COGNITIVE FUNCTIONING OF CHILDREN TREATED FOR ACUTE LYMPHOBLASTIC LEUKAEMIA WITH CHEMOTHERAPY-ONLY

Nathalie C.A.J. Jansen

COGNITIVE FUNCTIONING OF CHILDREN TREATED FOR ACUTE LYMPHOBLASTIC LEUKAEMIA WITH CHEMOTHERAPY-ONLY

N.C.A.J. Jansen

Stellingen

behorende bij het proefschrift

Cognitive functioning of children treated for acute lymphoblastic leukaemia with chemotherapy-only

Nathalie Jansen, april 2009

1. Betrouwbare afname van neuropsychologisch onderzoek bij kinderen met ALL vlak na diagnose is haalbaar. (*dit proefschrift*)

2. Een prospectieve onderzoeksopzet is bij uitstek het design om mogelijke cognitieve late effecten bij kinderen met ALL na behandeling met chemotherapie goed te onderzoeken. (*dit proefschrift*)

3. Kinderen die behandeld zijn voor ALL met chemotherapie functioneren 4½ jaar na diagnose op cognitief gebied normaal. (*dit proefschrift*)

4. Leereffecten zouden moeten worden verdisconteerd in prospectief, longitudinaal neuropsychologisch onderzoek bij kinderen. (*dit proefschrift*)

5. Men zou meer recht doen aan het kind door bij de indicatiestelling en criteria voor het speciaal onderwijs of het leerling gebonden budget ("het rugzakje") niet uit te gaan van strikte IQ getallen maar van betrouwbaarheidsintervallen.

6. Het testen van kinderen door het land heen, leert je de verschillende regionale culturen van Nederland goed kennen.

7. Men ontdekt meer over een persoon door een uur met hem te spelen dan door een jaar lang met hem te converseren. (*Plato 427-347 v. Chr.*)

8. Ontmoedig nooit iemand die blijvend vooruitgang maakt, hoe langzaam ook. (*Plato 427-347 v. Chr.*)

9. You are the elephant, you are the rider. (*Jonathan Haidt, 2008*)

Cognitive functioning of children treated for acute lymphoblastic leukaemia with chemotherapy-only

ISBN:	978-90-367-3778-4
Illustration cover:	Liek
Lay-out:	Vincent Driessen
Printed by:	Gildeprint BV, Enschede, The Netherlands

This study was granted by the Dutch Cancer Society (RUG 98-1737). The printing of this thesis was financially supported by the faculty of Medical Sciences of the University of Groningen and the Dutch Cancer Society.

cmb.

RIJKSUNIVERSITEIT GRONINGEN

COGNITIVE FUNCTIONING OF CHILDREN TREATED FOR ACUTE LYMPHOBLASTIC LEUKAEMIA WITH CHEMOTHERAPY-ONLY

Proefschrift

ter verkrijging van het doctoraat in de

Medische Wetenschappen

aan de Rijksuniversiteit Groningen

op gezag van de

Rector Magnificus, dr. F. Zwarts,

in het openbaar te verdedigen op

woensdag 8 april 2009

om 16.15 uur

door

Nathalie Chantal Anja José Jansen

geboren op 22 maart 1972

te Noordoostpolder

Promotores	:	Prof.dr. W.A. Kamps
		Prof.dr. J.M. Bouma
Copromotor	:	Dr. A. Kingma
Beoordelingscommissie	:	Prof.dr. O.F. Brouwer
		Prof.dr. W.H. Brouwer
		Prof.dr. R. Pieters

ISBN: 978-90-367-3778-4

Contents

CHAPTER 1

General introduction

Acute Lymphoblastic Leukaemia

Childhood leukaemia is the most common cancer in children and the second most common cause of death in children aged 1-14 years.[1] There are three main types of leukaemia in children, including acute lymphoblastic leukaemia (ALL), acute myelogenous leukaemia (AML) and chronic myelogenous leukaemia (CML). This thesis only refers to children with ALL, representing 80% of the childhood leukaemia's.

The incidence of ALL is 3 to 4 new cases per year per 100,000 children younger than 15 years, with a peak incidence between 2 to 5 years of age.[2] In the Netherlands, each year approximately 120 new cases of ALL are diagnosed in children younger than 15 years.

Leukaemia is a systemic malignancy characterized by abnormal proliferation and expansion of malignant lymphocytes (a specific type of white blood cells), developing in the bone marrow and spreading into the blood. Leukaemic cells may consequently infiltrate in organs such as lymphnodes, spleen, liver, testicles or the central nervous system (CNS).

Symptoms of a child with newly diagnosed ALL reflect the degree and extent of infiltration with leukaemia cells; the overproduction of these abnormal cells disrupts the normal function of several types of blood cells. Presenting signs and symptoms at the time of diagnosis may include fatigue, pallor, fever, easy bruising, infections, bone pain and enlarged lymphnodes, spleen or liver.

The pathogenesis of ALL is largely unknown. Both genetic, infectious, immunological and environmental factors are probably involved. There are various genetic subtypes of ALL which are associated with different clinical outcome.

During the past 3 decades, treatment has significantly changed and the prognosis of childhood leukaemia has dramatically improved. Treatment includes several phases, each with administration of multichemotherapy. The goal of the first phase, remission-induction therapy, is to eradicate most of the initial burden of leukaemia cells and to restore normal hematopoiesis. Secondly, CNS-prophylactic treatment is essential to prevent leukaemia cells to spread into the CNS. The elective or pre-symptomatic treatment of CNS involvement is an important component of treatment and has significantly contributed to improved prognosis. Thereafter, consolidation or intensification and maintenance therapy is required for 2 years to prevent a relapse (return) of the disease.

With current therapy, long-term survival rates are approximately 80 %.[3,4] With continued improvement in survival rates the focus has begun to change toward the reduction of unwanted acute and late effects of treatment.

Previously, CNS prophylaxis consisted of cranial irradiation (CI) and IT chemotherapy, usually methotrexate (MTX), or MTX combined with other drugs. Unfortunately, strong evidence emerged that this successful treatment was associated with persistent cognitive impairment. Numerous studies found deterioration of intelligence and impairment of memory and attention particularly if CI was given at a young age.[5] Therefore, CI has been largely replaced by IT therapy and high-dose systemic chemotherapy since the eighties. However, there is still controversy whether non-CI treatment, i.e. chemotherapy only (ChO), has or has no adverse effects on cognitive functioning. Although patients treated with ChO show undoubtedly better performance than irradiated children, there is still insufficient evidence whether cognitive development of ChO patients is equal to their healthy siblings.

Aim and design of the present study

We investigated both early and late neuropsychological effects of ALL treatment according to the DCOG (Dutch Childhood Oncology Group) ALL protocol 9. In this nationwide study, we applied a prospective, longitudinal design with healthy siblings as controls. In addition to intelligence, the domains of memory, learning, graphical construction, attention, cognitive flexibility and fine-motor functioning were investigated. We included 50 patients and 29 controls aged between 4 and 12 years to avoid test shift; i.e. the need to change tests with growing age because most tests have a limited age range.

Questions of the present study were:

1. Is neuropsychological assessment to establish base-line functioning feasible in seriously ill, newly and recently diagnosed patients with ALL?
2. Is change in cognitive functioning during and after 2 years of treatment the same in patients and sibling controls?
3. Is possibly long-term cognitive functioning in children treated for ALL with ChO equal to healthy controls and normative data, 4½ to 5 years after diagnosis?

4. Is possibly poorer performance or cognitive deterioration related to specific patients' characteristics?

Advanced statistical analyses including multilevel analyses were employed to correct for missing values during this longitudinal study.

Outline and scope of this thesis

Following this introduction, in **chapter 2**, the feasibility and results of neuropsychological assessment in leukaemia patients shortly after diagnosis are presented.[6]

In **chapter 3**, results of post-treatment evaluations of intellectual functioning are offered.[7]

Long-term (4 years after diagnosis) neuropsychological and intellectual outcome is described in **chapter 4** and **chapter 5**, respectively.[8,9]

Chapter 6 presents an overview of the literature on the neuropsychological functioning after ChO treatment in ALL patients. To date, studies yielded inconsistent results; methodological difficulties and limitations and differences among studies are discussed.[10]

Finally, in **chapter 7**, main findings and general conclusions and recommendations for future research are given.

Collaborative effort

This study was initiated in Groningen and started with a pilot of 10 subjects in March 1998. In The Netherlands, diagnosis and treatment of ALL is coordinated by the DOCG. Patients are treated on national protocols in specialized paediatric oncology centres. Participating paediatric oncology centres for this study were: Academic Medical Centre, Amsterdam, VU University Medical Centre, Amsterdam, Leiden University Medical Centre, Nijmegen St Radboud University Medical Centre, and University Medical Centre Utrecht.

This study was financially supported by a grant from the Dutch Cancer Society.

References

1. Linett MS, Hatch EE, Kleinerınann RA, et al. Residential exposure to magnetic fields and acute lymphoblastic leukemia in children. N Eng J Med 1997;337:1-7
2. Pizzo PA, Poplack DG: Principles and Practice of Pediatric Oncology, fourth ed. Philadelphia, PA: Lippincott Williams & Wilkins, 2001
3. Pui CH. Acute lymphoblastic leukemia in children. Curr Opin Oncol 2004;12:3-12
4. Chessells JM, Harrison G, Richards SM, Gibson BE, Bailey CC, Hill FGH et al. Failure of a new protocol to improve treatment results in paediatric lymphoblastic leukaemia: lessons from the UK Medical Research Council trials UKALL X and UKALL XI. Br J Haematol 2000;118:445-55
5. Roman DD, Sperduto PW. Neuropsychological effects of cranial radiation: current knowledge and future directions. Int J Radiat Oncol Biol Phys 1995;13(4):983-998
6. Jansen NC, Kingma A, Tellegen P, et al. Feasibility of neuropsychological assessment in leukaemia patients shortly after diagnosis: directions for future prospective research. Arch Dis Child 2005;90(3):301-4
7. Jansen NC, Kingma A, Schuitema A, et al. Post-treatment intellectual functioning in children treated for acute lymphoblastic leukaemia (ALL) with ChO: a prospective, sibling-controlled study. Eur J Cancer 2006;42(16):2765-72
8. Jansen NCAJ, Kingma A, Schuitema A, et al. Neuropsychological outcome in chemotherapy-only treated children with acute lymphoblastic leukemia. J Clin Oncol 2008;26(18):3025-30
9. Jansen NCAJ, Kingma A, Schuitema A, et al. Preserved Intellectual Functioning in Children treated for Acute Lymphoblastic Leukaemia with Chemotherapy-only. Submitted
10. Jansen NC, Kamps WA, Bouma A, et al. Cognitive functioning after chemotherapy-only in children with acute lymphoblastic leukaemia (ALL): A review, discussion and methodological issues.

CHAPTER 2

Feasibility of neuropsychological assessment in leukaemia patients shortly after diagnosis: directions for future prospective research

Nathalie C. Jansen, Annette Kingma, Peter Tellegen,
Rieneke I. van Dommelen, Anke Bouma, Anjo Veerman, Willem Kamps

Archives of Disease in Childhood 2005;90:301-304

Abstract

Aims: To study neuropsychological functioning of newly diagnosed children with acute lymphoblastic leukaemia (ALL) within two weeks after diagnosis in order to determine the feasibility of a sibling controlled prospective study design.

