aangebracht. Omschrijft men een systeem met wiskundige formules, zodat er gemakkelijker aan gerekend kan worden en het systeem op de computer kan worden gesimuleerd, dan spreekt men van mathematisch modelleren.

Voor het construeren van een model kan worden gebruik gemaakt van systeemidentificatietechnieken. Hierbij worden op basis van *à priori* kennis en aanvaarde theorie modelvergelijkingen opgesteld, die het systeem zouden moeten beschrijven. Meestal komt het neer op een stelsel algebraïsche, of ook wel differentiaalvergelijkingen. De numerieke waarde van parameters in deze vergelijkingen wordt geschat door ze zo in te stellen (dit gebeurt met een itererend algoritme), dat de als oplossing van de modelvergelijkingen berekende modelrespons zo goed mogelijk in overeenstemming wordt gebracht met uit input/output-metingen verkregen observaties. Kan geen goede overeenstemming worden bereikt, dan is dit een teken dat de modelvergelijkingen als systeembeschrijving niet voldoen, en aangepast moeten worden.

In dit proefschrift wordt getoond hoe, in een multidisciplinaire aanpak, mathematisch modelleren kan worden toegepast ter ondersteuning van de behandeling van leukemie, met chemotherapie in het bijzonder. Een aantal deelproblemen kan worden onderscheiden, waarvoor de oplossing steeds betekent het zoeken van het beste compromis tussen tegenstrijdige eisen. Met andere woorden, het oplossen van een optimaliseringsprobleem.

Door middel van inductie-chemotherapie kan een leukemiepatiënt meestal in remissie worden gebracht. Het aantal leukemiecellen is dan zodanig laag geworden, dat ze klinisch niet meer kunnen worden gedetecteerd. Vanuit deze toestand van "minimal residual disease" kunnen ze zich echter wel weer vermenigvuldigen, wat vroeg of laat uitmondt in een recidief. Hoe verloopt de teruggroei? Hoe lang en hoe intensief moet onderhoudstherapie worden gegeven om teruggroei te voorkomen, zonder de patiënt—gezien de ernstige bijwerkingen van de gebruikelijke cytostatica—odeloos te belasten?

Dit roept weer de vraag op naar dosis—effect relaties van cytostatica. En naar de farmacokinetiek van cytostatica. Hoe verspreidt een farmacon zich door het lichaam? Hoe (dosis en toedieningswijze) en in welke frequentie moet het worden toegediend om voldoende lang een voldoende hoge concentratie te bewerkstelligen om de leukemiecellen te doden, maar beneden het tolerantieniveau van normale weefselcellen te blijven?

Om dergelijke vragen te beantwoorden zijn gegevens verzameld uit proeven met de Brown Norway rat. Een experimentele leukemie (BNML) is transplanteerbaar in deze rassesoort, en vormt een relevant model voor humane acute myeloïde leukemie.

Na inleidende beschouwingen in *Hoofdstuk 1*, waarbij vooruitlopend op volgende hoofdstukken al enkele voorbeelden van modelleerresultaten worden gegeven, wordt in Hoofdstuk 2 en 3 de ongestoorde groei van leukemie en de teruggroei na chemotherapie *in vivo* behandeld.

In *Hoofdstuk 2* gaat het om de BNML. In *Hoofdstuk 2.1* worden enkele basisexperimenten vermeld, die voor de rest van het proefschrift van belang zijn. Onder andere wordt ingegaan op het detecteren van leukemische cellen.

In *Hoofdstuk 2.2* wordt getoond, dat ongestoorde groei, na leukemie-inductie met  $10^7$  BNML cellen *i.v.*, in de voornaamste doelorganen gekenmerkt wordt door een exponentiële fase (constante populatieverdubbelingstijd) die aansluitend overgaat in een fase van Gompertzgroei (exponentieel toenemende verdubbelingstijd). Voor de verschillende organen (lever, milt, beenmerg) gelden voor de groeicurve iets andere parameterwaarden. De curve voor de groei van de totale BNML celpopulatie wordt verkregen door de curven voor de genoemde organen op elk tijdstip te sommeren.

Chemotherapie werd gegeven als een enkele *i.p.* dosis cyclofosfamide (100 mg/kg). De uitwerking van de dosis op de celpopulatie werd beschreven met een aantal verschillende modellen. Een pulseffect, dat wil zeggen, een instantane reductie van de populatiegrootte, gevolgd door teruggroei analoog aan ongestoorde groei, bleek beter te conformeren aan de observaties dan de modellen die een meer geleidelijk effect beschrijven.

Groei van de BNML celpopulatie en invloed van twee verschillende AMSA-kuren, een enkelvoudige en een gefractioneerde toediening van een gelijke dosis van dit cytostaticum, wordt geëvalueerd in *Hoofdstuk 2.3*. Uit gemeten overleving van BNML cellen bleek variatie met de tijd in het effect van eenzelfde dosis AMSA op te treden. Met behulp van computersimulaties met een model voor leukemiegroei werd nagegaan of de BNML cellen in een bepaalde fase van hun celcyclus gevoeliger zouden zijn dan in andere fasen. Aangenomen werd, dat chemotherapie instantaan een fractie cellen doodt, waarna de rest doorgroeit op gelijke wijze als voor therapie. Hoewel een tijdsafhankelijkheid van het dosiseffect resulteerde, kon het experimenteel bepaalde patroon niet worden gereproduceerd. De extra gevoelige fase, blijkens de literatuur vermoed aan het eind van de S-fase, kon dus niet worden geïdentificeerd. In de appendix wordt aangetoond, dat een groeicurve kan worden beschreven met twee parameterstelsels, namelijk de celcyclustijd tezamen met óf een groeifractie en een celverliesfactor, óf een kans op sterfte en een kans op celverdubbeling na één cyclustijd.

Een groot probleem bij chemotherapie is dat leukemiecellen op den duur

resistent worden ten aanzien van de cytostatica. Het begin van *Hoofdstuk 2.4* bestaat uit een discussie over het ontstaan van resistentie en het ontwikkelen van een BNML subcellijn die resistent is tegen cyclofosfamide. Een simpel model voor groei van de gevoelige en resistente celpopulaties, waarbij gelijke groeikinetiek wordt verondersteld en mutatie van gevoelig naar resistent kan optreden, wordt gebruikt in samenhang met observaties aan levensduurverlenging van leukemische ratten, die op verschillende tijdstippen behandeld zijn met, groepsgewijs, verschillende doses cyclofosfamide. Een andere variabele hierbij is nog de verhouding gevoelig/resistent in de  $10^7$  BNML cellen, waarmee leukemie werd geïnduceerd. Curven worden afgeleid, die het verband aanduiden tussen potentieel dosiseffect (alle BNML cellen gevoelig) en netto dosiseffect (verminderd ten gevolge van resistentievorming), als functie van mutatiesnelheid, behandelingstijdstip en hoeveelheid initieel aanwezige resistente cellen. Vroege behandeling leidt tot langere overleving, wat overeenkomt met de klinische ervaring. Indien men geen resistentievorming aanneemt, dan moet, op grond van het 'log cell kill'-principe (eenzelfde cytostaticumdosis leidt altijd tot eenzelfde fractie gedode cellen) en het Gompertz-groei-model, juist het tegengestelde worden verwacht.

Het is te bezien, of het ontwikkelde model voor leukiegroei toepasbaar is op beschikbare klinische data. In *Hoofdstuk 3* wordt de kinetiek geanalyseerd van T-ALL (acute lymfatische leukemie) cellen in het perifere bloed van een aantal kinderen. Leukemische cellen werden gedetecteerd, gedurende remissie-inductie en onderhoudstherapie, met behulp van een gevoelige dubbele immunofluorescentietechniek. Op de onzekerheid in de observaties wordt in de appendix ingegaan. Eén patiëntje werd uitgebreid geanalyseerd. Ze werd gevolgd over een periode van 500 d, waarbij ze tweemaal een recidief ontwikkelde na evenzovele remissie-inducties. De derde inductiekuur mislukte. De verdubbelingstijd van de exponentieel groeiende populatie leukemiecellen (tijdens de onderhoudstherapiefasen) bleek steeds ongeveer gelijk te zijn, 6.5 d. Gedurende remissie-inductie nam de maligne celpopulatie af met een halveringstijd van 2.8 d. Celkinetische veranderingen door chemotherapie bleken dus niet op te treden. De toenemende resistentie tegen de chemotherapiekuuren kon getalsmatig tot uitdrukking worden gebracht in een daling van de therapie-effectiviteit.

In *Hoofdstuk 4* wordt een ander belangrijk aspect van chemotherapie behandeld, namelijk de distributie en het metabolisme van een toegediend cytostaticum. Met behulp van multicompartimentenmodellen met variabele structuur worden de farmacokinetische processen voor daunomycine (en metaboliet daunomycinol) in de rat geïdentificeerd. Hierbij worden diffusiesnelheden (de modelparameters) volgens het zogenoemde maximum likelihood principe geschat op basis van meetpunten in het concentratie—tijdsverloop in verschillende organen, volgend op een snelle *i. v.* injectie met daunomycine.

De gebruikte systeemidentificatietechniek wordt uitgebreid besproken (met in appendices een toelichting op elementaire matrixmanipulaties en Laplacetransformatie), evenals een aantal algoritmen om de log(likelihood)-functie te minimaliseren. Dit zijn een gemodificeerde Gauss-Newton (gradiënt) methode, een aanpak via eindige differenties en een directe bepaling van de tweede orde afgeleiden van de log(likelihood)-functie naar de parameters. De eerste methode bleek het best en snelst te convergeren. De uiteindelijke nauwkeurigheid van de geschatte parameterwaarden ligt maximaal in de orde van 5-10%.

Eerste orde kinetiek (passieve diffusieprocessen) blijkt de distributie en eliminatie van daunomycine redelijk goed te beschrijven. Metabolisme naar daunomycinol blijkt in alle organen, inclusief plasma, op te treden, maar in verschillend tempo. Echter, de berekende concentratie—tijdsverlopen komen voor de metaboliet slechter overeen met de meetgegevens dan voor daunomycine zelf, vooral in plasma. Dit duidt erop, dat nog modelverfijningen nodig zijn, waarbij wellicht niet-lineariteiten een rol moeten spelen.

Als systeemidentificatie mogelijk is met een minimum aantal observaties, betekent dit een besparing in het benodigde aantal proefdieren. De invloed van de aan-/afwezigheid van observaties in compartimenten, alsmede het aantal meetpunten in een compartiment en de verspreiding in de tijd, werd daarom geëvalueerd, met gesimuleerde observaties. Terugbrengen van het aantal observaties tot 25% doet de maximale nauwkeurigheid in de parameterschattingen teruglopen van 1% naar 5-10%. Met name vroege observaties konden worden gemist, zolang alle compartimenten geobserveerd blijven. Indien dit laatste niet meer het geval is, loopt de nauwkeurigheid snel terug naar 30% en meer. Het is daarom beter te beschikken over minder observaties in alle compartimenten, dan over vele in slechts enkele.