Methods: Fifty consecutive patients (median age at testing 6.6 years, range 4-12) were included in a prospective, longitudinal, nationwide study. Treatment would include intrathecal and systemic chemotherapy according to the DCLSG ALL-9 protocol. Children were evaluated with an extensive neuropsychological battery including measures of intelligence, memory, attention, language, visual-constructive function, and fine-motor abilities within two weeks after start of the chemotherapy. The control group consisted of 29 healthy siblings (median age at testing 8.2 years, range 4-12), who were tested < 4 weeks after the patients' assessment.

Results: Mean scores on Wechsler Intelligence Scales did not differ significantly between patients and siblings; mean IQ scores for both the patients and the controls were high average. To examine specific neuropsychological functions, norm scores based on the exact age were acquired by fitting procedures, but no significant differences were found.

Conclusions: Neuropsychological assessment of patients during early hospitalisation is feasible. The results indicate no adverse effect of illness and psychological factors on IQ and neuropsychological functioning of patients with recently diagnosed ALL. The prospective design of this study of cognitive late effects of chemotherapy will allow discrimination between adverse sequelae of disease and treatment.

Introduction

Acute lymphoblastic leukaemia (ALL) is the most common form of childhood cancer. Approximately 80% of newly diagnosed children with ALL are curable with modern treatment.[1,2] Following this improved survival rate, an increasing number of studies has focused on the quality life of the survivors. Patients who have been treated with cranial irradiation (CI) and additional chemotherapy have shown intellectual deterioration and specific neuropsychological deficits.[3-5] To date, prospective longitudinal studies on neuropsychological sequelae in children treated for ALL with chemotherapy only are rare or have yielded inconsistent results.[6-8] These inconsistencies may be understood from less suitable control groups, different ages at time of testing (age-effect), and selection of neuropsychological measures.

Moreover, pretreatment neuropsychological assessments are rarely conducted. It is commonly thought that testing shortly after diagnosis is not feasible because children diagnosed with ALL are seriously ill and have to cope with medical procedures and intensive treatment immediately after diagnosis. Leukaemia or leukaemia treatment can furthermore cause emotional, non-organic distress in patients and families, which may influence test behaviour of the children.

In 1999, we initiated a prospective longitudinal and nationwide study in the Netherlands, which includes siblings as controls, applies a comprehensive test battery, and has a broad age spectrum. Here we report the results of the neuropsychological assessment in patients shortly after diagnosis, and their healthy siblings. The results will eventually be used to investigate both early and late neuropsychological effects of chemotherapy according to the Dutch Childhood Leukemia Study Group (DCLSG) ALL-9 protocol. In this report, we review the results of neuropsychological assessment shortly after diagnosis of both patients and their healthy siblings.

Methods

Patients and sibling controls

From January 1999 to June 2001, 79 consecutive patients from six participating paediatric oncology centres in the Netherlands were eligible for this study. Criteria of eligibility were: newly diagnosed patients with high or standard risk ALL; age between 4 years and 12 years and 3 months; and Dutch as primary language. Informed consent was obtained according to each hospital's rules. Patients with initial CNS leukaemia and patients with pre-existent disorders that could interfere with normal cognitive development were excluded.

Sixteen (20 %) parents refused participation because of the expected burden, and 19 (24%) cases were missed due to logistical problems.

Between March 1998 and January 1999, six consecutive patients has been enrolled in a pilot study in the hospital which coordinated the study. These patients did not significantly differ from the children in the main study; hence, a combined group of 50 patients entered the study.

The control group consisted of 25 healthy siblings who met the same inclusion criteria as the patients. If the patient had more than 1 sibling, the child (1) closest in age to the patient and (2) the same sex was chosen. Table 1 shows the characteristics of patients and siblings.

Table 1: Characteristics of patients and siblings at the first neuropsychological evaluation shortly after diagnosis of ALL

Group	Male n (%)	Female n (%)	Age at testing Median (range)
Patients	30 (60)	20 (40)	6.6 (7.8)
Healthy controls (siblings)	11 (38)	18 (62)	8.2 (8.1)

Treatment

Patients had just started treatment according to the national chemotherapy only DCLSG-ALL-9 protocol, including vincristine, dexamethasone, daunorubicine, and triple intrathecal (IT) therapy as CNS prophylaxis. This protocol is similar to the DCLSG ALL-6 protocol.[9-11] The patients had received one cycle of vincristine, dexamethasone, daunorubicin, and triple IT therapy before their first assessment.

Study design

Patients were individually evaluated within two weeks after diagnosis and start of treatment. Siblings were individually assessed within four weeks after the patients' evaluation. Patients and siblings did not significantly differ with respect to age at testing and gender. There were no indications that patients with or without siblings differed in socioeconomic status (SES). To maximise standardisation, all participants were nationwide tested by one qualified child neuropsychologist who travelled to the hospitals were the children were treated.

Patients and healthy sibling controls were evaluated with an age appropriate comprehensive standardised neuropsychological test battery (table 2). Children aged 4-6 years were administered a developmental screening test and measures of intelligence, visual-motor integration, and if $\geq$ 5 years, fine-motor functioning. Participants aged 6-12 years were assessed with a more extensive test battery. The neuropsychological assessment of these children took about three hours, including measures of intelligence and specific cognitive functions as verbal-auditory and visual memory, visual-motor integration, attention, cognitive flexibility, verbal fluency, and fine-motor functioning. If necessary, the assessment was split into two sessions.

Table 2: Neuropsychological battery

Neuropsychologic domain	Measures	Age	No. patients	No. siblings
Mental, motor and social development	Denver Developmental Scales[13]	4-6	17	6
Intelligence	Wechsler Pre-School and Primary Scales of Intelligence (WPPSI), 10 subtests[14]	4-6		
	FS-IQ		12	6
	V-IQ		13	6
	P-IQ		12	6
Intelligence	Wechsler Intelligence Scale for Children (WISC-R), 10 subtests[15]	6-13		
	FS-IQ		26	23
	V-IQ		30	23
	P-IQ		26	23
	Concentration factor		28	23
	Perceptual Organisation		25	23
Verbal-auditory learning and memory	Dutch version of Rey´s Auditory-Verbal Learning Test (RAVLT)[16]	6-13	26	23
Visual memory	Rey-Osterreith Complex Figure Test delayed (CFT) recall1[7]	6-13		
Verbal fluency	Animal-naming Fluency Test[17]	6-13	30	23
Sustained attention/ speed	Bourdon-Vos ; self-paced, continuous performance cancellation task[12]	6-13	20	21
Cognitive flexibility	Wisconsin Card Sorting Test (WCST)[18]	6-13	26	23
Perceptual-motor skills	Beery Developmental Test of Visual-Motor Integration (VMI)[19]	4-13	47	28
Visuospatial constructional ability/ planning	Rey-Osterreith Complex Figure Test (CFT) copy[16]	6-13	26	23
Fine-motor function	Purdue Pegboard[20]	5-13	29	27

Statistical analysis

Performances of patients were compared to those of sibling controls using non-directional two tailed Student's *t* tests for paired groups.

For the Wechsler Pre-School and Primary Scales of Intelligence (WPPSI-R), Experimental Dutch-Flemish version, and for the Wechsler Intelligence Scale for Children-revised (WISC-R, Dutch version), mean norm scores are 100 (SD = 15). For the remaining tests, norms have been acquired by fitting procedures based on the raw scores and the exact ages resulting in norm scores (mean = 50; SD = 10). The fitting procedures were based on the published norm data (means and standard deviations for different age groups) in the respective test manuals or other publications. This procedure enables comparisons of standardised scores between subjects of any specific age.[21,22] Significance levels were established at $p < 0.05$. Statistical analyses were performed using version 10 of the Statistical Package for the Social Sciences (SPSS Inc, Chicago, IL, U.S.A.).

Results

Included patients (n = 50) did not significantly differ from missing patients (n = 35) in terms of sex, age at diagnosis, and initial characteristics of disease and prognostic risk group. We had no indication of differences in socioeconomic status between included and missing patients. The latter mainly emanated from two hospitals; patients were missed due to illness of the psychologists who should have referred eligible patients. Patients and siblings aged 4-6 years at diagnosis were assessed as essentially normal on the Denver Developmental Scales. Patients aged 4-6 years scored significantly higher than siblings on WPPSI-R FS-IQ and WPPSI-R VIQ (table 3). Comparing patients aged 6-13 and siblings, no significant differences were found for any WISC-R factor. IQs were high average for patients on the WPPSI-R and both patients and siblings on the WISC-R.

Table 3: Results of intelligence testing, comparison of patients to siblings at the first evaluation shortly after diagnosis of ALL

Wechsler scales	Patients mean (SD)	Siblings mean (SD)	t value	P value	95% CI	
4-6 years						
WPPSI-R FS-IQ	114.7 16.7)	101.5 (7.3)	2.32	.034	1.2	25.2
WPPSI-R V-IQ	116.1 (14.5)	99.5 (8.4)	2.59	.019	3.1	30.1
WPPSI-R P-IQ	108.8 (13.9)	106.0 (19.6)	.36	.726	-14.0	19.7
6-12 years						
WISC-R FS-IQ	108.1 (15.4)	107.1 (10.9)	.26	.799	-6.8	8.8
WISC-R V-IQ	106.9 (14.1)	107.0 (10.8)	-.04	.970	-7.3	7.0
WISC-R P-IQ	109.0 (17.5)	108.0 (13.3)	.22	.825	-8.0	10.0
WISC-R Perceptual organisation	109.5 (15.9)	106.0 (14.1)	.80	.428	-5.3	12.2
WISC-R Concentration	106.5 (15.0)	108.0 (12.1)	-.38	.707	-9.2	6.3

WPPSI-R; V-IQ (information, similarities, arithmetic, vocabulary, digit span), P-IQ (object assembly, block design, mazes, picture completion, animal pegboard).
WISC-R; V-IQ (information, similarities, arithmetic, vocabulary, digit span), P-IQ (picture completion, block design, object assembly, coding, mazes), Perceptual organisation factor (picture completion, block design, object assembly, mazes), Concentration factor (arithmetic, digit span, coding).

Table 4 shows results for the remaining cognitive measures. No significant differences between the groups were found for any test. Overall, patients and siblings had average scores.

Table 4: Results of neuropsychological tests comparing ALL group with siblings at the first evaluation shortly after diagnosis

Test-measures	Patients Mean (SD)	Siblings Mean (SD)	t value	P value	95% CI	
RAVLT						
Immediate recall	50.6 (8.0)	52.2 (8.7)	-.67	.506	-6.4	3.2
Delayed recall	52.4 (11.0)	49.2 (9.8)	1.06	.295	-2.9	9.2
Fluency Test: Animal-naming	63.7 (10.5)	60.4 (9.7)	1.17	.248	-2.4	9.0
Bourdon-Vos						
Speed	53.8 (14.3)	48.8 (8.7)	1.36	.183	-2.5	12.4
Accuracy	51.0 (10.4)	49.7 (7.1)	.45	.658	-4.4	6.8
WCST						
Errors	49.4 (11.1)	50.7 (8.8)	-.44	.664	-7.1	4.6
Perseverations	49.3 (10.7)	50.8 (8.6)	-.54	.591	-7.1	4.1
Trials administered	50.0 (10.7)	50.7 (8.9)	-.28	.784	-6.5	4.9
Beery VMI	47.6 (9.7)	51.4 (11.6)	-1.52	.133	-8.7	1.2
Rey-Osterreith CFT						
Copy	54.3 (5.9)	54.7 (5.5)	-.26	.795	-3.7	2.9
Delayed recall	44.5 (8.4)	46.6 (6.1)	-.91	.370	-6.5	2.5
Purdue Pegboard						
Dominant hand	45.5 (10.5)	45.7 (10.1)	-.11	.916	-5.8	5.2
Non-dominant hand	48.3 (8.3)	46.4 (8.6)	.81	.424	-2.8	6.4
Both hands	49.0 (11.1)	48.3 (7.3)	.25	.804	-4.5	5.7
Assembly	52.5 (11.7)	51.9 (9.1)	.20	.840	-5.1	6.3

Discussion

We have shown that recently diagnosed children with the life threatening disease ALL can be reliably assessed with an extensive standardised neuropsychological test battery shortly after diagnosis. An important observation in this study was that the majority, even the very young children, enjoyed the assessment, which was rather a distraction among numerous medical procedures than an emotional burden. Moreover, this study is strengthened by the inclusion of healthy siblings as controls, who were also pleased to be involved in the study and enjoyed the special attention. This control group enables appropriate comparison with the healthy population. Decrements in test results within the patient group can be detected, even if the results are still above average. For accurately assessing changes, precise standardised age scores are essential. Therefore the validity of this study is enhanced by using a fitting procedure for the construction of test norms which provides standardised scores based on the exact age of the subjects.[21,22]

Our data correspond with the few other studies offering baseline pretreatment in newly diagnosed patients which also showed no obvious different results.[7]

The present study can be criticized for the high number of missing patients, which could possibly account for bias in these test results. However, this is unlikely because included patients did not significantly differ from missed patients concerning demographic and initial disease characteristics. Missed patients should mainly have been referred by two ill psychologists. Fortunately, patients in these hospitals represent a random patient population, hence we have no indication that characteristics of the missed children differed from those who could be included. With the current numbers we would detect IQ differences of 0.7 SD (10.5 IQ points) to obtain an adequate power of 80%. To illustrate the meaning of 10.5 SD, a difference between 105 and 95 would be significant, but both IQs are considered average and children with both IQs would be in a regular school class. We could not control for SES. Given the overall average results, bias does not seem likely. In general, patients showed greater standard deviation both on the intelligence tests and neuropsychological tests. However, differences between patients and siblings did not result from one or few individuals with extreme scores.

The norm scores of the Experimental Dutch-Flemish version of the WPPSI-R were recently evaluated as disputable, which could explain the above-average IQs in the young

patients (table 3). However, if the patients' IQs are overrated, we could expect above-average IQs in the siblings as well. There were no demographic differences explaining the IQ differences between patients and siblings aged 4-6 years. The scores of the children tested with the WISC-R are high average as a result of the Flynn effect, accounting for an IQ rise of about 6 points since test norms were collected in the early 1980s.[23] If evaluated with more recent test norms these children would probably have average results.

Generally, it is often suggested that emotional, non-organic distress influences the test results. However, such an effect is very unlikely given the normal outcome. Even measures of attention and memory, known to be sensitive for emotional distress,[24] did not differ between patients and siblings.

Conclusion

The present data strongly suggest that patients do not suffer from neuropsychological deficits related to acute disease or early treatment. In the future, patients' base-line scores can be used to discriminate between possible adverse sequelae of disease and/ or treatment and eventually, to optimise treatment protocols compromising between high cure rate and good quality of life. Ideally, neuropsychological assessment early after hospitalisation also selects patients who need early intervention for mental or academic deficits, but this was not the aim of this study.

Neuropsychological assessment of children with ALL shortly after diagnosis with sibling controls is feasible and essential to discriminate between adverse sequelae of treatment. Prospective, longitudinal study designs should become the standard for evaluating possible treatment effects.

Acknowledgements

This study is a collaborative multicentre investigation, which is coordinated by the Paediatric Oncology Centre, Groningen University Hospital. Participating centres: Wilhelmina Children´s Hospital, Utrecht; St Radboud University Medical Centre, Department of Paediatrics, Nijmegen; Vrije Universiteit University Medical Centre, Department of Paediatrics, Amsterdam; University Medical Centre, Emma Children´s Hospital, Amsterdam; University Medical Centre, Willem-Alexander Children and Youth Centre, Leiden. This study was supported by a grant from Dutch Cancer Society.

The authors would like to thank the children and families who participated in this study, and dr. A. Jennekens-Schinkel for reviewing this paper.

References

1. Pui CH. Acute lymphoblastic leukemia in children. Curr Opin Oncol 2000;12:3-12.
2. Chessells JM, Harrison G, Richards SM, et al.
   Failure of a new protocol to improve treatment results in paediatric lymphoblastic leukaemia: lessons from the UK Medical Research Council trials UKALL X and UKALL XI. Br J Haematol 2002;118:445-55.
3. Mulhern RK, Fairclough D, Ochs J. A prospective comparison of neuropsychologic performance of children surviving leukemia who received 18-Gy, 24 Gy, or no cranial irradiation. J Clin Oncol 1991;9:1348-1356.
4. Langer T, Martus P, Ottensmeier H, et al. CNS late-effects after ALL therapy in childhood. Part III: neuropsychological performance in long-term survivors of childhood ALL: impairments of concentration, attention, and memory. Med Pediatr Oncol 2002;38:320-8.
5. Kingma A, Van Dommelen RI, Mooyaart EL, et al. Slight cognitive impairment and MRI abnormalities but normal academic career in children treated for acute lymphoblastic leukemia with chemotherapy only. J Pediatr 2001;3:413-420.
6. Copeland DR, Moore BD, Francis DJ, et al. Neuropsychologic effects of chemotherapy on children with cancer: A longitudinal study. J Clin Oncol 1996;14:2826-2835.
7. Kingma, A, Van Dommelen, RI, Mooyaart, EL, et al. No major cognitive impairment in young children with acute lymphoblastic leukemia using chemotherapy only: a prospective longitudinal study. J Pediatr Hematol Oncol 2002;24:106-113.
8. Espy KA, Moore IM, Kaufmann PM, et al. Chemotherapeutic CNS prophylaxis and neuropsychologic change in children with acute lymphoblastic leukemia: a prospective study. J Pediatr Psychol 2001;26:1-9.
9. Kamps WA, Bokkerink JP, Hakvoort-Cammel FG, et al. BFM-oriented treatment for children with acute lymphoblastic leukemia without cranial irradiation and treatment reduction for standard risk patients: results of DCLSG protocol ALL-8 (1991-1996). Leukemia 2002;16:1099-111.
10. Kamps WA, Veerman AJ, van Wering ER, et al. Long-term follow-up of Dutch Childhood Leukemia Study Group (DCLSG) protocols for children with acute lymphoblastic leukemia, 1984-1991. Leukemia 2000;14:2240-6.
11. Veerman AJ, Hahlen K, Kamps WA, et al. High cure rate with a moderately intensive treatment regimen in non-high-risk childhood acute lymphoblastic leukemia. Results of protocol ALL VI from the Dutch Childhood Leukemia Study Group. J Clin Oncol 1996;14:911-8.
12. Bourdon-Vos test manual (3rd edn.). Lisse, The Netherlands: Swets & Zeitlinger; 1998.
13. Cools ATM, Hermanns JMA. Handleiding Denver Ontwikkeling Screeningtest (DOS). Lisse, The Netherlands: Swets & Zeitlinger, 1986.
14. Vander Steene G, Bos A. Wechsler Preschool and Primary Scales of Intelligence-revised, Vlaams-Nederlandse aanpassing (Dutch manual). Lisse, The Netherlands :Swets & Zeitlinger, 1997.
15. Vander Steene G, van Haasen PP, de Bruyn EEJ, et al. Wechsler Intelligence Scale for Children-revised (Dutch manual). Lisse, The Netherlands: Swets & Zeitlinger, 1986.

16. Van der Burg W, Kingma A. Performance of 225 Dutch school children on Rey's Auditory Verbal Learning Test (AVLT): Parallel test-retest reliabilities with an interval of 3 months: normative data. Arch Clin Neuropsychol 1999;14:545-559.
17. Lezak MD. Neuropsychological assessment (3rd ed.). New York: Oxford University Press, 1995.
18. Chelune GJ, Baer RA. Developmental norms for the Wisconsin Card Sorting test. J Clin Exp Neuropsychol 1986;8:219-28.
19. Beery KE. Revised administration, scoring and searching manual for the development test of visual-motor integration. Cleveland: Modern Curriculum Press, 1982.
20. Gardner RA, Broman M. The Purdue Pegboard; normative data on 1334 schoolchildren. J Clin Child Psychol 1979;8:156-162.
21. Tellegen PJ, Laros JA. The construction and validation of a nonverbal test of intelligence: the revision of the Snijders-Oomen tests. Europ J Psychol Assess 1993;9:147-157.
22. Snijders JTh, Tellegen PJ, Laros JA.SON-R 5½-17, Snijders-Oomen nonverbal intelligence test. Manual and research report. Groningen: Wolters-Noordhoff, 1989.
23. Flynn JR. Massive IQ gains in 14 nations: what IQ tests really measure. Psychol Bull 1984;101:171-191.
24. Kusché CA, Cook ET, Greenberg MT. Neuropsychological and cognitive functioning in children with anxiety, externalizing and comorbid psycho pathology. J Clin Child Psychol 1993;2:172-195.

CHAPTER 3

Post-treatment intellectual functioning in children treated for acute lymphoblastic leukaemia (ALL) with chemotherapy-only: a prospective, sibling-controlled study

Nathalie C. Jansen, Annette Kingma, Arnout Schuitema, Anke Bouma, Jaap Huisman, Anjo Veerman, Willem Kamps

European Journal of Cancer 2006;42(16):2765-72

Abstract

Intellectual functioning (verbal, performance and full-scale IQ) in 43 children treated for acute lymphoblastic leukaemia (ALL) with chemotherapy-only was evaluated in a nationwide, prospective, sibling-controlled study. Intellectual assessment was performed at diagnosis and repeated shortly after cessation of 2 years treatment, including intrathecal and systemic chemotherapy. Using hierarchical regression analysis, patients' and siblings' (n = 27) scores were longitudinally analysed and compared to assess possible changes and differences over time. At both assessments, before and after treatment, patients showed average scores on intelligence tests compared to population norms. Longitudinal analysis and cross-sectional comparisons revealed no significant differences between patients and controls. Young patients showed a small relative decline, albeit not significant, on performance-IQ compared to healthy siblings. Despite intensive and potentially neurotoxic treatment, no evident negative effects on intelligence were found. However, it can not be precluded that younger patients are at risk for a small decline in PIQ.

Introduction

Acute lymphoblastic leukaemia (ALL) is the most common form of childhood cancer with a cure rate approaching 80%. To achieve this good outcome, prophylactic treatment of the central nervous system (CNS) in addition to systemic chemotherapy is essential.[1,2] Former treatment protocols using cranial radiation (CR) as prophylaxis have shown lower intelligence scores and specific neuropsychological deficits post-treatment in patients compared to controls.[3-5] To avoid these adverse late effects, children with ALL have been treated with chemotherapy-only regimens since the early 1980s.