Door de beperkte nauwkeurigheid van de parameterschattingen dreigen opzettelijk in de modelstructuur aangebrachte kleine fouten niet te worden opgemerkt.

Analyse van DNA histogrammen, verkregen met een flowcytometer, is het onderwerp van *Hoofdstuk 5*. Hier wordt systeemidentificatie, analoog aan het in *Hoofdstuk 4* behandelde, toegepast op een laboratoriuminstrument waarmee inzicht kan worden verkregen in de kinetiek (proliferatie) van celpopulaties. Op zeker tijdstip kan namelijk de populatiesamenstelling, met betrekking tot de fractie cellen in de  $G_1$ -, S- of  $G_2+M$ -fase van de celcyclus, worden bepaald uit een door de flowcytometer geproduceerd histogram, waarin celaantal is uitgezet als functie van gemeten fluorescentie (=DNA inhoud) per cel. Uit zo'n histogram is de gewenste informatie niet rechtstreeks af te lezen, maar wel af te leiden via een model voor de totstandkoming van dat histogram. Het kenmerkende van het hier gebruikte model is, dat de S-fase in het histogram wordt samengesteld uit een aantal harmonische functies (sinussen en cosinussen), waarmee grote flexi-



biliteit in de vorm van de resultante wordt verkregen, terwijl toch slechts een beperkt aantal parameters hoeft te worden geschat.

De analysemethode werd getest met een serie gesimuleerde standaardhistogrammen uit de literatuur, en bleek in vergelijking met andere gepubliceerde DNA histogramanalysemethoden goed te voldoen. Ook bij het analyseren van histogrammen van uitzonderlijke celpopulaties, zoals bijvoorbeeld met relatief zeer veel cellen in  $G_2+M$ -fase. Uit deze resultaten mag worden afgeleid dat de methode onder andere geschikt zou zijn voor analyse van histogrammen van celpopulaties die met celcyclusfasespecifieke cytostatica behandeld zijn.

Als toepassingsvoorbeeld werden twee reeksen DNA histogrammen geanalyseerd, die de ontwikkeling van een populatie B-cel lymfocyten in de tijd reflecteren, wanneer wel of niet een groeifactor wordt toegevoegd. Zonder groeifactor blijven veel cellen in de  $G_1$ -fase (geen proliferatie); met groeifactor worden veel meer DNA synthetiserende cellen (S-fase) waargenomen. Histogrammen van normale en maligne cellen in het beenmerg, de lever en de milt van de BN rat, op verschillende tijdstippen na inoculatie met  $10^7$  BNML cellen, werden eveneens geanalyseerd. Een lichte daling van het aantal BNML cellen in S-fase werd in het tijdsverloop waargenomen. Gompertzgroei hangt samen met verminderde celproliferatie, naast verhoogd celverlies.

Beenmergtransplantatie (*Hoofdstuk 6*) biedt perspectieven ten aanzien van verhoogde genezingskans bij leukemie. Een behandeling met hogere doses wordt mogelijk, met grotere kans de populatie leukemiecellen te vernietigen, omdat geen rekening meer gehouden hoeft te worden met beperkte tolerantie van het beenmerg (de producent van onmisbare gezonde bloedcellen). Het beenmerg mag nu meevernietigd worden, omdat het via transplantatie na de behandeling wordt vervangen. In *Hoofdstuk 6.1* wordt een beknopte inleiding tot beenmergtransplantatie gegeven.

In *Hoofdstuk 6.2* komt allogene beenmergtransplantatie aan de orde. De patiënt krijgt dan beenmerg van een gezonde donor. Een probleem hierbij is, dat het lichaamsvreemde transplantaat afweersverschijnselen oproept gericht tegen de ontvanger (graft versus host disease, GvHD). Men kan kiezen uit drie opties voor een extra aanval op residuele leukemiecellen in de patiënt. A) Er worden geen speciale maatregelen getroffen—zoals verwijdering van T-cellen uit het transplantaat—om GvHD te voorkomen. Men profiteert van de dan eveneens optredende GvL-reactie (graft versus leukemia). De met GvHD samenhangende complicaties moet men dan accepteren. Indien wel wordt getracht GvHD te voorkomen, waarbij de GvL-reactie verloren gaat, dan kan men B) de conditioneringstherapie (voorafgaand aan beenmergtransplantatie) verzwaren, of C) na transplantatie extra chemotherapie met lage doses cytostatica toedienen in een poging een leukemierecidief te voorkomen.

Uit analyse van overlevingscurven kan worden afgeleid dat de GvL-reactie

ongeveer overeen moet komen met 1 log kill van leukemische cellen (dat wil zeggen, een reductie van de leukemiecelpopulatie met een factor  $10^1 = 10$ ). Het wordt aangetoond, dat hiermee de geobserveerde stijging van de genezingskans van 40% naar 90% kan worden verklaard. Bepleit wordt, dat 1 log cell kill ook met een van beide andere opties bereikt kan worden. Dit heeft het voordeel, dat tevens volledige GvHD profylaxe toegepast kan worden, zodat de genezingskans net zo sterk verhoogd wordt, maar nu zonder bijkomende complicaties van GvHD.

*Hoofdstuk 6.3* is toegespitst op autologe beenmergtransplantatie. Hierbij wordt beenmerg afgenomen bij de patiënt zelf, wanneer hij in remissie is, en gecryopreserveerd. Dan volgt de intensieve conditioneringstherapie, gericht tegen de residuele leukemie, waarna het beenmerg wordt gereïnfundeerd. Een analyse wordt gepresenteerd, gericht op het beantwoorden van de vraag, in welke mate aan een recidief wordt bijgedragen door enerzijds de toch nog in de patiënt overgebleven leukemiecellen en door anderzijds de waarschijnlijk in het transplantaat aanwezige leukemiecellen. Op basis van informatie verkregen uit de BNML wordt beargumenteerd dat, gezien het feit dat na *i.v.* inoculatie met leukemiecellen slechts een klein aantal werkelijk uitgroeit (1-2%), alsmede de waarneming dat cryopreservatie van het transplantaat het aantal levensvatbare leukemiecellen sterk reduceert (tot 0.1-1%), de bijdrage uit het transplantaat minder dan 10% zal zijn. Dientengevolge zou grotere aandacht moeten worden besteed aan het bestrijden van residuele ziekte in de patiënt, dan aan het verwijderen van leukemiecellen uit het transplantaat voorafgaand aan teruggave.

*Hoofdstuk 7* wordt begonnen met een overzicht van hoe de in voorgaande hoofdstukken beschreven mathematische modellen als componenten passen in een groter geheel, ressorterend onder een modelleringsaanpak die is gericht op het optimaliseren van de behandeling van leukemie. Voordat echt aandacht gegeven kan worden aan optimaliseren, uiteindelijk om therapie aan te passen aan de behoeften van de individuele patiënt, is nog aanzienlijke inspanning vereist voor de identificatie van de werkingsmechanismen van een aantal complexe biologische systemen, waaronder *in vivo* dosis—effect relaties voor verschillende cytostatica bij verschillende toedieningswijzen, resistentievorming en farmacokinetiek bij verschillende stadia van leukemie.

Een aantal nieuwe ontwikkelingen worden genoemd, onder andere mathematische modellen gebaseerd op fuzzy logic, simulatiemodellen ter verbetering van de therapeutische index door de behandelingsfrequentie optimaal te kiezen (Z-methode), en simulatiemodellen voor visualisering van tumorgroei, alsmede verfijnde methoden voor detectie van lage aantallen leukemiecellen, en het gebruik van flowcytometrie voor celkinetiek bij maligne, met chemotherapie behandelde celpopulaties.

## APPENDIX A

### Experimental Determination of the f-Factor, Relating the Number of Spleen Colonies Counted to the Number of Leukemic Cells Injected

Let a cell suspension consist of  $C$  cells, among which is an unknown amount,  $L$ , of leukemic cells with the capability to form colonies in the spleen (LCFU-S). A fraction,  $k \cdot C$  cells, is injected into a recipient rat. Upon examination of the spleen nineteen days later, let the number of spleen colonies counted be  $p$ . These colonies are due to  $p$  LCFU-S out of the  $k \cdot L$  inoculated ones (present in the  $k \cdot C$  cells inoculated) having reached the spleen. The dilution factor  $f$  is the ratio  $p/(k \cdot L)$ .

The remaining fraction  $(1-k) \cdot C$ , containing  $(1-k) \cdot L$  LCFU-S, are injected into another recipient rat. After a certain time its spleen is taken out and a cell suspension is made. This suspension must contain  $f \cdot (1-k) \cdot L$  LCFU-S, i.e., the  $f$ -factor—assumed to be a constant—times the number of LCFU-S inoculated.

The suspension is injected into a secondary recipient. The spleen of this recipient then should contain  $f \cdot f \cdot (1-k) \cdot L$  LCFU-S, or spleen colonies, nineteen days later. Let the number of colonies counted be  $q$ . Now, two equations are available,  $f \cdot f \cdot (1-k) \cdot L = q$  and  $f \cdot k \cdot L = p$ , from which the two unknown variables,  $f$  and  $L$ , can be found, e.g.,  $f = k \cdot q / ((1-k) \cdot p)$ .

## APPENDIX B

### Two Parameter Sets to Describe Cell Population Growth<sup>1</sup>

#### B.1 A Simple Model for Population Growth

A cell population grows through proliferation, i.e., each newborn cell goes through a process of maturation till it divides and produces two daughter cells (Fig. 1.6). During maturation four phases can be distinguished, which together make up the cell's cycle. In the  $G_1$ -phase the cell synthesizes a multitude of products vital to its existence and functioning. In the S-phase duplication of the cell's DNA (containing the genetic information to be passed on) occurs. The  $G_2$ -phase is similar to  $G_1$ . During the fourth phase, M, processes leading to the actual cell division take place.

In general, cells—even when they are of the same type—will not complete the cycle in a fixed time; especially the variation in  $G_1$  duration may be large. The cycle time is distributed continuously between a minimum and a maximum that may be so large that, apparently, the cell is in a resting (non-proliferating) state.

In a simple model this distribution of cell cycle times is approximated with a system of classes. A cell of class  $i$  ( $i=1,2,..$ ) needs  $i$  times an average cycle time ( $T_c$ ) before it divides.

In a proliferating cell population there will in general any time be more young cells (in earlier phases of the cycle) than old ones (far advanced in the cycle). Therefore, in the model the average cycle time  $T_c$  is divided into a number of age (or maturity) classes and the cells of the population are distributed in the age classes at random according to an exponential distribution. In this way the probability of putting a cell in the first age class is twice that of putting it in the last one (Fig. 1.7A).