While patients treated with chemotherapy-only perform better on intelligence tests than those who received CR,[6-8] it is yet unclear if the intellectual achievement in the former patients is equal to healthy peers. In a recent review by Moore, it was concluded that studies of the effects of chemotherapy in isolation are far less frequent compared to the effects of CR on cognitive outcome.[9] Some studies report lower scores on intelligence tests compared to controls,[10,11] while others find no deterioration of intellectual abilities over time.[12,13] Methodological problems including lack of proper control groups, missing base-line assessments prior to treatment, limited age range, test-shift and small sample size may have added to the inconsistent results found in chemotherapy-only late effect studies.[14-16]

To overcome such methodological problems, we started a nation-wide, prospective sibling-controlled study, in which patients were assessed with Wechsler intelligence scales within two weeks after start of chemotherapeutic treatment and after cessation of therapy. We recently showed that testing of patients shortly after diagnosis is feasible and reliable.[17] In that study, no adverse effect of illness and psychological stress on the IQ of patients with recently diagnosed ALL was found. In the present report, we focus on patients' intellectual functioning over time until shortly after cessation of treatment, in comparison to their base-line performance and to sibling-controls.

Patients and methods

Subjects

From 1999 to June 2001, 85 consecutive patients from six paediatric oncology centres in the Netherlands were eligible for the study. Inclusion criteria for this study were newly diagnosed children with ALL, age between 4.0 years and 12.3 years, and Dutch as their primary language. The age criterion was chosen because the follow-up time was 4.6-5.0 years (two windows of 3 months for two re-assessments). Most psychometric tests for children and youngsters can be used till the age of 17; hence, the upper age threshold at the time of inclusion could not exceed 12 years. Patients with initial CNS leukaemia or pre-existent disorders that could interfere with normal cognitive development were excluded. Sixteen out of 85 parents refused to participate because of the expected burden and 19 patients did not enter the study due to a long absence of the referring psychologist in two centres. Those two psychologists should have referred the patients in two of the participating hospitals to the single neuropsychologist who had done all assessments. Fortunately, a very low number of patients were lost during follow-up. The 50 participating children represented 59% of potentially eligible patients and did not differ significantly from non-recruited children with regard to age, gender, SES and disease characteristics. Moreover, the missing patients from the two afore-mentioned centres did not differ from the national cohort of paediatric leukaemia patients.

Twenty-nine patients had healthy siblings who could serve as a control with the same inclusion criteria as the patients with respect to age, language and normal development. If a patient had more than one eligible sibling, the child closest in age to the patient was chosen. The characteristics of patients and sibling controls are given in Table 1. Parental informed consent was obtained according to institutional rules. Full details of the group of children included in the first neuropsychological assessment (NPA-I) were described in Jansen and colleagues.[17]

Measurements could be repeated in 44/50 patients, but only 43 were included in the analysis. Reasons for attrition were relapse of ALL or death (n = 3), refusal to further participation (n = 2), or switch to another treatment protocol (n = 1). Retrospectively, there were strong indications for pre-morbid mental retardation in one child, who could not complete the intelligence test at NPA-I. After the second neuropsychological assessment

(NPA-II), it was decided to exclude this patient from the study, so that an aggregate group of 43 patients who were in complete continuous remission could be assessed twice. To preclude any bias, we excluded the results of 11 patients who had incomplete IQ tests at the first assessment from the longitudinal analysis but these children could be re-subjoined in the post-treatment comparison with the siblings. Mean FS-IQ at NPA-II of these patients was lower (101.3), compared to the main patient group, but these relatively lower scores were caused by only two individual patients who had an IQ < 90. The 32 patients who were included in the longitudinal analysis were comparable in respect to age, gender and disease characteristics to the rest of the study population. Of the 29 siblings, 2 (7%) refused to participate at the NPA-II; hence a total of 27 could be assessed.

Treatment

Patients were treated according to the national chemotherapy-only Dutch Childhood Oncology Group (DCLOG) ALL-9 protocol, including systemic chemotherapy (vincristine, dexamethasone, L-asparaginase, medium dose methotrexate (MTX), 6-MP and repeated triple IT (MTX, PD, ara-C) therapy as CNS prophylaxis. This protocol is similar to the DCLSG ALL-6 protocol.[18-20] High risk patients (16/50) received additional systemic chemotherapy including daunorubicin, cyclofosfamide and cytosine-arabinoside. All patients had received one dose of vincristine, dexamethasone, (daunorubicin in case of high risk patients) and triple IT therapy before their first assessment. Total duration of treatment was 108 weeks.

Study design

At NPA-I, patients were individually evaluated within two weeks after diagnosis. Siblings were individually assessed within four weeks after the patients' evaluation.[17] NPA-II was repeated three to six months after cessation of therapy, that is, 2.3 – 2.6 years after NPA-I (Md = 2.4 years).

To optimise standardization, all participants were tested nationwide by one qualified child neuropsychologist. The patients were tested either at home or at the hospital but no difference in IQ-scores was found between these two sites.

Test materials and procedures

At NPA-I, not all children could be tested with the same IQ-test. Children aged 4-6 years were assessed with the Experimental Dutch-Flemish version of the Wechsler Pre-School and Primary Scales of Intelligence (WPPSI-R) (10 subtests; extrapolating for the subtests comprehension and geometric design).[21] The children ≥ 6 years were tested with the Wechsler Intelligence Scale for Children-revised (WISC-R, Dutch version) (10 subtests; extrapolating for the subtest comprehension and picture arrangement).[22] At NPA-I, the patients (N = 11) who missed > 2 subtests were excluded from the longitudinal analysis. The reasons for missing data were an infusion in the (dominant) hand, illness, pain, tiredness or a combination of these factors. At NPA-II, all children were tested with the WISC-R. Twenty-one patients and 21 siblings were assessed twice with the WISC-R. For the WPPSI-R and for the WISC-R, mean norm-scores are 100 (SD = 15).

Statistical Methods

To compare patients to siblings, regression analyses were preferred to more traditional ANCOVA because an interaction effect between age and group was expected. Moreover, this statistical technique is often used such that a stepwise control can be performed for possible confounding effects.[23]

We conducted three different analyses (Table 1). First, longitudinal comparisons were made by hierarchical regression analysis for the group of patients and siblings who were assessed with the same test (WISC-R) at both NPA-I and NPA-II. In this model, the differences between NPA-II and NPA-I for FS-IQ, verbal IQ (VIQ) and performance IQ (PIQ) performances were predicted with age, group and interaction between age and group as explanatory variables.

Second, data of a subgroup of patients and siblings who had been tested with the WPPSI-R at NPA-I and with the WISC-R at NPA-II are described. For the given small sample size, only descriptive statistics are given for this group.

Last, IQ's of all patients (N = 43) were analysed at NPA-II by hierarchical regression analysis to investigate the largest possible group of patients and compare them with the siblings while controlling for age and test-shift (if a child was tested at NPA-I with WPPSI-R, and at NPA-II with the WISC-R).

For analysis 1 and 3, the results in patients and siblings are presented by: (a) descriptive results; (b) scatter plots; and (c) hierarchical regression analysis.

A comparison of characteristics of patients and siblings showed that there were slight, insignificant differences between patients and siblings with respect to gender and age at testing; median age for siblings (Md = 8.2) was 1.8 years higher than for patients (Md = 6.4).

The influence of gender as a possible confounding effect on the test scores was found nil. Finally, no strong regression assumption violations were found.[23]
Significance levels were established at $p < .05$.

The Statistical Package for the Social Sciences (SPSS Inc, Chicago, IL, USA) WINDOWS 10.0 was used for the statistical analyses.

Table 1: Numbers and characteristics of patients and siblings

	Patients	**Siblings**
Base-line inclusion 17	50	29
Incomplete IQ measures at NPA-I	(11[a])	(0)
Drop outs	7	2
Longitudinal analysis (N total)	32	27
Longitudinal analysis (>6 years at NPA-I; analysis 1)	*21*	*21*
Median age at diagnosis (range)	7.6 (6.2-11.7)	8.3 (6.8-12.6)
Female, %	38%	57%
Longitudinal analysis (<6 years at NPA-I; analysis 2)	*11*	*6*
Median age at diagnosis (range)	5.0 (4.0 – 5.9)	5.6 (4.5 – 5.9)
Female, %	18%	67%
Cross-sectional, post-treatment analysis (all ages; analysis 3) (32+11[a])	43	27
Median age at diagnosis (range)	6.4 (4.0-11.7)	8.2 (4.5-12.6)
Female, %	41%	62%

[a] Patients who had incomplete measures at NPA-I but full measures at NPA-II and then added to the post-treatment analysis.

Results

1. Longitudinal analysis (analysis 1; subjects > 6 years at NPA-I)

a. Descriptive results

Fig. 1 offers the results of 21 patients and 21 siblings who were assessed twice with the WISC-R. The mean scores of siblings and patients at NPA-I and NPA-II were high average compared to population norms.

Table 2 shows small differences over time if subtracting scores at NPA-I from NPA-II (Table 2). Hence, the positive differences indicate higher scores at NPA-II, and the negative differences indicate lower scores at the second assessment.

Figure 1: The descriptive bar graphs of WISC-R results of 21 patients and 21 siblings (>6 at NPA-I), at NPA-I and NPA-II with error bars indicating the 95% confidence interval

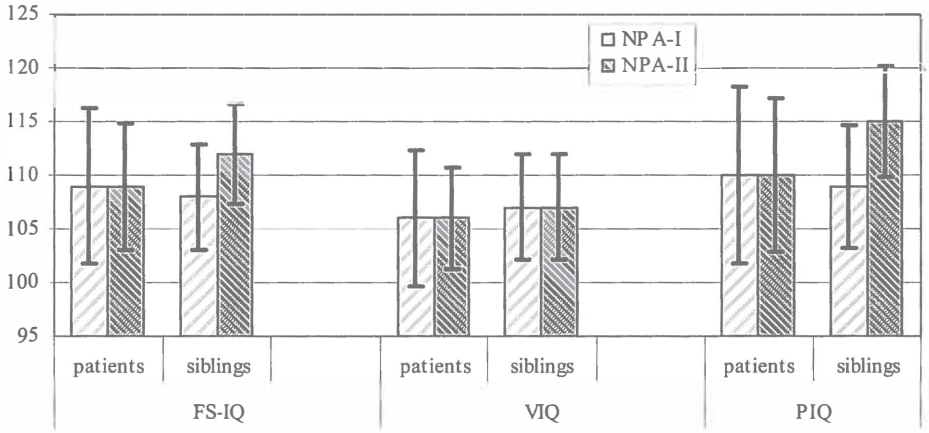


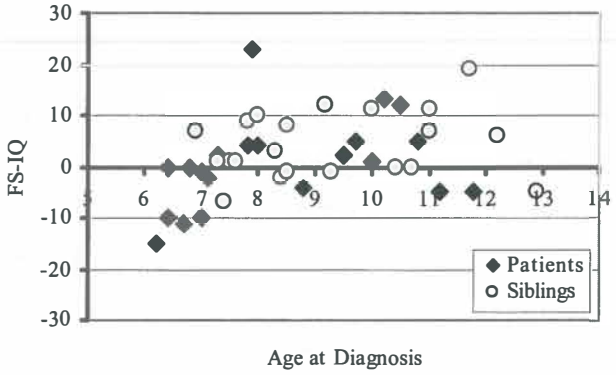
Table 2: Descriptive WISC-R difference-scores (NPA-II - NPA-I) in 21 patients and 21 siblings >6 at NPA-I (analysis 1)

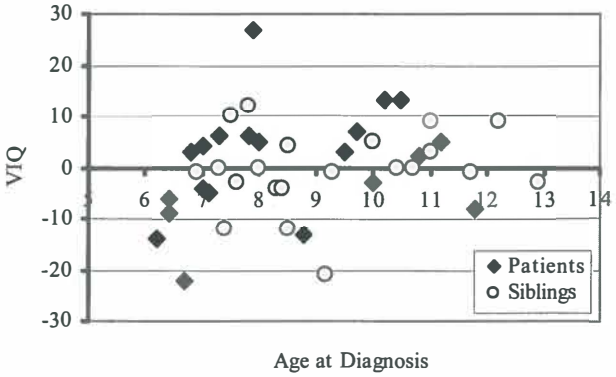
	Patients (N = 21)				Siblings (N = 21)			
	Mean	SD	95% CI		Mean	SD	95% CI	
			Low	High			Low	High
Full-scale IQ (FS-IQ)	.4	8.8	-3.6	4.4	4.3	6.5	1.3	7.2
Verbal IQ (VIQ)	.5	10.8	-4.5	5.4	-.5	7.9	-4.1	3.1
Performance IQ (PIQ)	.0	9.0	-4.0	4.1	5.9	8.9	1.9	10.0

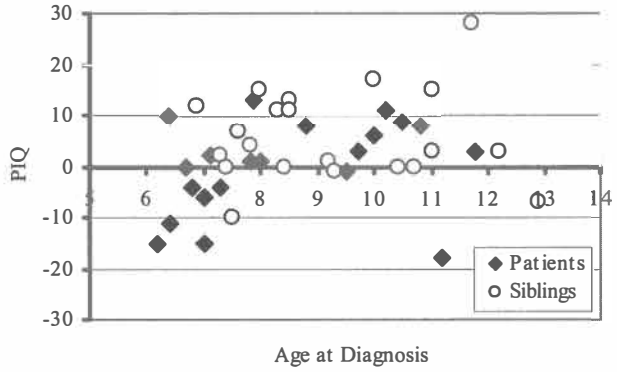
b. Scatter plots

Difference-scores for FS-IQ, VIQ and PIQ plotted against age mainly overlap in both groups (Fig. 2a). However, two young patients show a decline of > 10 IQ points at FS-IQ, versus none of the siblings (Fig. 2a). For VIQ, 3 patients and 3 siblings had a decline of > 10 IQ points. Fig. 2b shows that particularly older patients and siblings scored positive. Three young patients and one older patient against none of the healthy siblings show a decline of > 10 IQ points at PIQ (Fig. 2c). Fig. 2a and c indicate that difference-scores for FS-IQ and PIQ are mainly positive, except for younger patients, who had slightly lower mean scores at NPA-II (Table 2). However, the number of young patients is small and there are even less young siblings; therefore, the possible age effect can not be precluded.

Figure 2 (a-c): Scatter plots of WISC-R difference scores (NPA-II - NPA-II) in 21 patients and 21 siblings >6 at NPA-I (analysis 1).







c. Hierarchical regression analysis

Additionally, hierarchical regression analysis was conducted to compare the difference-scores for patients and siblings to detect the changes over time (Table 3). In the first step of hierarchical regression analysis, the age was included to control for a possible confounding effect and in the second step, the group factor was added to determine if this would lead to a meaningful change in the amount of explained variance (ΔR^2). In the third step, the interaction between age and group factor was included.

Table 3 shows that the factor age explains 10% of the variance of changes in FS-IQ and 5% in VIQ and PIQ; thus, age is probably a significant predictor of changes in FS-IQ ($\Delta R^2 = 0.100$; $p = 0.041$) if doing the WISC-R for the second time. Adding group as explanatory variable to the model does not lead to an important change in the amount of explained variance. So, there are no indications that the patients scored significantly lower than siblings. The results of hierarchical regression analysis do not demonstrate an interaction between age and group.

Table 3: Comparison of WISC-R difference-scores (NPA-II –NPA-I) in 21 patients and 21 siblings (> 6 at NPA-I) by hierarchical regression (analysis 2)

	df change		**Full scale-IQ (FS-IQ)**		**Verbal IQ (VIQ)**		**Performance IQ (PIQ)**	
Model			ΔR2	*p*-value	ΔR2	*p*-value	ΔR2	*p*-value
Age	1	40	.100	.041	.051	.152	.054	.137
Age, Group	1	39	.032	.235	.012	.486	.073	.079
Age, Group, Interaction Age and Group.	1	38	.014	.438	.010	.516	.016	.410

2. Longitudinal analysis (analysis 2; subjects < 6 years at NPA-I)

The data of a subgroup of children who were < 6 years and who had been tested with the WPPSI at NPA-I and with the WISC-R at NPA-II were analysed, to investigate the possible group × age interaction effect as mentioned before. Given the small sample size, only descriptive statistics will be given of this group.

Table 4 presents mean IQ-scores for patients and siblings at NPA-I (WPPSI-R) and NPA-II (WISC-R).

The performances of both patients and siblings remained average over time. However, the tests results of WPPSI-R and WISC-R cannot be compared, which will be discussed later.

Table 4: Mean scores of patients and siblings at NPA-I (WPPSI-R) and NPA-II (WISC-R) (analysis 2)

	Patients (N = 11)				**Siblings (N = 6)**			
	Mean NPA-I (WPPSI-R)	SD	Mean NPA-II (WISC-R)	SD	Mean NPA-I (WPPSI-R)	SD	Mean NPA-II (WISC-R)	SD
Full scale-IQ (FS-IQ)	113.2	16.6	110.1	8.9	101.5	7.3	105.2	9.1
Verbal-IQ (VIQ)	115.6	15.5	115.5	9.9	99.5	8.4	103.5	8.4
Performance-IQ (PIQ)	107.1	13.1	101.7	10.8	106.0	19.6	105.7	11.7

3. Cross-sectional, post-treatment analysis (analysis 3; all ages)

a. Descriptive results

Table 5 offers WISC-R results (FS-IQ, VIQ, PIQ) of patients and siblings after cessation of treatment. IQs are high average for both groups on the WISC-R in comparison to population norms.

Table 5: Cross-sectional results of WISC-R at NPA-II in 43 patients and 27 siblings (all ages) after cessation of therapy (analysis 3)

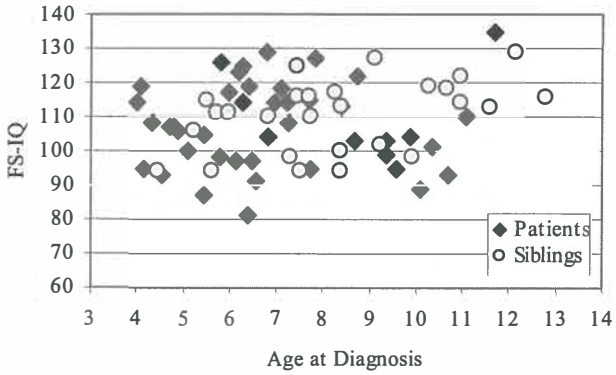
	Patients (N = 43)				Siblings (N = 27)			
	Mean	SD	95% CI		Mean	SD	95% CI	
			Low	High			Low	High
Full-scale IQ (FS-IQ)	107.4	12.5	103.6	111.3	110.4	10.5	106.3	114.6
Verbal IQ (VIQ)	108.1	11.5	104.6	111.7	106.0	10.5	101.8	110.1
Performance IQ (PIQ)	104.8	14.8	100.3	109.4	113.0	12.3	108.1	117.9

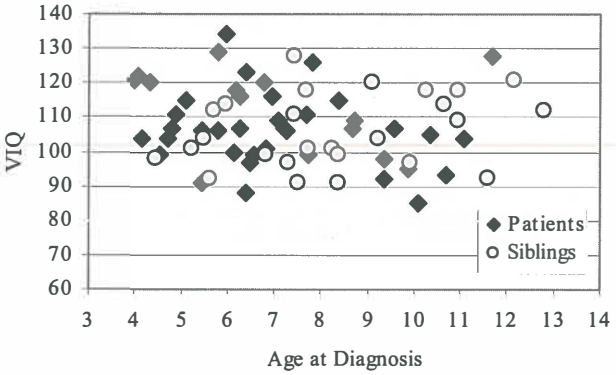
WISC-R; VIQ (Information, Similarities, Arithmetic, Vocabulary, Digit Span), PIQ (Picture Completion, Block Design, Object Assembly, Coding, Mazes); (CI=Confidence Interval)

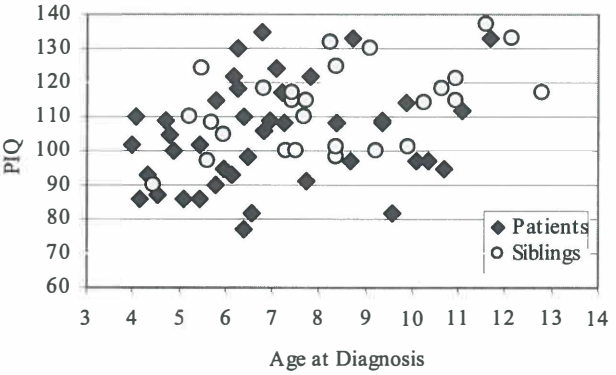
b. Scatter-plots

The scatter plots with IQ's for patients and siblings plotted against age indicate that scores in both groups mainly overlap (Fig. 3). As shown in Fig. 3, 3 patients have a FS-IQ < 90, whereas none of the siblings scored < 90 (Fig. 3a). For VIQ, 2 patients and none of the healthy sibling scored < 90 (Fig. 3b). For PIQ, 6 younger patients (younger than 8 years at diagnosis) and 1 older patient (older than 8 years at diagnosis) against none of the healthy siblings score < 90 (Fig. 3c). These findings support a group × age interaction effect for PIQ, indicating that younger patients have poorer performance. But again, this applies to a small number of patients.

Figure 3 (a-c): The cross-sectional scatter plots of WISC-R at NPA-II in 43 patients and 27 siblings (all ages) after cessation of therapy (analysis 3)







c. Hierarchical regression analysis

Hierarchical regression analysis was conducted to compare performances of patients to siblings on WISC-R (Table 6). In the first step of the hierarchical analysis, age was included as a control variable. In the second step, test-shift was added to the regression analysis to find out if this factor leads to a meaningful change in the amount of explained variance (ΔR^2), which will be discussed later. In the third step, the group factor was added. In the fourth step, the interaction between age and the group factor was included. As Table 6 shows, the factor age explained a low amount of variance, except for PIQ in which age explains a moderate but significant amount of variance ($\Delta R^2 = .142$; $p = .001$). By adding group, no significant effects were found; thus, no indications were found that the patients scored lower than the siblings. For VIQ, a significant group × age interaction effect was detected ($\Delta R^2 = .056$; $p = .049$), with the younger patients scoring higher than the older patients and all siblings. For PIQ and FS-IQ, no significant effects were found.