Having distributed the cells of a population of initial size  $C(0)=C_0$  in the age classes, the growth of the population is simulated as follows. After each time step the cells in an age class are transferred to the next one. If they were in the last age class two selections are made. First, a cell is considered for being lost from the population (e.g., natural cell death), which event has been assigned the probability  $p_2$ . If the cell is lost it "disappears"; if not, it will either split

---

<sup>1</sup>reproduced from Schultz FW, Hagenbeek A (1991) Simulation studies on the regrowth of acute myeloid leukemia after autologous bone marrow transplantation. In: Arino O, Axelrod D, Kimmel M (eds). Proceedings of the Second International Conference on Mathematical Population Dynamics, Rutgers University, New Brunswick, NJ, 17-20 May 1989. Marcel Dekker Inc, New York, pp 689-709

into two cells (probability  $p_1$ ) or it remains unchanged. In both cases the resulting cell(s) go(es) to the first age class (Fig. 1.7B).

When the (initial) population is small, its growth can be monitored by following each individual cell, whose fate (death, division) can be determined by using a random number generator. When the population grows large such a procedure is too time consuming. Fortunately, for large numbers frequencies will tend to their probabilities. So, from say 5000 cells, the population size after time step  $n$  can be derived from the size after time step  $n-1$  (Fig. 1.7C). As  $p_2 \cdot C_{n-1}$  cells die,  $p_1 \cdot (1-p_2) \cdot C_{n-1}$  cells double, and  $(1-p_1) \cdot (1-p_2) \cdot C_{n-1}$  cells remain unchanged, the increase in  $C$  is given by  $C_n - C_{n-1} = p_1 \cdot (1-p_2) \cdot C_{n-1} - p_2 \cdot C_{n-1}$ , hence:

$$C_n = C_{n-1} \cdot (1 + p_1 + p_2 + p_1 \cdot p_2). \quad (B.1)$$

The probabilities  $p_1$  and  $p_2$  may of course be functions of time or population size.

## B.2 Relations between $p_1$ , $p_2$ and $T_c$ , GF and $\Phi$

To describe tumor growth kinetics several parameters and their interrelationships have been introduced before [Steel, 1977; see references, section 2.5]. For a clear understanding a few of the derivations are repeated below. Very generally, population growth during a time step  $\Delta t$  can be described with:

$$C(t+\Delta t) = C(t) + \Delta t \cdot \frac{dC(t)}{dt}, \quad C(0) = C_0. \quad (B.2)$$

If the fractions of cells being born and dying during time step  $\Delta t$  are proportional to the population size  $C(t)$  (birth rate  $B$ , death rate  $D$ ), the growth rate becomes:

$$\frac{dC(t)}{dt} = (B-D) \cdot C(t) = K \cdot C(t), \quad (B.3)$$

from which follows, should  $B$  and  $D$  remain constant, that the population would double its size in time:

$$T_d = \frac{\ln(2)}{K}. \quad (B.4)$$

The minimum value of the doubling-time is the (average) cell cycle time, meaning that all cells would proliferate without cell loss and divide after  $T_c$ . In

other words, the growth fraction, GF, then equals 1; the average number of daughter cells per cell per time length  $T_c$ ,  $A$ , equals 2; and the cell loss factor,  $\Phi$ , equals zero. The maximum or potential production rate then is:

$$K_{pot} = \frac{\ln(2)}{T_c}. \quad (B.5)$$

If on average  $A$  daughter cells per cell per time length  $T_c$  are produced ( $1 \leq A \leq 2$ ) because many cells need more than time  $T_c$  to complete their cycle, then the growth fraction (the fraction of the population that produces 2 daughter cells per cell per time length  $T_c$ ) is  $GF = A - 1$  ( $0 \leq GF \leq 1$ ). The production rate then is:

$$K_{prod} = \frac{\ln(A)}{T_c}, \quad (B.6)$$

and the cell loss factor:

$$\Phi = 1 - \frac{K}{K_{prod}}; \quad (0 \leq \Phi \leq 1). \quad (B.7)$$

The defined GF thus reflects the difference between  $K_{pot}$  and  $K_{prod}$  (potential and actual production), while  $\Phi$  is a measure of the difference between  $K$  (the resultant of production and loss) and  $K_{prod}$ .

When, at maximum, the doubling-time tends to infinity ( $T_d \rightarrow \infty$ ) a stationary phase of constant population size is found, which is—mathematically—undetermined. GF may be zero while no cell loss occurs (all cells resting); GF may be one (maximum) while maximum cell loss occurs (in time span  $T_c$  all cells produce two daughter cells of which one is lost); or GF and  $\Phi$  may have intermediate values within their ranges, which, however, are strictly related.

If  $K$  can be considered a constant during the  $n^{\text{th}}$  time span  $T_c$ , then Eq.(B.2) for population growth can be solved as:

$$C_n = C_{n-1} \cdot \exp \left[ \frac{T_c \cdot \ln(2)}{T_d} \right]. \quad (B.8)$$

Comparison with Eq.(B.1) yields the relation between the probabilities  $p_1, p_2$  on the one hand and the cell cycle variables on the other hand:

$$p_1 = -1 + \exp\left[\frac{T_c \cdot \ln(2)}{T_d}\right] / (1 - p_2). \quad (\text{B.9})$$

In steady state,  $T_d \rightarrow \infty$ , Eq.(B.9) reduces to:

$$p_1 = \frac{p_2}{(1 - p_2)}. \quad (\text{B.10})$$

The cell cycle parameters can be expressed in terms of  $p_1$ ,  $p_2$  and  $T_c$ . From Eq.(B.9) it is derived that

$$T_d = T_c \cdot \ln(2) / \ln(1 + p_1 - p_2 - p_1 \cdot p_2). \quad (\text{B.11})$$

With Eq.(B.4) it follows that

$$K = \ln(1 + p_1 - p_2 - p_1 \cdot p_2) / T_c, \quad (\text{B.12})$$

which in combination with Eq.(B.6) yields

$$K_{prod} = \ln(1 + p_1) / T_c \quad (\text{B.13})$$

and

$$A = 1 + p_1, \quad \text{so,} \quad GF = p_1, \quad (\text{B.14})$$

for  $K$  equals  $K_{prod}$  if no cell loss occurs ( $p_2=0$ ).

The cell loss factor then can be written (Eq.(B.7), Eq.(B.12) and Eq.(B.13)) as

$$\Phi = 1 - \frac{\ln(1 + p_1 - p_2 - p_1 \cdot p_2)}{\ln(1 + p_1)}. \quad (\text{B.15})$$

The other way around,  $p_1$  and  $p_2$  can be expressed in  $GF$  and  $\Phi$ :

$$p_1 = GF \quad \text{and} \quad p_2 = 1 - (1 + GF)^{-\Phi}. \quad (\text{B.16})$$

So, either  $p_1$  and  $p_2$  or  $GF$  and  $\Phi$  determine, together with  $T_c$ , the dynamic behavior of the cell population.

## APPENDIX C

### Detection of T-ALL Cells using Immunofluorescence (IF) Technique; Theoretical Considerations about the Uncertainty in the Observations

Malignant T-ALL cells that possibly are present in a sample of patient material, e.g., peripheral blood (PB) or bone marrow (BM), can be marked by certain fluorescent dyes. This enables detection by fluorescence microscopy as well as—after some corrections—quantification of the numbers of malignant cells. Each dye is bound to a monoclonal antibody (McAb; this explains the term immunofluorescence technique). One McAb is able to bind T-lymphocytes specifically. Another one attaches to a receptor that is characteristic for the malignant lymphocyte. Thus, double stained cells are marked as malignant T-ALL cells (further denoted as L-cells; in contrast to normal cells, N-cells).

In several ALL patients—children of various ages and both sexes—the quantity of L-cells in PB and/or BM was determined using the mentioned IF-technique. This happened at various time points during remission-induction (RI) chemotherapy and during the follow-up phase (when the patient in remission was treated with a maintenance therapy (MT) regimen). The samples were always taken just before the start of a course or the administration of a drug dose. The datapoints yielded interesting information about the course with time of the magnitude of the L-cell population.

#### *Poisson Statistics*

With the following Poisson formula it is possible to compute the probability,  $Pr$ , that there are a certain number of L-cells ( $X$ , say  $X = K$ ) in a sample of  $U$  cells (where  $U$  is large), given the condition that the chance that a cell is leukemic,  $p$ , is very small.

$$Pr\{X=K\} = \exp(-M) \cdot \frac{M^K}{K!}, \quad (C.1)$$

where  $M = p \cdot U$  and  $K!$  ( $K$  factorial) =  $1 \times 2 \times 3 \times \dots \times K$  ( $0! = 1$ ,  $1! = 1$  by definition).

From the Poisson distribution a mean value,  $\mu$ , and a standard deviation,  $\sigma$ , are derived:

$$\mu = M \quad \text{and} \quad \sigma = M^{1/2}. \quad (C.2)$$

For certain values of  $U$ ,  $p$  and  $K$  the corresponding Poisson probabilities can be found in Table C-1.



TABLE C-1 POISSON PROBABILITY,  $P_R\{X=K\}$  (%), OF FINDING K L-CELLS IN A SAMPLE OF U CELLS FOR VARIOUS L-CELL FREQUENCIES,  $F_T$

| U (cells)<br>(mm <sup>3</sup> ) | 3000        |           |           | 30,000    |           |           |
|---------------------------------|-------------|-----------|-----------|-----------|-----------|-----------|
|                                 | $F_T$       | $p$       |           | $F_T$     | $p$       |           |
|                                 | $1:10^3$    | $1:10^4$  | $1:10^5$  | $1:10^3$  | $1:10^4$  | $1:10^5$  |
|                                 | $10^{-3}$   | $10^{-4}$ | $10^{-5}$ | $10^{-3}$ | $10^{-4}$ | $10^{-5}$ |
| K                               | $Pr\{X=K\}$ |           |           |           |           |           |
| 0                               | <1          | 5         | 74        | 5         | 74        | 7         |
| 1                               | <1          | 15        | 22        | 15        | 22        | 3         |
| 2                               | <1          | 22        | 3         | 22        | 3         | 1         |
| 3                               | <1          | 22        | <1        | 22        | <1        | 1         |
| 4                               | <1          | 17        | <1        | 17        | <1        | 1         |
| $\mu$ (mm <sup>-3</sup> )       | 3           | 0.3       | 0.0       | 0.3       | 0.03      | 0.003     |
| $\sigma$ "                      | 2           | 0.5       | 0.0       | 0.5       | 0.17      | 0.05      |
| $\mu$ (sample <sup>-1</sup> )   | 30          | 3         | 0.3       | 3         | 0.3       | 0.03      |
| $\sigma$ "                      | 6           | 2         | 0.5       | 2         | 0.5       | 0.17      |

( $\mu$ ,  $\sigma$ : mean value of X, respectively, standard deviation)

It is assumed that a PB sample contains 3000 cells per mm<sup>3</sup>. Sample sizes of 1 or 10 mm<sup>3</sup> are considered, which thus contain  $U = 3000$  or  $U = 30,000$  cells. Various true frequencies of L-cells,  $F_T$ , are considered; they are directly related to the probability,  $p$ , that a cell in the sample is an L-cell. The corresponding mean values of the L-cells in the samples, and their standard deviation, can then be calculated from Eq.(C.2).

For instance, take a 1 mm<sup>3</sup> (3000 cells) sample. If 1 in 10,000 cells is an L-cell, then  $p = 10^{-5}$  and the chance that 0 L-cells will be found in the sample is 74%. The chance that 1 L-cell or 2 L-cells are found are 22% and 3%, respectively. The chance that more than 2 L-cells are found is only 1%.

If many such samples are regarded, then a mean value of  $0.3 \pm 0.5$  L-cells per sample will be observed.

Note that it is assumed that all L-cells that are present in the sample—if any—will be actually detected. Thus, the limitation in detecting leukemia in a patient when L-cell frequencies are very low is only a matter of insufficiently large sample sizes.

It is a different story should an L-cell that is present in the sample go unnoticed, e.g., due to instrumental inaccuracies, fluctuations in immunofluorescence intensities, etc. In that case it still holds in the above example that there is a 22% chance of having 1 L-cell in the sample. The chance of actually observing that L-cell may be, e.g., 7% only, if in reality—apparently—on average two in three stained cells go unnoticed.

TABLE C-2 POISSON PROBABILITY,  $PR\{X=K\}$  (%), OF FINDING K L-CELLS IN SAMPLES THAT CONTAIN ON AVERAGE M L-CELLS

| $\mu =$ | 10          | 1     | 0.1   | 0.01  |
|---------|-------------|-------|-------|-------|
| K =     | $Pr\{x=K\}$ |       |       |       |
| 0       | <0.01       | 37    | 90    | 99    |
| 1       | 0.05        | 37    | 9     | 0.99  |
| 2       | 0.2         | 18    | 0.5   | <0.01 |
| 5       | 3.6         | <0.01 | <0.01 |       |
| 8       | 11.3        |       |       |       |
| 9       | 12.5        |       |       |       |
| 10      | 12.5        |       |       |       |
| 11      | 11.4        |       |       |       |
| 12      | 9.5         |       |       |       |
| 15      | 3.5         |       |       |       |
| 20      | 0.2         |       |       |       |

Inversely, suppose that the true frequency of L-cells amounts to:

1 in  $10 \text{ mm}^3$  ( $0.1/\text{mm}^3$ ); 1 in  $1 \text{ mm}^3$  ( $1/\text{mm}^3$ ); 1 in  $0.1 \text{ mm}^3$  ( $10/\text{mm}^3$ );  
in other words,  $p =$

$1/30,000$ ;  $1/3000$ ;  $1/300$ ;

then on average the number of L-cells observed in samples of U cells:

U = 3,000:  $\mu = 0.1$ ; 1; 10;

30,000: 1; 10; 100.

The probability of observing K L-cells in such samples is given in Table C-2.

Example: Take a sample of  $U = 3000$  cells ( $1 \text{ mm}^3$ ); if the number of L-cells on average present in such a sample is  $\mu = 0.1/\text{mm}^3$  ( $= 0.1/\text{sample}$ ), then  $K = 0$  L-cells will be observed in 9 out of 10 cases. Letting the sample size increase to  $U = 30,000$  cells (so,  $10 \text{ mm}^3$  and  $\mu = 1$  L-cell/sample) then  $K = 0$  is observed still in 4 out of 10 cases.

The other way round, if  $K = 0$  is the result of measuring  $U = 3000$  cells, then it is very well possible that in the patient there are on average  $\mu = 0.1$  L-cell per  $\text{mm}^3$  (this will be so in 9 out of 10 cases). It is even possible that there are on average  $\mu = 1$  L-cell present per  $\text{mm}^3$  (4 against 6). It would be very unlikely that  $\mu = 10$  L-cells per  $\text{mm}^3$  are present (probability less than 0.01%).

N.B., it is assumed that only the sample size is limiting for the detection probability.

## APPENDIX D

### A Few Elementary Matrix Manipulations

#### Definitions.

An  $N \times M$  matrix,  $A$ , consists of *elements*,  $a_{ij}$ , arranged in  $N$  rows and  $M$  columns as follows:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1M} \\ a_{21} & a_{22} & \dots & a_{2M} \\ \cdot & \cdot & \cdot & \cdot \\ a_{N1} & a_{N2} & \dots & a_{NM} \end{pmatrix}.$$

So, the element  $a_{ij}$  (a so-called *scalar*, i.e., a real or complex number) is to be found in the position where the  $i^{\text{th}}$  row crosses the  $j^{\text{th}}$  column.

In the *transposed* matrix of  $A$ ,  $A^T$ , the columns and rows are interchanged. Therefore,  $A^T$  is an  $M \times N$  matrix:

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{N1} \\ a_{12} & a_{22} & \dots & a_{N2} \\ \cdot & \cdot & \cdot & \cdot \\ a_{1N} & a_{2N} & \dots & a_{NM} \end{pmatrix}.$$

There are various types of matrices. If the numbers of rows and columns are equal,  $N = M$ , the matrix is called *square*. If a square matrix has  $a_{ij} = a_{ji}$ , then  $A^T = A$  and  $A$  is called *symmetrical*. A *diagonal*

matrix is square and has all elements  $a_{ij} = 0$ , except its elements,  $a_{ii}$ , on the diagonal from top left to bottom right. It also is written as  $diag(a_{11}, a_{22}, \dots, a_{NN})$ . If, in this case, all diagonal elements are equal to one,  $a_{ii} = 1$ , the matrix is called an  $N$ -dimensional *identity* matrix,  $A = I$  or  $A = I_N$ . If  $A$  consists of 1 column only ( $N \times 1$  matrix) it is called a  $N$ -dimensional *vector*, e.g.,

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ x_N \end{pmatrix},$$

also written as  $col(x_1, x_2, \dots, x_N)$ . Its transpose, the row vector  $\underline{x}^T = [x_1 \ x_2 \ \dots \ x_N]$  is a  $1 \times N$  matrix.

#### Operations.

Two matrices  $A$  and  $B$  of the same dimension can be *added* or *subtracted* by performing such an operation element by element.

They can be *multiplied* if  $A$  is an  $N \times M$  and  $B$  is an  $M \times P$  matrix. The result,  $C = A \cdot B$  is an  $N \times P$  matrix, e.g.,

Thus, the element  $c_{ij}$  is obtained by multiplying in sequence the  $i^{\text{th}}$  row elements of  $A$  and the  $j^{\text{th}}$  column elements of  $B$ , and adding the results.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11} \cdot b_{11} + a_{12} \cdot b_{21} + a_{13} \cdot b_{31} & a_{11} \cdot b_{12} + a_{12} \cdot b_{22} + a_{13} \cdot b_{32} \\ a_{21} \cdot b_{11} + a_{22} \cdot b_{21} + a_{23} \cdot b_{31} & a_{21} \cdot b_{12} + a_{22} \cdot b_{22} + a_{23} \cdot b_{32} \end{pmatrix}.$$

Multiplication of matrix A by a scalar, p, means that every element  $a_{ij}$  is multiplied by p.

### Determinant.

If a matrix A is square, its *determinant*  $|A|$  or  $\det(A)$  can be defined. For a 2x2 matrix:

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}.$$

For a 3x3 matrix:

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \cdot \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \cdot \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \cdot \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

Higher order determinants can be found by repeating the basic routine of decomposition into lower order determinants.

Determinants play a role in solving sets of simultaneous algebraic equations (see Cramer's rule; Table 4-3) and in matrix inversion. The *inverted* matrix,  $A^{-1}$ , of the square matrix A is that matrix, which after the multiplication  $A^{-1} \cdot A$  will yield an identity matrix:  $A^{-1} \cdot A = A \cdot A^{-1} = I$ .

For instance, an unknown matrix B may be solved from  $A \cdot B = C$  by multiplying both sides by  $A^{-1}$ :  $A^{-1} \cdot A \cdot B = A^{-1} \cdot C \Rightarrow I \cdot B = A^{-1} \cdot C \Rightarrow B = A^{-1} \cdot C$ .

### Inversion.

To calculate the inverse of A, define the *minor* of element  $a_{ij}$  as the determinant  $|E_{ij}|$ , where  $E_{ij}$  is the matrix A with its *i*th row and *j*th column deleted. Define

the *cofactor*,  $c_{ij}$ , of  $a_{ij}$  as its minor multiplied by, depending on the values of  $i$  and  $j$ , the factor of  $+1$  or  $-1$ :  $c_{ij} = (-1)^{i+j} |E_{ij}|$ . Put all cofactors into a matrix,  $C$ , the transpose of which is called the *adjugate* of  $A$ , or *adj A*. Then,  $A^{-1} = \text{adj } A / |A| = C^T / |A|$ .

Differentiation.

Let  $s$  be a scalar,  $M$  be a  $P \times Q$  matrix,  $\underline{V}$  a  $K$ -dimensional vector and  $\underline{W}$  a  $L$ -dimensional vector. Following the same convention as Schweppe [1973], the derivative of the matrix with respect to the scalar,  $\partial A(s)/\partial s$ , is obtained by taking all  $P \times Q$  derivatives  $\partial a_{ij}(s)/\partial s$ . Consequently, the derivative of the vector with respect to the scalar,  $\partial \underline{V}(s)/\partial s$ , is the vector with  $K$  elements  $\partial v_i(s)/\partial s$ .

The derivative of the scalar with respect to the vector,  $\partial s(\underline{V})/\partial \underline{V}$ , is the  $1 \times K$  gradient vector (row vector by definition!):  $[\partial s/\partial v_1, \dots, \partial s/\partial v_K]$ .

The derivative of the scalar with respect to the matrix,  $\partial s(A)/\partial A$ , is the  $Q \times P$  gradient matrix:

$$\begin{bmatrix} \partial s/\partial a_{11} & \dots & \partial s/\partial a_{p1} \\ \dots & \dots & \dots \\ \partial s/\partial a_{1Q} & \dots & \partial s/\partial a_{pQ} \end{bmatrix}$$

Finally, the derivative of the vector  $\underline{V}$  with respect to the vector  $\underline{W}$ ,

$\partial \underline{V}(\underline{W})/\partial \underline{W}$ , is the  $K \times L$  matrix:

$$\begin{bmatrix} \partial v_1/\partial w_1 & \dots & \partial v_1/\partial w_L \\ \dots & \dots & \dots \\ \partial v_K/\partial w_1 & \dots & \partial v_K/\partial w_L \end{bmatrix}$$

The second derivative, for instance of the scalar  $s$  with respect to the

vector  $\underline{V}$  becomes, as  $\partial^2 s/\partial \underline{V}^2 = \partial \{\partial s/\partial \underline{V}\}/\partial \underline{V}$ , the  $K \times K$  matrix:

$$\begin{bmatrix} \partial^2 s/\partial v_1^2 & \dots & \partial^2 s/\partial v_k \partial v_1 \\ \dots & \dots & \dots \\ \partial^2 s/\partial v_1 \partial v_k & \dots & \partial^2 s/\partial v_k^2 \end{bmatrix}$$

Reference:

-Schweppe FC (1973) Uncertain dynamic systems. Prentice Hall Inc., Englewood Cliffs, NJ

## APPENDIX E

### Laplace Transformation and Analytical Solution of the Model Equations

#### Laplace Transformation.

By definition the Laplace transform (LT) of the function  $F(t)$ ,  $\mathcal{L}\{F(t)\}$ , is given by:

$$\mathcal{L}\{F(t)\} = F(s) = \int_0^{\infty} F(t) \cdot e^{-s \cdot t} dt.$$

Backward transformation yields:  $\mathcal{L}^{-1}\{F(s)\} = F(t)$ .