Table 6: Cross-sectional comparison of WISC-R results at NPA-II of 43 patients to 27 siblings (all ages) by hierarchical regression analysis after cessation of therapy (analysis 3)

	df change		**Full scale-IQ (FS-IQ)**		**Verbal IQ (VIQ)**		**Performance IQ (PIQ)**	
Model			ΔR2	*p*-value	ΔR2	*p*-value	ΔR2	*p*-value
Age	1	68	.038	.106	.005	.545	.142	.001
Age, Test-shift	1	67	.001	.789	.021	.228	.004	.559
Age, Test-shift, Group	1	66	.005	.576	.009	.433	.034	.102
Age, Test-shift, Group, Interaction Age and Group	1	65	.031	.144	.056	.049	.007	.473

Discussion

The purpose of the present study was to investigate the effects of chemotherapy-only on intelligence by means of the most optimal study design and sophisticated statistical procedures. The highlights of our study include a prospective design with a sibling control-group, strict methodology, little loss of patients during follow-up, a homogeneous patient group and a single neuropsychologist who did all the assessments.

The main conclusion of the longitudinal and post-treatment analysis is that there are no differences between patients and siblings. Moreover, both groups scored high average at both assessments compared to population norms. This could be a result of the Flynn effect, accounting for an IQ rise of about 6 points since test norms were collected in the early 80s.[24] If evaluated with more recent norms, these children would probably have average results.

As shown by the more extended hierarchical regression analysis and scatter-plots, the patients > 8 years at diagnosis did not decline on intelligence over time. It seems that the group of younger patients (< 8 years at diagnosis) might slightly decline on PIQ compared to older patients and the group of healthy siblings although not significant.

Our results agree with the longitudinal studies of Kingma and colleagues, Copeland and colleagues and Brown and colleagues who found no major differences between patients and controls on intelligence tests.[11,13,25] However, recent findings of the retrospective study by Montour-Proulx and colleagues, showed a significant decline in PIQ.[26] Different outcome can be explained by variability in interval and frequency of assessments, and the lack of a control group. Moreover, their patient group was younger than our group, but results of their's and our study are not contradictory considering the youngest age groups.

There are two possible explanations for specifically young patients showing a relative decline in PIQ compared to siblings. A first explanation is that both the older patients and the siblings profit more from earlier testing than younger children.[27] Theoretically, practise effect could be stronger for older children than for the younger children. Secondly, younger patients might score lower on PIQ at follow-up than their siblings as a true consequence of the higher susceptibility for the negative effects of chemotherapy in the immature brain.[28] Lastly, a general explanation for the differences

between younger and older patients in our study could be sample fluctuation, with no relation to treatment effect. The true meaning of a possible difference between the young and old patients cannot fully be established because of a relative lack of young siblings in the present study. Given the peak incidence of ALL (3-5 years), minimum age for the Wechsler scales (4 years) and the fact that the average number of children in Dutch families is < 2, this problem can not be solved. To establish a larger study population it would be interesting to run a multi-national study, however, differences in test versions (languages) and treatment protocols among countries would make comparisons difficult.

A few other remarks on methodology have to be made. First, the number of patients who refused to participate at NPA-I, and hence for the longitudinal comparison, could possibly yield bias in these study results. However, included patients did not differ in respect to age, sex, and disease characteristics from missing patients. Moreover, our percentage of missing patients is generally accepted for this kind of research.[29] Another group of patients was never referred, unrelated to patients' or parents' characteristics. Second, a number of patients (N = 11) could not complete the full intelligence test at NPA-I and these patients might have profited less of earlier testing at the second assessment. To preclude any bias, we excluded these cases for the longitudinal analyses. Last, inevitable test-shift from the WPPSI-R to the WISC-R, as a consequence of the limited age range of intelligence tests, may have influenced the outcome for the younger group. Tests scores of WPPSI-R and WISC-R can not be compared, given different statistical properties and norm-scores. Bos and De Sonneville found a difference of approximately 7.5 IQ points between older versions of WPPSI and WISC but their findings could not be used in our study.[30]

In summary, children with ALL treated with chemotherapy-only have normal intellectual functioning after cessation of two-year' intensive chemotherapeutic treatment including systemic and intrathecal MTX. Our findings also suggest that most children do not decline intellectually but it can not be precluded that a few individual cases might deteriorate on PIQ. Further research must reveal whether the subtle differences found in young patients are temporary or will persist or even increase either by global damage or selective impairment to the brain. In clinical practice, it is important to monitor those individual patients who suffer from intellectual impairment to further improve their quality of live and to develop support programs. For future research it remains a challenge to

elucidate the phenomenon of chemotherapy induced impairment in some individual patients who might have a higher sensibility for negative sequelae.

Acknowledgements

This study is a collaborative multicentre investigation which is coordinated by the Paediatric Oncology Centre, Groningen University Hospital. Participating centres: University Medical Centre Utrecht, Utrecht; St Radboud University Medical Centre, Department of Paediatrics, Nijmegen; Vrije Universiteit University Medical Centre, Department of Paediatrics, Amsterdam; University Medical Centre, Emma Children's Hospital, Amsterdam; University Medical Centre, Willem-Alexander Children and Youth Centre, Leiden. This study was supported by a grant from the Dutch Cancer Society.
The authors would like to thank the children and families who participated in this study.

References

1. Pui CH, Relling MV, Downing JR. Acute lymphoblastic leukemia. N Engl J Med 2004;350:1535-1548.
2. Brenner H, Kaatsch P, Burkhardt-Hammer T, Harms DO, Schrappe M, Michaelis J. Long-term survival of children with leukemia achieved by the end of the second millennium. Cancer 2001,92:1977-1983.
3. Spiegler BJ, Bouffet E, Greenberg ML, Rutka JT, Mabbott DJ. Change in neurocognitive functioning after treatment with cranial radiation in childhood. J of Clin Onc 2004;22:706-713.
4. Meadows AT, Gordon J, Massari DJ, Littman P, Ferguson J, Moss K. Declines in IQ scores and cognitive dysfunctions in children with acute lymphocytic leukaemia treated with cranial irradiation. Lancet 1981;2:1015-8.
5. Jankovic M, Brouwers P, Valsecchi MG, et al. Association of 1800 cGy cranial irradiation with intellectual function in children with acute lymphoblastic leukaemia. Lancet 1994;344:224-27.
6. Langer T, Martus P, Ottensmeier H, Hertzberg H, Beck JD, Meier W. CNS late-effects after ALL therapy in childhood. Part III: neuropsychological performance in long-term survivors of childhood ALL: impairments of concentration, attention, and memory. Med Pediatr Oncol 2002;38:320-8.
7. Butler RW, Hill JM, Steinherz PG, Meyers PA, Finlay JL. Neuropsychologic effects of cranial irradiation, intrathecal methotrexate, and systemic methotrexate in childhood cancer. J Clin Oncol 1994;12:2621-9.
8. Moleski M. Neuropsychological, neuroanatomical, and neurophysiological consequences of CNS chemotherapy for acute lymphoblastic leukaemia. Arch Clin Neuropsy 2000;15:603-630.
9. Moore BD 3rd. Neurocognitive outcomes in survivors of childhood cancer. J Pediatr Psychol 2005;30:51-63.
10. Kingma A, Van Dommelen RI, Mooyaart EL, Wilmink JT, Deelman BG, Kamps WA. Slight cognitive impairment and MRI abnormalities but normal academic career in children treated for acute lymphoblastic leukemia with chemotherapy only. J Pediatr 2001;3:413-420.
11. Giralt J, Ortega JJ, Olive T, Verges R, Forio I, Salvador L. Long-term neuropsychologic sequelae of childhood leukemia: comparison of two CNS prophylactic regimens. Int J Radiat Oncol Biol Phys 1992;24:49-53.
12. Copeland DR, Moore BD 3rd, Francis DJ, Jaffe N, Culbert SJ. Neuropsychologic effects of chemotherapy on children with cancer: a longitudinal study. J Clin Oncol 1996;14:2826-2835.
13. Brown RT, Madan-Swain A, Pais R, Lambert RG, Sexson S, Ragab A. Chemotherapy for acute lymphocytic leukemia: Cognitive and academic sequelae. J Pediatr 1992;121:885-889.
14. Eiser C, Tillmann V. Learning difficulties in children treated for acute lymphoblastic leukaemia (ALL). Pediatr Rehabil 2001;4:105-18.
15. Kingma A, Van Dommelen RI, Mooyaart EL, Wilmink JT, Deelman BG, Kamps WA. No Major Cognitive Impairment in Young Children With Acute Lymphoblastic Leukemia Using Chemotherapy Only: A Prospective Longitudinal Study. J Pediatr Hematol Oncol 2002;24:106-113.
16. Raymond-Speden E, Tripp G, Lawrence B, Holdaway D. Intellectual, neuropsychological, and academic functioning in long-term survivors of leukemia. J Pediatr Psychol 2000;25:59-68.
17. Jansen NC, Kingma A, Tellegen P, et al. Feasibility of Neuropsychological Assessment in Leukemia Patients Shortly After Diagnosis: Directions For Future Prospective Research. Arch Dis Child 2005;90:301-304.

18. Kamps WA, Veerman AJ, van Wering ER, van Weerden JF, Slater R, van der Does-van den Berg A. Long-term follow-up of Dutch Childhood Leukemia Study Group (DCLSG) protocols for children with acute lymphoblastic leukemia, 1984-1991. Leukemia 2000;14:2240-2246.
19. Veerman AJ, Hahlen K, Kamps WA, et al. High cure rate with a moderately intensive treatment regimen in non-high-risk childhood acute lymphoblastic leukemia. Results of protocol ALL VI from the Dutch Childhood Leukemia Study Group. J Clin Oncol 1996;14:911-918.
20. Kamps WA, Bokkerink JP, Hakvoort-Cammel FG, et al. BFM-oriented treatment for children with acute lymphoblastic leukemia without cranial irradiation and treatment reduction for standard risk patients: results of DCLSG protocol ALL-8 (1991-1996). Leukemia 2002;16:1099-111.
21. Vander Steene G, Bos A: Wechsler Preschool and Primary Scales of Intelligence-revised, Vlaams-Nederlandse aanpassing (Dutch manual). Lisse, The Netherlands: Swets & Zeitlinger, 1997.
22. Vander Steene G, van Haasen PP, de Bruyn EEJ, et al: Wechsler Intelligence Scale for Children-revised (Dutch manual). Lisse, The Netherlands: Swets & Zeitlinger, 1986.
23. Stevens, J: Applied Multivariate Statistics for the Social Sciences (3rd ed.), Hillsdale, NJ: Lawrence Erlbaum Associates, 1996.
24. Flynn JR. Massive IQ gains in 14 nations: what IQ tests really measure. Psychol Bull 1984;101:171-191.
25. Brown RT, Sawyer MB, Antoniou G, et al. A 3-year follow-up of the intellectual and academic functioning of children receiving central nervous system prophylactic chemotherapy for leukemia. J Dev Behav Pediatr 1996;17:392-398.
26. Montour-Proulx I, Kuehn SM, Keene DL, et al. Cognitive changes in children treated for acute lymphoblastic leukemia with chemotherapy only according to the Pediatric Oncology Group 9605 protocol. J Child Neurol 2005;20:129-33.
27. Vanderploeg, RD: Clinician's guide to neuropsychological assessment, Hillsdale, NJ: Lawrence Erlbaum Associates, 1994.
28. Ochs J, Mulhern RK, Fairclough D, et al. Comparison of neuropsychologic functioning and clinical indicators of neurotoxicity in longterm survivors of childhood leukemia given cranial irradiation of parenteral methotrexate: a prospective study. J Clin Oncol 1991;1:145-151.
29. Waber DP, Silverman LB, Catania L, et al. Outcomes of a Randomized Trial of Hyperfractionated Cranial Radiation Therapy for Treatment of High-Risk Acute Lymphoblastic Leukemia : Therapeutic Efficacy and Neurotoxicity. J Clin Oncol 2004;22:2701-2704.
30. Bos H, De Sonneville LMJ. De Nederlandse WISC-R. Kind Adolesc 1989;10:72-9.