Application of the LT to the function  $F(t) = K \cdot e^{-M \cdot t}$ , for instance, yields:

$$F(s) = \mathcal{L}\{K \cdot e^{-M \cdot t}\} = \int_0^{\infty} K \cdot e^{-M \cdot t} dt = K \cdot \frac{1}{(s+M)} \cdot e^{-(s+M) \cdot t} \Big|_0^{\infty} = \frac{K}{s+M}.$$

The other way around, the inverse LT of  $F(s) = K/(s+M)$  is  $F(t) = K \cdot e^{-M \cdot t}$ . For many functions  $F(t)$  the tabulated LTs, vice versa, can be found in the literature (e.g., [Godfrey, 1983]).

A useful property of LT is that it converts a differential equation into an algebraic equation, which often is easier to solve. E.g., using partial integration:

$$\begin{aligned} \mathcal{L}\left\{\frac{dF(t)}{dt}\right\} &= \int_0^{\infty} \frac{dF(t)}{dt} \cdot e^{-s \cdot t} dt = \\ &F(t) \cdot e^{-s \cdot t} \Big|_0^{\infty} + \int_0^{\infty} s \cdot F(t) \cdot e^{-s \cdot t} dt = -F(0) + s \cdot \mathcal{L}\{F(t)\}. \end{aligned}$$

Thus, the equation:

$$\frac{dF(t)}{dt} = -K \cdot F(t), \quad F(0) = M$$

can be solved through LT:

$$\mathcal{L}\left\{\frac{dF(t)}{dt}\right\} = \mathcal{L}\{-K \cdot F(t)\} \Rightarrow -M + s \cdot F(s) = -K \cdot F(s) \Rightarrow (s+K) \cdot F(s) = M \Rightarrow F(s) = \frac{M}{s+K} \Rightarrow F(t) = \mathcal{L}^{-1}\{F(s)\} = M \cdot e^{-Kt}.$$

Analytical Solution of State Equations of a 2x3 Compartment Model.

The state equations for the model as shown in Fig. 4.7, in terms of amounts of drug as function of time,  $Q(t)$ , are:

$$\begin{pmatrix} \dot{Q}_1(t) \\ \dot{Q}_2(t) \\ \dot{Q}_3(t) \\ \dot{Q}_4(t) \\ \dot{Q}_5(t) \\ \dot{Q}_6(t) \end{pmatrix} = \begin{pmatrix} -(p_1+p_3+p_8) & p_2 & 0 & 0 & 0 & 0 \\ p_1 & -(p_2+p_7) & 0 & 0 & 0 & 0 \\ p_3 & 0 & 0 & 0 & 0 & 0 \\ p_8 & 0 & 0 & -(p_4+p_6) & p_5 & 0 \\ 0 & p_7 & 0 & p_4 & -p_5 & 0 \\ 0 & 0 & 0 & p_6 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} Q_1(t) \\ Q_2(t) \\ Q_3(t) \\ Q_4(t) \\ Q_5(t) \\ Q_6(t) \end{pmatrix},$$

by applying the rules of mass conservation, and of matrix-vector notation as explained in Appendix D. The initial condition is the vector  $Q(0) = \text{col}(D,0,0,0,0,0)$ , where D is the total *i. v.* drug dose.

These six state equations contain six unknown variables, assuming that the transfer rate constants,  $p_1$  through  $p_8$ , have been given a value. They can be split into two sets of three equations each, i.e., a set for the parent drug ( $Q_1$  through  $Q_3$ ) and a set for the metabolite. The former does not depend on the latter, because metabolism is a one-way process. The former set, therefore, can be solved separately first. After substituting the solution in the second set, the three remaining unknown variables can be solved too.

A convenient way of solving both sets is by Laplace transformation. After some rearrangement the following equations are derived:

$$\begin{pmatrix} s+p_1+p_3+p_8 & -p_2 & 0 \\ -p_1 & s+p_2+p_7 & 0 \\ -p_3 & 0 & s \end{pmatrix} \cdot \begin{pmatrix} Q_1(s) \\ Q_2(s) \\ Q_3(s) \end{pmatrix} = \begin{pmatrix} D \\ 0 \\ 0 \end{pmatrix}, \text{ respectively,}$$

$$\begin{bmatrix} s+p_4+p_6 & -p_5 & 0 \\ -p_1 & s+p_5 & 0 \\ -p_3 & 0 & s \end{bmatrix} \cdot \begin{bmatrix} Q_4(s) \\ Q_5(s) \\ Q_6(s) \end{bmatrix} = \begin{bmatrix} p_8 \cdot Q_1(s) \\ p_7 \cdot Q_2(s) \\ 0 \end{bmatrix}.$$

As said, from these sets  $Q_1(s)$  through  $Q_3(s)$  are solved first. Then,  $Q_4(s)$  through  $Q_6(s)$  are solved. Next, inverse LT yields  $Q_1(t)$  through  $Q_6(t)$ . This results in:

$$Q_1(t) = D \cdot \left[ \frac{p_2+p_7-a}{b-a} \cdot e^{-a \cdot t} - \frac{p_2+p_7-b}{b-a} \cdot e^{-b \cdot t} \right],$$

$$Q_2(t) = D \cdot \frac{p_1}{b-a} \cdot (e^{-a \cdot t} - e^{-b \cdot t}) \quad \text{and}$$

$$Q_3(t) = D \cdot p_3 \cdot \left[ \frac{p_2+p_7}{a \cdot b} - \frac{a-(p_2+p_7)}{a \cdot (a-b)} \cdot e^{-a \cdot t} + \frac{b-(p_2+p_7)}{b \cdot (a-b)} \cdot e^{-b \cdot t} \right],$$

where  $a = \frac{1}{2} \cdot \{ (p_1+p_3+p_8+p_2+p_7) + [(p_1+p_3+p_8+p_2+p_7)^2 - 4 \cdot (p_2p_7+p_3p_2+p_3p_7+p_8p_2+p_8p_7)]^{1/2} \}$   
and  $b = \frac{1}{2} \cdot \{ (p_1+p_3+p_8+p_2+p_7) - [(p_1+p_3+p_8+p_2+p_7)^2 - 4 \cdot (p_2p_7+p_3p_2+p_3p_7+p_8p_2+p_8p_7)]^{1/2} \}$

Subsequently,

$$Q_4(t) = P_1 \cdot e^{-a \cdot t} + P_2 \cdot e^{-b \cdot t} + P_3 \cdot e^{-c \cdot t} + P_4 \cdot e^{-d \cdot t},$$

where the  $P_s$  follow from:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ b+c+d & a+c+d & a+b+d & a+b+c \\ bc+bd+cd & ac+ad+cd & ab+ad+bd & ab+ac+bc \\ bcd & acd & abd & abc \end{bmatrix} \cdot \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{bmatrix} = \begin{bmatrix} 0 \\ K_1 \\ K_2 \\ K_3 \end{bmatrix},$$

while

$$K_1 = D \cdot p_8, \quad K_2 = D \cdot p_8 \cdot (p_2+p_7+p_5) \quad \text{and} \quad K_3 = D \cdot p_5 \cdot \{ p_8 \cdot (p_2+p_7) + p_2 \cdot p_7 \},$$

$$c = \frac{1}{2} \cdot \{ (p_4+p_6+p_5) + [(p_4+p_6+p_5)^2 - 4 \cdot p_6 \cdot p_5]^{1/2} \},$$

$$d = \frac{1}{2} \cdot \{ (p_4+p_6+p_5) - [(p_4+p_6+p_5)^2 - 4 \cdot p_6 \cdot p_5]^{1/2} \}.$$



In a similar way,

$$Q_5(t) = P_1 \cdot e^{-a \cdot t} + P_2 \cdot e^{-b \cdot t} + P_3 \cdot e^{-c \cdot t} + P_4 \cdot e^{-d \cdot t},$$

but here the Ps follow from the same set of formulae, except that the vector  $\text{col}(0, K_1, K_2, K_3)$  must be replaced by the vector  $\text{col}(0, 0, K_1, K_2)$ , in which now

$$K_1 = D \cdot (p_7 \cdot p_1 + p_8 \cdot p_4) \text{ and } K_2 = D \cdot \{p_7 \cdot p_1 \cdot (p_4 + p_6) + p_8 \cdot p_4 \cdot (p_2 + p_7)\}.$$

Finally,

$$Q_6(t) = D - Q_1(t) - Q_2(t) - Q_3(t) - Q_4(t) - Q_5(t).$$

#### Parameter estimation.

Thus, the time courses of the amounts of DAU and DOL in the 6 compartments can be calculated for certain parameter values,  $\mathbf{p}$ . If compartment volumes,  $\underline{V}$ , are known, the concentrations can be calculated as well. The most appropriate values for the parameters must be estimated from observed data. Observed compartments are: plasma DAU and DOL and excretion DAU and DOL. A sequential, 3-step, method can be used.

1) Using a nonlinear least squares routine (commonly used routines can be found in e.g., [Bevington, 1969; Press et al., 1985], fit a bi-exponential curve to the plasma DAU data:

$$C_1(t) = A \cdot e^{-\alpha \cdot t} + B \cdot e^{-\beta \cdot t}.$$

This will yield the best values for A, B,  $\alpha$  and  $\beta$ , while for the compartment's volume,  $V_1 = D/C_1(0) = D/(A+B)$  is found. Also,  $a = \alpha$ ,  $b = \beta$  and  $(p_2 + p_7) = (A \cdot \beta + B \cdot \alpha)/(A + B)$  are found.

2) With the values from step 1, fit the curve

$$Q_3(t) = D \cdot p_3 \cdot \left[ \frac{p_2 + p_7}{a \cdot b} - \frac{a - (p_2 + p_7)}{a \cdot (a - b)} \cdot e^{-a \cdot t} + \frac{b - (p_2 + p_7)}{b \cdot (a - b)} \cdot e^{-b \cdot t} \right]$$

to the excretion DAU data; this will yield the best value for the only unknown parameter,  $p_3$ .