CHAPTER 4

Preserved intellectual functioning in children treated for acute lymphoblastic leukaemia with chemotherapy-only

Nathalie CAJ Jansen, Annette Kingma, Arnout Schuitema, Anke Bouma, Anjo J Veerman, Willem A Kamps

Submitted

Abstract

Objective: To study intellectual functioning over time in children treated for acute lymphoblastic leukaemia (ALL) with chemotherapy-only.

Design: Nation-wide, prospective-longitudinal and sibling-controlled study.

Patients and methods: Forty-nine consecutive patients (Md age at diagnosis = 6.8 years; range 4.0-11.8) who were treated with multi agents systemic chemotherapy were repeatedly evaluated. Intellectual functioning (verbal (VIQ), performance (PIQ) and full-scale IQ (FS-IQ)) was assessed three times: within 2 weeks after diagnosis (T1), 3-6 months after cessation of 2 years therapy (T2) and two years after the second assessment, i.e. 4½ years after diagnosis (T3). Patients' performances were compared to performances of 29 healthy siblings (Md age at diagnosis = 8.2 years; range 4.5-12.6) and to normative data.

Results: Multilevel regression analyses indicated no major IQ differences between patients and siblings over time, with both groups mainly performing in the normal range. Although patients scored significantly lower compared to siblings on PIQ at T2, their score had normalized at T3. Additionally, patients who reported physical complaints (i.e. pain and/or tiredness) at the T1-evaluation scored significantly lower than older siblings on FS-IQ and PIQ after treatment at T2 and T3. Overall, significant practice effects were found for both patients and siblings on FS-IQ and PIQ.

Conclusions: Children treated for ALL with chemotherapy-only show normal intellectual functioning over time. However, patients who uttered physical symptoms at T1 show poorer performance on FS-IQ and PIQ compared to siblings. Therefore, it is important to continue monitoring these children to detect possible adverse effects on the long term.

Introduction

In children with acute lymphoblastic leukaemia (ALL), cranial irradiation (CI) as elective central nervous system (CNS) treatment has been associated with cognitive toxicity.[1-2] To minimize adverse cognitive late-effects, intrathecal and systemic chemotherapy were introduced, with equivalent success in the prevention of CNS leukaemia and relapse.[3]

It has generally been accepted that patients treated with chemotherapy-only protocols perform significantly better on cognitive measures compared to patients treated with CI and chemotherapy.[4-5] However, fewer studies have addressed the cognitive functioning of ALL-patients treated with chemotherapy-only in a longitudinal study design[6-10] and controlled studies are even scarcer[6-8]. Moreover, methodological differences and limitations have resulted in considerable indistinctness about the cognitive effects of chemotherapy-only protocols. Some longitudinal studies report that CNS directed chemotherapy is associated with slight impairment in verbal IQ[6,9], but other longitudinal studies could not confirm such deterioration.[7,10-12]

Therefore, the effects of chemotherapy-only protocols on cognitive functioning in long-term cancer survivors remain controversial. The purpose of the present study was to investigate the late-effects of chemotherapy-only treatment of children with ALL on intelligence. It extends our previous 2½-year prospective-longitudinal study on intellectual functioning of children with ALL[8,13] to 4½-years after diagnosis. This extended follow-up is important to assess whether patients might deviate from normal intellectual development on the longer term.

Patients and methods

Patients and healthy siblings

Forty-nine consecutive newly diagnosed children with ALL between the ages of 4.0 and 12.6 years and Dutch as primary language were eligible. Patients with initial overt CNS leukaemia or pre-existent disorders that could interfere with normal cognitive development (developmental or psychiatric disorders or Down syndrome) were ineligible. Patients were

treated according to the national chemotherapy-only Dutch Childhood Oncology Group (DCOG) ALL-9 protocol, including systemic chemotherapy (vincristine, dexamethasone, L-asparaginase, medium dose methotrexate (MTX), 6-MP) and repeated triple IT (MTX, prednisone, ara-C) therapy as CNS prophylaxis (see Table 1 for details). This protocol is similar to the DCOG ALL-6 protocol.[14-16] High risk patients (16/49) received additional systemic chemotherapy including daunorubicin, cyclofosfamide and cytosine-arabinoside. Treatment for all patients was continued for 108 weeks.

Table 1: DCOG-ALL-9 protocol for children with acute lymphoblastic leukaemia: cumulative drug doses per m^2 body surface area in 108 weeks' total treatment duration; all administrations i.v. except oral dexamethasone, 6-mercaptopurine, methotrexate (in non-high risk patients) and triple intrathecal therapy (TIT).

Cytostatic drug	Non-high risk (n = 32)	High risk (n = 17)
Vincristine	68 (max. 85) mg	62 (max. 77) mg
Dexamethasone	1365 mg	1238 mg
l-asparaginase	24000 IU	114000 IU
6-mercaptopurine	17500 mg	24850 mg
Methotrexate	2100 mg (po)	1650 mg (iv)
High-dose methotrexate	3 (weekly) x 2000 mg	4 (biweekly) x 3000 mg
(Leucovorin rescue)	3 x 15 mg per course	3 x 15 mg per course
TIT age-adjusted	13 x	15 x
Daunorubicin		175 mg
Cytosine-arabinoside		1920 mg
Cyclophosphamide		1920 mg

Twenty-nine patients had a healthy sibling who could serve as a control with the same inclusion criteria as the patients with respect to age, language and normal development. If a patient had more than one eligible sibling, the child closest in age to the patient was chosen.

At T1, patients were evaluated shortly after diagnosis, within two weeks after start of the chemotherapeutic treatment.[13] The second evaluation (T2) was repeated three to six months after cessation of therapy, that is, 2.3 – 2.6 years after T1 (Md = 2.4 years).[9] The last follow-up (T3) was repeated 1.8-2.3 years after T2 (Md = 2.2 years), i.e. 4½ years after diagnosis.[17]

Reasons for attrition from T1 to T2 were relapse of ALL or death (n = 3), refusal to further participation (n = 2), or switch to another treatment protocol (n = 1). From T2 to

T3, 2 patients relapsed and 1 patient refused further participation. So, an aggregate group of 40/49 patients who were in complete continuous remission could be assessed three times. However, the results of all initial 49 patients could be analysed by using multilevel analyses.

Of the initial 29 siblings, 2 refused to participate at T2 and 1 agreed to participate again at T3.

Characteristics of patients and healthy siblings are given in Table 2. Parental informed consent was obtained according to institutional rules.

Table 2: Numbers and characteristics of patients and siblings

	Patients	**Siblings**
Base-line T1[13]	**49**	**29**
Median age at diagnosis (range)	6.4 (4.0-11.8)	8.2 (4.5-12.6)
Female, %	41%	62%
Attrition from T1 to T2	*6*	*2*
T2[9]	**43**	**27**
Median age at follow-up (range)	9.0 (6.4-14.1)	10.5 (6.9-15.2)
Female, %	40%	60%
Attrition from T2 to T3	*3*	*+1**
T3[17]	**40**	**28**
Median age at follow-up (range)	11.2 (8.2 – 16.3)	12.7 (9.1 – 17.0)
Female, %	40%	61%

*) one sibling decided to participate again at T3

Siblings' evaluations took place within four weeks after the patients' evaluations.[13] Patients and siblings were individually tested either at home or at the hospital but no difference in IQ-scores was found between these two sites. To optimise standardization, all participants were tested nationwide by one qualified child neuropsychologist.

Measures

The experimental Dutch-Flemish version of the Wechsler Pre-School and Primary Scales of Intelligence (WPPSI-R) was used for children aged 4-6 years at T1.[18] Later, it appeared that the norm scores of the WPPSI-R were unreliable, and was decided to exclude the WPPSI-R results. Therefore, smaller groups of both young patients and young siblings have no

IQ-score at T1. Children ≥ 6 years were tested with the Wechsler Intelligence Scale for Children-revised (WISC-R, Dutch version) (10 subtests; extrapolating for the subtest comprehension and picture arrangement).[19] At T2 and T3 all children were tested with the WISC-R.

Assessments included neuropsychological tests but these results are beyond the scope of this paper and have been be published separately.[17]

Statistics

Analyses were performed using three multivariate MLwiN analyses, which enables the use of all outcome measures for an individual patient at any assessment.[20] The following analyses were performed:

a. Multilevel models were constructed in which controlling variables were tested stepwise, analogous to hierarchical regression analysis. Controlling variables were respectively: (1) general practice effects (an improvement of performances as a result of repeated assessments), (2) age and gender, (3) age < 6 versus > 6 years at T1 as a dichotomous variable as not all children were old enough to perform the WISC-R at T1, and (4) "sick" i.e. uttering physical symptoms (i.e. pain and/or tiredness) at T1 yes/no. We previously found that sick patients scored lower at T1 if uttering physical symptoms only. Therefore, "sick" was included as a variable in the model. If the estimation of the regression weights of one of these factors in a specific step of testing was significantly larger than zero, the factor was kept in the model.

b. After controlling for these factors, differences between patients and siblings at every NPA were tested. After this, the interaction patients × < 6 years at T1 was tested.

c. For the descriptive analyses, patients and siblings were divided into 5 subgroups. Based on the analyses of the missing values, 5 subgroups were formulated.

1. The young patient group (< 6 years at T1; N = 16) was established for two reasons: young age has been considered as a risk factor for adverse cognitive sequalae[13, 21] and young children could not complete the WISC-R because of age restrictions. Therefore, they could not profit equally from practice effects at T2 and T3 compared with older children. Hence, the variable "< 6" was included in the model.

2. Patients > 6 years at T1 with missing values due to physical complaints including pain, sickness and tiredness were defined as the sick group (N = 13). Missing values occurred on the WISC-R and/or on neuropsychological measures (> 2 measures).[17] A few patients < 6 years also uttered physical symptoms at T1 in combination with missing values and lower scores but this subgroup was too small to analyze separately.
3. Patients > 6 years at T1 who could be assessed completely and who had no test shift considering their age, were defined as the older group (N = 20).
4. Healthy siblings were defined as young i.e. age < 6 at T1 (N = 6); like the young patients, they could not complete the WISC-R because of age restrictions.
5. The older sibling group (age > 6 at T1; N = 23) could be assessed fully at every assessment and encountered no test shift.