3) Now, assuming that  $V_4 = V_1$ , fit a tetra-exponential curve

$$C_4(t) = \{P_1 \cdot e^{-a \cdot t} + P_2 \cdot e^{-b \cdot t} + P_3 \cdot e^{-c \cdot t} + P_4 \cdot e^{-d \cdot t}\} / V_4$$

to the plasma DOL data. This yields values for  $P_1$  through  $P_4$ , and for  $c$  and  $d$ . Substituting these values in the above equations for  $P_s$ ,  $a$ ,  $b$ ,  $c$ ,  $d$  and  $K_s$ , the latter ( $K_1$  through  $K_3$ ) can be solved. In turn, these variables will yield the parameters  $p_8$ ,  $p_4$ ,  $p_6$  and  $p_5$ .

Next,  $p_1 = a + b - \{(p_3 + p_8) + (p_2 + p_7)\}$  and  $p_2 = - \{ab - (p_2 + p_7) \cdot (p_1 + p_3 + p_8)\} / p_1$  and  $p_7 = (p_2 + p_7) - p_2$  can be calculated.

With best values of all transfer rates now known, the remaining time course,  $Q_2(t)$ ,  $Q_5(t)$  and  $Q_6(t)$  can be calculated as well. As excretion DOL observations are available, the calculated  $Q_6(t)$  can be checked for goodness of fit.

It is important to note that this way of estimating the 9 parameters (8 transfer rate constants and 1 volume) requires at least 5 observations in the plasma DAU compartment, 2 observations in the excretion DAU compartment, and 7 observations in the plasma DOL compartment.

References:

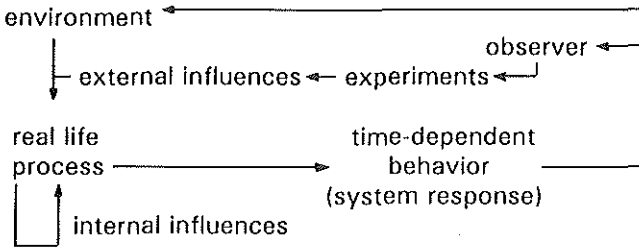
- Bevington PhR (1969) Data reduction and error analysis. McGraw Hill, New York
- Godfrey K (1983) Compartmental models and their application. Academic Press, London
- Press WH, Flannery BP, Teukolsky SA and Vetterling WT (1985) Numerical recipes: the art of scientific computing. Cambridge University Press, New Rochelle, NY

## APPENDIX F

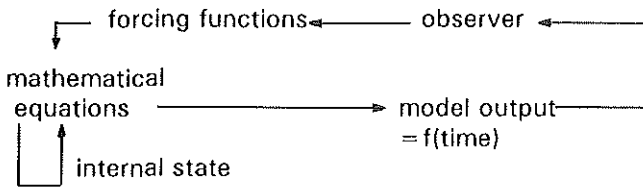
### System Identification Approach and Beyond

The main concepts of system identification are briefly summarized below. Distinguish I) the dynamical system or process to be studied, and II) a mathematical model that—to some extent—describes this system.

#### I physical/biological system:



#### II mathematical model:



dynamical system: the system shows a time-dependent behavior, due to external perturbations or to internal transitions

system identification: to find the best (correct) model for the description of the system (behavior), based on input/output measurements

parameter estimation: to find the best values of characteristic constants or variables in the current model equations

the procedure: based on *a priori* knowledge and/or reasonable assumptions a hypothetical mechanism of action of the real life process is stated and "translated" into mathematical equations; in- and external influences are transformed into mathematical forcing functions. The model output is calculated, by running a computer simulation program that solves the model equations, with preliminary estimates for the model parameters substituted. This model output is compared to the system behavior that has been experimentally observed under comparable circumstances (the observer has the freedom to set these circumstan-

ces in his experiments). If necessary, the parameter values are changed (automatically by an optimization program) to yield a better fit (e.g., in the least squares sense) of calculated model output to measured system response. If the "best" fit, thus obtained, is not satisfactory, then another model (based on an alternative hypothesis) must be stated and tested. This is repeated until a good match has been achieved.

If two hypotheses explain the experimental observations equally well, then new experiments must be designed and conducted to find the more probable of the two. In this respect, by exposing areas of insufficient knowledge, modeling shows the direction into which experimental research must go.

If the system has been properly identified, and it is described by the correct model, then by performing computer simulations with the latter the dependence of the model response on the forcing functions can be easily evaluated. This means that, at this stage, except for confirmation purposes no actual experiments (with sacrifice of laboratory animals) need to be conducted, but that the (future) behavior of the real system under analogous circumstances can be predicted in advance. In this way, by carrying out simulations it is possible to learn how to set the forcing functions so as to obtain the desired model response. In other words, for the present case, how to adapt the chemotherapy schedule so as to arrive at the desired clinical response.

advantages of the SYSTEM IDENTIFICATION approach: once a model has been built simulation results will quickly become available; it will mean fewer costs with respect to laboratory work and animal lives if interaction between laboratory work and theoretical developments is maintained, both before (model is tested for explaining a given observed system response) and after (system has been identified, model is used to predict future system behavior) the proper model has been fully developed; there is flexibility in model design and test set-ups.

disadvantages: due to scarce data and system complexity biological/physical knowledge is often insufficient, or existing knowledge is scattered and cannot be managed to serve as a strong basis for making *à priori* reasonable assumptions and stating sound hypotheses; numerical problems, yielding inaccurate estimates, may be encountered during the solving of the model equations and the running of the optimization programs when large and ill-defined models are used. However, it has to be mentioned here that a purely experimental approach without modeling will suffer from the same disadvantages.

## Abbreviations and Symbols

|              |  |
|--------------|--|
| $\mathbf{0}$ | null vector  |
| <b>A</b>     |  |
| A            | adenosine; age distribution; amount of drug; constant, or Gaussian white noise; cytostatic agent; fluorescence intensity; Fourier coefficient; Gompertz' retardation constant; intercept of regression line; mean number of daughter cells per cell per cycle time; rate constant; start position rectangle; system matrix |
| A50          | amount of drug, required for 50% effect  |
| ABMT         | autologous bone marrow transplantation   |
| ADR          | adriamycin   |
| AIC          | Akaike's information criterion (Eq.4.54)   |
| ALDH         | aldehyde dehydrogenase   |
| ALL          | acute lymphoblastic leukemia   |
| AML          | acute myelocytic leukemia  |
| AMSA         | acridinyl anisidide  |
| AOS          | advanced operating system  |
| ARA-C        | arabinoside cytosine   |
| AUC          | area under the curve   |
| a            | constant; integer; slope; width of rectangle   |
| $a_r$        | rate of mutation from drug resistant to drug sensitive phenotype   |
| $a_s$        | rate of mutation from drug sensitive to drug resistant phenotype   |
| ara-C        | arabinoside cytosine   |
| <b>B</b>     |  |
| B            | birth rate; cytostatic agent; end position rectangle; fluorescence intensity; Fourier coefficient; input matrix; mean number of daughter cells per cell per cycle time; slope of regression line   |
| BB           | biology based  |
| $B_j$        | system matrix differentiated with respect to jth parameter   |
| BM           | bone marrow  |
| BMT          | bone marrow transplantation  |
| BN           | Brown Norway   |
| BN/Bi/Rij    | Brown Norway rat, raised in the Rijswijk colony  |
| BNML         | Brown Norway rat acute myelocytic leukemia   |
| BrdUrd       | bromodeoxyuridine  |
| b            | birth rate; blood; constant; integer; intercept  |
| <b>C</b>     |  |
| C            | concentration; cytosine; number of cells; unit of DNA contents   |
| $C_0$        | initial concentration; initial number of cells   |
| $C_1$        | control matrix   |
| $C_2$        | distribution matrix  |

|                   |  |
|-------------------|--|
| $C_{max}$         | maximum cell population size                                 |
| $C_n$             | number of cells after n time steps                           |
| $C^+$             | number of cells just after treatment                         |
| $C^-$             | number of cells just before treatment                        |
| CBS               | Centraal Bureau voor de Statistiek                           |
| CC                | committed cell   |
| CD15              | CD15 membrane antigen  |
| CD5               | CD5 membrane antigen   |
| CF                | intercept of line  |
| CFA               | cyclophosphamide   |
| CFA/R             | cyclophosphamide resistant cell line                         |
| CFA/S             | cyclophosphamide sensitive cell line                         |
| CFU-C             | granulocytic progenitor cell                                 |
| CFU(-S)           | colony forming unit(-spleen); pluripotential progenitor cell |
| CLL               | chronic lymphoblastic leukemia                               |
| CML               | chronic myelocytic leukemia                                  |
| CMT               | continuous maintenance therapy                               |
| CPU               | central processing unit                                      |
| CR                | complete remission   |
| CSC               | committed stem cell  |
| CT                | conditioning therapy   |
| CV                | coefficient of variation                                     |
| Cy                | cyclophosphamide   |
| CyA               | cyclosporin A  |
| c                 | constant   |
| calc              | calculated   |
| cmpt              | compartment  |
| cor               | correlation coefficient                                      |
| <b>D</b>          |  |
| D                 | dose; death rate; growth inhibition                          |
| $D_z$             | residual process error (= observation - prediction)          |
| $\underline{D}_z$ | random variable measurement error                            |
| DAU               | daunomycin   |
| DDH               | direct determination of Hessian matrix                       |
| DMBA              | dimethyl-benzanthracene                                      |
| DOL               | daunomycinol   |
| DOX               | doxorubicin  |
| DSF               | disulfiram   |
| d                 | death rate   |
| $d_r$             | death rate for drug resistant cells                          |
| $d_s$             | death rate for drug sensitive cells                          |
| dev               | observed value - calculated value                            |

|                                |   |
|--------------------------------|---|
| E                              | effect; Fisher information matrix   |
| EC                             | error criterion (Eq.4.55)   |
| ED <sub>50</sub>               | number of inoculated leukemic cells yielding disease in 50% of the recipients |
| EG                             | contiguous exponential and Gompertz curves                                    |
| EGT                            | contiguous exponential and Gompertz curves with time delay                    |
| E.coli                         | Escherichia coli bacteria   |
| Ep                             | expected value  |
| Eq                             | equation  |
| e                              | break-off criterion for parameter step  |
| e <sub>i</sub>                 | eigenvalue  |
| eff                            | effective   |
| eps                            | very small number   |
| F                              |   |
| F                              | constant; fluorescence  |
| F(p)                           | function of parameters  |
| F <sub>1</sub>                 | fraction of G <sub>1</sub> cells  |
| F <sub>2</sub>                 | fraction of G <sub>2</sub> +M cells   |
| F <sub>S</sub>                 | fraction of S cells   |
| F <sub>t</sub>                 | true frequency of leukemic cells  |
| FACS                           | fluorescence activated cell sorter  |
| FCM                            | flow cytometer (-metry)   |
| FD                             | finite differences  |
| FISH                           | fluorescent in situ hybridization   |
| FITC                           | fluorescein isothiocyanate  |
| f                              | dilution factor; female; fraction of drug sensitive cells; weighting factor   |
| f <sub>i</sub>                 | fraction of patients with leukemic cell load of 10 <sup>i</sup>               |
| G                              |   |
| G                              | Gaussian curve; Gompertz curve; guanosine                                     |
| G <sub>1</sub> ,G <sub>2</sub> | phases of the cell cycle when cells contain 2C,4C DNA                         |
| GF                             | growth fraction   |
| GF <sub>r</sub>                | constant growth fraction  |
| GN                             | Gauss-Newton algorithm  |
| GOF                            | goodness of fit   |
| GRAD(F)                        | gradient of function F  |
| GSH                            | glutathione   |
| GST                            | glutathione dependent enzyme  |
| Gr                             | gradual drug effect   |
| GvHD                           | graft versus host disease   |
| GvL(R)                         | graft versus leukemia (reaction)  |
| g                              | numerical constant  |

|            |  |
|------------|--|
| gdt        | gradient   |
| gr         | granulocytes   |
| grad(F)    | gradient of function F   |
| H          |  |
| H          | constant; Hessian matrix; observation matrix   |
| $H_0$      | initial therapy level  |
| HGF        | (biochemical) growth factor  |
| HPLC       | high pressure liquid chromatography  |
| HSC        | hemopoietic stem cell  |
| h          | multiharmonic function; proportionality constant   |
| I          |  |
| I          | electric current; identity matrix; input   |
| IBMT       | isologous bone marrow transplantation  |
| IMSL       | International Mathematical & Statistical Libraries Inc, software firm  |
| IF         | immunofluorescence   |
| IFA        | ifosfamide   |
| ILS        | increase in lifespan   |
| Ig         | immunoglobulin   |
| In         | instantaneous drug effect  |
| i          | channel number; integer  |
| i.p.       | intraperitoneal  |
| i.v.       | intravenous  |
| J          |  |
| J(p)       | function of parameters   |
| $J_0$      | environmental drug concentration   |
| j          | channel number; integer  |
| j'         | channel number   |
| jp         | channel number   |
| K          |  |
| K          | actual number of leukemic cells in sample; actual number of leukemic cells surviving (conditioning) therapy; difference between cell birth and loss rate; kilobyte; (proportionality) constant; scaling constant |
| $K_{j,i}$  | metabolic rate constant, from compartment i to compartment j   |
| $K_{pot}$  | potential production rate  |
| $K_{prod}$ | cell production rate   |
| $K1_{j,i}$ | Michaelis-Menten constant  |
| $K2_{j,i}$ | Michaelis-Menten constant  |
| k          | constant; difference between cell birth and loss rate; integer; numerical constant   |
| $k_1$      | constant; rate of drug uptake  |
| $k_2$      | constant; rate of drug efflux  |



|           |  |
|-----------|--|
| $k_3$     | constant; drug transfer rate   |
| $k_4$     | rate of drug inactivation/clearance  |
| $k_e$     | growth fraction in exponential phase   |
| $k_{i,j}$ | transfer rate constant, to compartment i, from compartment j   |
| $k_{j,i}$ | transfer rate constant, from compartment i to compartment j  |
| L         |  |
| L         | constant; likelihood function; (number of) leukemic cells; therapy level   |
| LCFU-S    | leukemic colony forming unit spleen  |
| LCK       | log cell kill  |
| LF        | likelihood function  |
| LLF       | log of likelihood function   |
| lac-Z     | <i>Escherichia coli</i> $\beta$ -galactosidase gene  |
| lin       | linear   |
| ly+bl     | lymphocytes + blast cells  |
| M         |  |
| M         | constant; ED <sub>50</sub> value; leukemic cell burden of patient in remission; metabolite; mitosis (phase of the cell cycle; number of cells); number of datapoints; total number of channels |
| MCA       | monoclonal antibody  |
| MDR       | multidrug resistance   |
| MFA       | mafosfamide  |
| MGN       | modified Gauss-Newton  |
| MH        | multiharmonic  |
| ML        | maximum likelihood   |
| MLE       | maximum likelihood estimation  |
| MNC       | mononuclear cell   |
| MRD       | minimal residual disease   |
| MST       | mean survival time   |
| MT        | maintenance therapy  |
| MTT       | 3-4,5-dimethylthiazol-2yl-2,5-diphenyltetrazoliumbromide (assay)   |
| MTX       | methotrexate   |
| McAb      | monoclonal antibody  |
| MdST      | median survival time   |
| MeCCNU    | methyl-cyclohexyl-chlorethyl nitrosourea   |
| MoAb      | monoclonal antibody  |
| m         | male; mutation rate; number of model parameters  |
| m'        | reduced number of parameter values   |
| $m_1$     | G <sub>1</sub> channel   |
| $m_2$     | G <sub>2</sub> +M channel  |
| m-AMSA    | amsacrine  |

|                   |   |
|-------------------|---|
| N                 | Gaussian probability density function; normal cell; number of cells; number of inoculated cells; number of observation times; number of spikes, rectangles or harmonic functions              |
| $N_{\text{calc}}$ | calculated number of cells  |
| $N_g$             | number of cells at transition from exponential to Gompertz growth   |
| $N_{\text{max}}$  | maximum population size   |
| $N_{\text{obs}}$  | observed number of cells  |
| NC                | total number of cells   |
| NR                | number of observations used in parameter estimation   |
| NT                | non-target compartment  |
| n                 | number of cells, compartments, rats   |
| neo <sup>R</sup>  | neomycin resistance gene  |
| nr                | number  |
|                   | O   |
| O                 | output  |
| $O(\Delta p^3)$   | rest term of order of magnitude of 3d power of the parameter step   |
| obs               | observation, observed   |
|                   | P   |
| P                 | parent drug; polynomial   |
| $P_r$             | probability function  |
| PB                | peripheral blood  |
| PBS               | phosphate buffered saline   |
| PC                | personal computer; slope  |
| PCC               | premature chromosome condensation   |
| PCR               | polymerase chain reaction   |
| PD                | pharmacodynamic   |
| PF                | performance index (Eq.4.24)   |
| PI                | propidium iodide  |
| PK                | pharmacokinetic   |
| PSC               | pluripotent stem cell   |
| PTT               | post-transplantation therapy  |
| Pr                | probability of  |
| p                 | birth probability; chance that cell in sample is leukemic; chance that cell survives treatment; number of parameters; number of spleen colonies; polynomial function; potential log cell kill |
| $p_1$             | probability of cell division  |
| $p_2$             | probability of cell death   |
| $p_{CT}$          | chance that cell survives conditioning therapy  |
| $p_s$             | chance that cell survives additional therapy  |
| $p_i, p'_i$       | ith parameter   |
| $p_m$             | parameter value for minimum function value  |
| $\mathbf{p}$      | parameter vector  |

|          |  |
|----------|--|
| $D_0$    | initial parameter vector   |
| $D_{ML}$ | vector containing maximum likelihood parameter values  |
| pdf      | probability density function   |
| p.o.     | per os (orally)  |
| Q        |  |
| Q        | variance(covariance matrix) of the measurement error(s)                                      |
| q        | net log cell kill; number of spleen colonies   |
| q24h     | 24 h time interval   |
| R        |  |
| R        | decay rate; number of drug resistant cells; resistance; variance of residuals                |
| $R_0$    | initial number of drug resistant cells   |
| $R^+$    | number of drug resistant cells just after treatment  |
| $R^-$    | number of drug resistant cells just before treatment   |
| $R_i$    | ith remission (induction) period; number of cells resistant against drug i                   |
| RI(T)    | remission induction (therapy)  |
| RK       | Runge-Kutta  |
| RM124    | monoclonal antibody for BNML cell detection  |
| RNG      | random number generator  |
| r        | correlation coefficient  |
| $r_j$    | uniformly redistributed cells from tail of Gaussian curve                                    |
| rpe      | residual process errors  |
| S        |  |
| S        | number of drug sensitive cells; phase of the cell cycle when cells synthesize DNA; stem cell |
| $S_0$    | initial number of drug sensitive cells   |
| $S^+$    | number of drug sensitive cells just after treatment  |
| $S^-$    | number of drug sensitive cells just before treatment   |
| S        | sensitivity vector   |
| SCA      | spleen colony assay  |
| SD       | standard deviation   |
| SE       | standard error   |
| SF       | surviving fraction   |
| SSQ      | sum of squared errors  |
| SSR      | sum of squared residuals (Eq.4.53)   |
| $SSR_i$  | sum of squared residuals in compartment i  |
| SVD      | singular value decomposition   |
| s        | value of standard deviation  |
| s.c.     | subcutaneous   |
| sd       | standard deviation   |
| seq      | sequential curve fitting method  |

|                   |  |
|-------------------|--|
| sim               | simultaneous solution method   |
| <b>T</b>          |  |
| T                 | eigenvector matrix; halftime or doubling time; period; target compartment; thymidine; time; treatment time |
| T'                | reduced eigenvector matrix   |
| T(N)              | time till extinction for normal cells  |
| T(T)              | time till extinction for tumor cells   |
| $T_{1/2}$         | halftime   |
| $T_2$             | doubling time  |
| $T_c$             | cell cycle time  |
| $T_d$             | doubling time  |
| $T_{e5}$          | time of reaching a level of $10^5$ leukemic cells  |
| $T_h$             | halftime   |
| $T_s$             | drug sensitive period of the cell cycle  |
| $T_t$             | treatment time   |
| $\underline{T}_i$ | eigenvector  |
| T-ALL             | T cell acute lymphoblastic leukemia  |
| TBI               | total body irradiation   |
| TCC               | total correlation coefficient (Eq.4.51, Eq.5.43)   |
| $TCC_t$           | total correlation coefficient, all compartments (Eq.4.52)  |
| TM                | transition matrix (matrices)   |
| TNC               | total nucleated cells  |
| TPC               | theoretical probability of cure  |
| TRITC             | tetramethylrhodamine-isothiocyanate  |
| TdT               | terminal deoxynucleotidyl transferase  |
| t                 | DNA contents per cell; tissue; (survival) time   |
| $t_0$             | starting time  |
| $t_e$             | doubling time in exponential growth phase  |
| $t_f$             | constant   |
| $t_g$             | time of transition from exponential to Gompertz growth   |
| $t_k$             | observation time k   |
| $t_t$             | treatment time   |
| $t^+$             | time point just after treatment  |
| $t^-$             | time point just before treatment   |
| <b>U</b>          |  |
| U                 | magnitude of log cell kill; number of cells in sample  |
| $\underline{U}$   | input vector   |
| u                 | drug input rate  |
| $\underline{u}$   | input vector   |
| <b>V</b>          |  |
| V                 | magnitude of log cell kill; voltage  |
| $\underline{V}$   | volume vector, noise vector  |

|                                  |  |
|----------------------------------|--|
| $V_i$                            | volume of compartment $i$  |
| $V_p$                            | variance-covariance matrix of estimated parameter values   |
| $V_x$                            | variance-covariance matrix of the state vector   |
| $V_{ZP}$                         | variance-covariance matrix of predicted histogram  |
| $\underline{y}$                  | measurement noise vector   |
| var                              | variance   |
| $\text{var}_i$                   | variance of residuals in compartment $i$   |
|                                  | W  |
| W                                | weight factor  |
| WBC                              | white blood cell count   |
| w                                | frequency; number of observed compartments   |
|                                  | X  |
| X                                | number of $ED_{50}$ units; number of leukemic cells in sample; number of leukemic cells surviving (conditioning + additional) therapy; percentage of CD15 positive cells per number of mononucleated cells |
| $X'$                             | number of leukemic cells surviving conditioning therapy  |
| x                                | absolute value of the logarithm of the ratio of true and estimated parameter values  |
| $x_i$                            | drug concentration in compartment $i$  |
| $\underline{x}$                  | state vector   |
|                                  | Y  |
| Y                                | eigenvalue matrix; percentage of myeloid cells per total number of leukocytes  |
| $Y'$                             | reduced eigenvalue matrix  |
| $y_j(t, \underline{p})$          | model response in compartment $j$ for parameter values in $\underline{p}$ at time $t$  |
| $y_{mj}(t_k)$                    | observation in compartment $j$ at time $t_k$   |
| $\underline{y}$                  | system response vector   |
| $\underline{y}_m$                | observation vector   |
|                                  | Z  |
| Z                                | percentage of T-ALL cells per number of mononucleated cells; periodical function; treatment efficacy   |
| $Z_M$                            | observed histogram   |
| $Z_P$                            | predicted histogram  |
| $Z_T$                            | theoretical histogram  |
| $Z_t$                            | true histogram   |
| $\underline{z}$                  | vector containing parameters of metabolism   |
|                                  | Greek  |
| $\Delta\text{Pr}\{\text{cure}\}$ | increase in cure probability   |
| $\underline{\Delta p}$           | parameter step vector  |
| $\Delta t$                       | time interval; time step   |

|                    |  |
|--------------------|--|
| $\Delta_j Z_p[jp]$ | predicted number of cells in channel jp that have arrived from channel j |
| $\delta$           | cut-off criterion; drug pulse duration; unit pulse                       |
| $\Theta$           | matrix for parameter step calculation                                    |
| $\mu$              | mean value   |
| $\sigma$           | standard deviation   |
| $\mu_X$            | mean value of X  |
| $\sigma_X$         | standard deviation of X  |
| $\tau$             | time delay   |
| $\Phi$             | cell loss factor; transition matrix                                      |
| $\Psi_j$           | transition matrix  |
| $\omega$           | drug free interval   |

## Naschrift

Het is terecht de gewoonte een proefschrift af te sluiten met een woord van dank. Immers, al prijkt slechts de naam van één enkele auteur op het omslag, aan het tot stand komen van het 'boekje' hebben velen—in mindere of meerdere mate belangeloos—een bijdrage geleverd. Al was het maar met het scheppen van een geschikte werksfeer, met een blijk van belangstelling of het geven van een aanmoediging. Want naast directe hulp en medewerking is, herhaaldelijk gegeven, ook dat soort steun belangrijk, zo niet noodzakelijk.

Het ligt niet in mijn bedoeling iemand ten achter te stellen. Maar het is ondoenlijk een ieder, die er enig recht op zou mogen laten gelden, hier ook persoonlijk te vermelden. In de loop der jaren zijn bij het RBI/ITRI gewoon (te)velen verschenen en verdwenen.

Laat ik in elk geval beginnen met het bedanken van mijn promotoren. Zij hebben het zaag-, timmer- en schaafwerk begeleid. Bij Ton Hagenbeek zelfs in de laatste tijd ettelijke malen thuis, onder het genot van koffie en koekjes, wat ruimschoots vergoedde dat het 'wel een uur gaans de polder in' was. Ton, bedankt dát je me hebt betrokken bij het leukemie-onderzoek. En voor de manier waarop. Inspirerend, onderhoudend, met geduld onderwijzend. En, ondanks de werkdruk, altijd opgewekt. Ook bedankt dat ik via jou contacten mocht leggen met voor dit werk belangrijke wetenschappers, in het buitenland (op congressen van Californië tot China) en het binnenland (DdHK, EUR). Soms leidde dit tot vruchtbare en heel plezierige samenwerkingsverbanden, zoals bijvoorbeeld met Jacques van Dongen en medewerkers. Bob Mulder ken ik al heel lang als iemand die enthousiasme weet op te wekken door op het interessante in problemen te wijzen en—na even scherp denken—vaak ook de richting aan te duiden, waarin je de oplossing zou moeten zoeken. Bob, bedankt dat je me wegwijs maakte op het gebied van systeemidentificatietechnieken. En bedankt voor je voorspraak om me deze bij TNO te laten toepassen op problemen in de farmacokinetiek, alsmede voor je behulpzaamheid daarbij (ook met het beschikbaar stellen van TUD faciliteiten).

De overige leden van de promotiecommissie wil ik bedanken voor het kritisch doorlezen van het manuscript. Het polijstwerk aan het proefschrift kon geschieden aan de hand van hun waardevolle opmerkingen en suggesties. Ik ben Dick van Bekkum, directeur van het voormalige RBI, dankbaar voor het vertrouwen, dat vliegtuigbouwers—met niet meer dan 'middelbare school'-kennis aan biologie en chemie—een de moeite waard zijnde bijdrage aan het gezondheidsonderzoek kunnen leveren. En voor het geven van de kans om dit daadwerkelijk te gaan bewijzen. Van Johan Broerse leerde ik, dat ook bij tegenslag een gulle lach niet verloren hoeft te gaan. Had iedereen een tiende van zijn werklust, verantwoordelijkheidsgevoel, maatschappelijk (politiek) inzicht en organisatie-talent, dan had 'science' een goede kans 'to prevail'.

Vervolgens moet ik een aantal sleutelfiguren noemen, dat voor de voorbereiding van dit proefschrift onontbeerlijk was. Natuurlijk Ton's Leukemie Team (TLT), ondermeer Ger Arkesteijn, Kees de Groot, Ted Kloosterman. En vóór Anton Martens. Met deze BNML-specialist kon altijd met voldoening van gedachten worden gewisseld over allerlei wetenschappelijk, en anderzins interessante zaken. Niet alleen het ontwerp, de uitvoering en evaluatie van experimenten voor celpopulatiodynamica, waaronder flow cytometry, waren hier in goede handen. Ook de organisatie van buiten-werktijd-activiteiten voor beenmergkamerleden en burenen—van Jan Visser's Lab—was hem wél toevertrouwd. Brainstormen met Kees Nooter, als opvolger van Pieter Sonneveld, over zaken betreffende farmacokinetiek, cytostatica en al dan niet leukemische ratten, leverde vaak vruchtbare ideeën op. Een groot deel van de beschreven biomedische experimenten werd met kundigheid uitgevoerd door een aantal analisten, van wie ik Hans Vollebregt, Arjan de Vries en vooral Carla Ophorst-Van Marrewijk wil noemen.

Voor het toegankelijk maken en houden van de rekenfaciliteiten gaat mijn dank naar de Computerafdeling (Wytze Oostenbrug en medewerkers). Ook mogen, wat dit aspect betreft, meneer Verkerk en Henk Lindenburg (TUD, Vliegtuigbouw) niet worden vergeten. Een aantal illustraties werd vervaardigd door de Afdeling Fotografie (Jan de Kler, Eric van der Reijden en Henk van Westbroek). Medewerkers van de bibliotheek (Paul van Rossum, Bert van der Wurff, Hans van Hoeven) hielpen bij het opzoeken van literatuur. De familie Delfgauw wordt bedankt voor het beschikbaar stellen van de kopieermachine.

Dan wil ik mijn grote waardering uiten voor het geboden huiselijk verkeer bij de Afdeling Barendsen. In het bijzonder voor Ab van Rotterdam. Niet alleen als mijn chocolabrouwende modelbouwmaat, die me de vreugd liet smaken van het delen van gezamenlijke interesses als het ontwerpen van rekenalgorithmen, pogen inzicht in het gedrag van systemen te vergroten, toepassen van statistiek. Maar ook voor het warme onthaal op wetenswaardigheden over filosofie, klassieke muziek, politiek, (klein)kinderen, Nederlandse letterkunde, enz. En voor het fungeren als universeel smeermiddel bij de zakelijke en minder zakelijke contacten tussen klanten en collega's.

Veel dank ben ik verschuldigd aan Hans Zoetelief, voor alle ruimte die ik kreeg om het proefschrift af te schrijven, na de overgang naar een nieuw TNO-onderdeel. Een aantal taken zou zijn blijven liggen, als niet collega's, in het bijzonder Jan Jansen, zonder morren werkzaamheden hadden overgenomen. Ook hiervoor dank.

Tot slot moet worden bedacht, dat de (on)regelmaat van een 'not-925-job' een extra belasting vormt voor het huishouden thuis. Bedankt ma, dat ik altijd mijn gang heb kunnen gaan, zonder dat 'de verzorging' in het gedrang kwam. Ik hoop in de komende tijd een aantal verwaarloosde vriendschapsbanden weer aan te mogen halen.



## Curriculum Vitae

Frank Willem Schultz werd op 25 juni 1951 geboren te Djakarta.

Na het behalen van het diploma h.b.s.-B in 1968 aan het Stevin Lyceum te Den Haag werd in datzelfde jaar begonnen met de studie Luchtvaart- en Ruimtevaarttechniek aan de, toen nog, Technische Hogeschool Delft. Deze studie werd in 1975 afgesloten met het ingenieursexamen bij de vakgroep Stabiliteit en Besturing (prof.dr.ir. O.H.Gerlach), projectgroep Vliegproeven.

Na het vervullen van de militaire dienstplicht (1976-1977) werd een tijdelijk dienstverband aangegaan met het Radiobiologisch Instituut TNO te Rijswijk, om werkzaamheden te verrichten op het gebied van mathematisch modelleren van farmacokinetiek van cytostatica onder leiding van drs. P.Sonneveld en ir. J.A.Mulder.

Dit dienstverband werd in 1980 omgezet in één voor onbepaalde tijd. Het onderwerp farmacokinetiek, nu onder dr. K.Nooter, werd uitgebreid met mathematisch modelleren van celkinetiek. Gaandeweg kwam de nadruk te liggen op het analyseren van gegevens die, onder supervisie van dr. A.Hagenbeek, werden verkregen uit experimenten met het preklinische BNML model voor humane acute myeloïde leukemie, zoals in dit proefschrift beschreven.

Reorganisaties binnen TNO leidden in 1990 tot overplaatsing naar de afdeling Stralingsfysica van dr. J.Zoetelief, welke als onderdeel van een op te richten Centrum voor Stralingsbescherming en Dosimetrie met ingang van 1994 werd ondergebracht binnen het TNO Instituut voor Milieu- en Energietechnologie, Apeldoorn. Met behulp van computersimulatie van stralingstransport wordt hier, onder meer, onderzoek verricht naar de stralingsbelasting van en dosisverdeling in de patiënt bij röntgendiagnostiek en radiotherapie.