Significance levels were established at $p < .05$. Data were analysed using MlwiN.[22,23] For *a.* and *b.*, only main effects and significant confounding variables will be presented. With the descriptive statistics (c), the pattern from T1 to T3 is illustrated.

Results

Results for 3 IQ measures will be presented by *a)* testing possible confounding variables (practice effects, age, gender and being sick at T1) using MLwiN-analyses, *b)* further MLwiN-analyses testing possible significant differences between patients and siblings with respect to significant confounding variables and patients < 6 × time interactions, to get more insight into adverse effects of chemotherapy-only on IQ and *c)* graphical presentation of mean test-scores of the 5 subgroups according to line charts. Additionally, a scatter plot of FS-IQ at T3 will be presented for all subgroups for visual inspection.

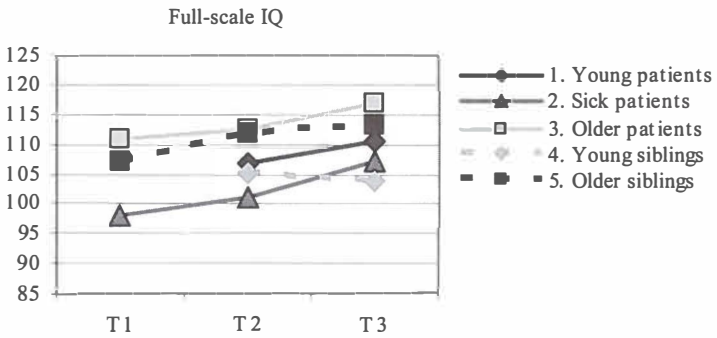
1. Full-scale IQ (FS-IQ)

1-a. Multilevel analyses of FS-IQ indicated that test scores of patients and siblings increased over time, probably as a consequence of general practice effects for the performance subtests. Being sick at NPA-I (sick patients scoring lower) was found to be a significant controlling variable.

1-b. No significant differences between patients and siblings were found after controlling for possible confounding effects. Multilevel analyses indicated no significant patients < 6 years × time interaction effects.

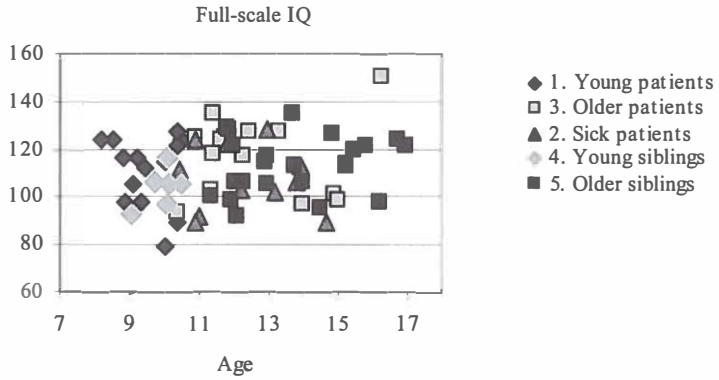
1-c. The pattern from T1 to T3 was analogous for all participants (Fig. 1a), except for young siblings. IQ scores for this subgroup slightly declined from T2 to T3. However, the number of young siblings is small. The scatter plot (Fig. 1b) does not indicate that IQ-scores of patients are clearly lower than siblings at T3. Nevertheless, 2 young patients and 2 sick patients versus none of the siblings have a FS-IQ < 90 at T3.

Figure 1a: Mean scores (mean, 50; standard deviation, 10) of FS-IQ for five subgroups at T1, T2 and T3.



Legend
T1 = First assessment, within 2 weeks after diagnosis
T2 = Second assessment; 3-6 months after cessation of 2 years therapy
T3 = Third assessment; two years after the second assessment, i.e. 4½ years after diagnosis

Figure 1b: Scatter plot of FS-IQ scores (WISC-R) at T3 (mean, 50; standard deviation, 10)
T3 = Third assessment; two years after the second assessment, i.e. 4½ years after diagnosis



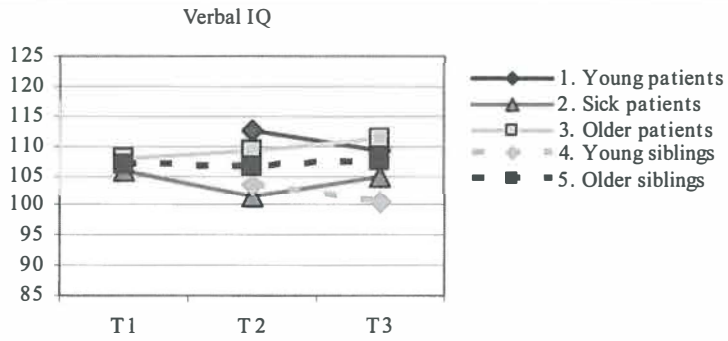
2. Verbal IQ (VIQ)

2-a. Multilevel analyses of VIQ scores did not show increase or decrease over time for both patients and siblings. Young age was found to be a significant confounding variable with higher scores at T2 and lower scores at T3.

2-b. No significant differences between patients and siblings were found after controlling for possible confounds. Multilevel analyses indicated no significant patients younger 6 × time interactions. At T3, no significant differences between patients and siblings were found.

2-c. The pattern from T1 to T3 is not equal for all subgroups with mean scores of both young patients and young siblings declining from T2 to T3 (Fig. 2). Scores of older patients slightly increased over time, while scores of older siblings remained stable. Scores of sick patients decreased from T1 to T2, but increased again at T3.

Figure 2: Mean scores (mean, 50; standard deviation, 10) of VIQ for five subgroups at T1, T2 and T3.



T1 = First assessment, within 2 weeks after diagnosis
T2 = Second assessment; 3-6 months after cessation of 2 years therapy
T3 = Third assessment; two years after the second assessment, i.e. 4½ years after diagnosis

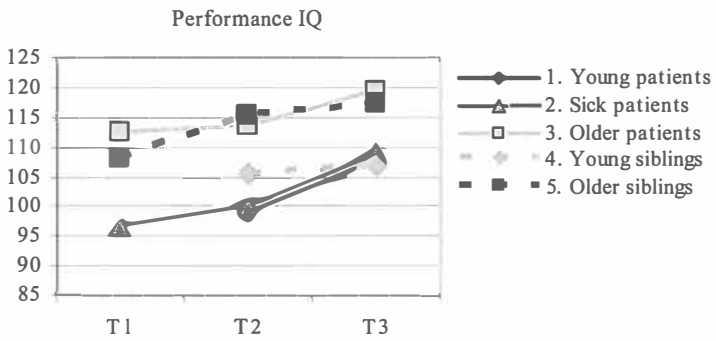
3. Performance IQ (PIQ)

3-a. Multilevel analyses of PIQ indicated that test-scores of patients and siblings increased over time, probably as a consequence of general practice effects. Age (with older children scoring higher) and being sick at T1 (with sick patients scoring lower) were found to be significant confounding variables.

3-b. Patients scored significantly lower at T2 compared to both sibling subgroups after controlling for possible confounds, but their scores normalized at T3. Multilevel analyses indicated no significant patients younger 6 × time interaction effects.

3-c. The pattern from T1 to T3 was analogous for all subgroups, except for young siblings (Fig. 3). IQ scores for this subgroup increased less from T2 to T3 compared with other subgroups. However, the number of young siblings is small.

Figure 3: Mean scores (mean, 50; standard deviation, 10) of PIQ for five subgroups at T1, T2 and T3.



T1 = First assessment, within 2 weeks after diagnosis

T2 = Second assessment; 3-6 months after cessation of 2 years therapy

T3 = Third assessment; two years after the second assessment, i.e. 4½ years after diagnosis

Discussion

In our 4½-year follow-up study multilevel analyses revealed that ALL-patients generally show normal intellectual functioning both compared with healthy siblings and normative data.

Our findings are consistent with the longitudinal, controlled study of Kingma in which no major IQ differences were found.[7] ALL-survivors showed intellectual performances within the normal range and comparable to those of healthy controls.[10] Other studies in contrast, did report adverse effects of chemotherapy-only on IQ.[6,9] Differences among studies might be explained by ignoring influence of practice-effects and test-shift in longitudinal designs. Also, the use of different or heterogeneous study group could cause distortion. Particularly, dose and duration of MTX courses and ITC therapy may determine the absence or presence of treatment-related cognitive deficits.[24]

Four findings in our study need special attention.

1) Lower scores were found on PIQ at the assessment after cessation of therapy. However, this could be explained by the relative low scores of young patients and patients who uttered physical symptoms at T1 (sick patients). These patients reduce the mean IQ of the whole patient group. Moreover, vincristine neuropathy may cause fine motor impairments and could therefore have an adverse effect on the scores of the performance subscale as it includes subtests with a time-factor.[9, 25]

2) Sick patients show less increase of scores on PIQ and hence FS-IQ on at T2 and T3. This finding could be explained by reduced mental reserve capacity as they still perform poorer on PIQ 4½ years later. Another explanation could be differences in practice effects as the PIQ is more susceptible for these effects.[26] Hypothetically, sick patients could have profited less of a general practice effect for the IQ test, following their poorer physical condition. Additionally, sick patients might indicate a risk group for true adverse effects of chemotherapeutic treatment. In our earlier study, we postulated that the brain of some individual patients might be more sensitive for adverse effects of chemotherapy.[9] Another explanation could have been that some individuals cope differently with physical symptoms, i.e. in times of stress. This could explain a lower IQ at T1 and T2 but is not in line with a lower PIQ at T3.

3) Both patients and siblings scored (high) average compared to population norms. This can be explained by the Flynn effect, accounting for an IQ rise of about 6 points since test norms were collected in the early eighties.[27] If evaluated with more recent norms most children would probably have average results. The addressed Flynn effect emphasizes the significance of including a healthy control-group with assessments in a fixed timeframe.

4) Last, significant differences were found for both patients and siblings on performance IQ and hence FS-IQ, due to practice effects. As stated earlier, this finding could be explained by the fact that PIQ is more susceptible for practice effects. Even with a 2-years retest interval, practice effects have been found.[26]

To avoid most methodological problems encountered in earlier published studies, we applied a prospective, longitudinal study-design with healthy siblings as controls. We have shown that intellectual assessment shortly after diagnosis is feasible and reliable, despite hospital admission and diagnosis with a life threatening disease.[13] We were able to look for several potential risk factors and therefore could avoid ignoring differences among subgroups that would bias results for all patients. Opposite to other studies, we did not exclude patients with incomplete data set. The use of multilevel techniques has also greatly improved the ability to conduct this type of research by enlarging the study group and reducing bias from missing data.

We acknowledge that the modest sample size is a possible confounder because small differences might not have been detected. The exclusion of the IQs of young patients at T1 as a consequence of unreliable norm scores of the WPPSI-R also reduced the power of our study. Finally, the number of young siblings in our study is small, therefore, no firm conclusions can be drawn on young patients.

In conclusion, children treated for ALL with chemotherapy-only generally showed normal intellectual functioning 4½ years after cessation of chemotherapy. However, sick patients show less increase of scores on performance IQ compared with healthy siblings. Therefore, it is important to continue monitoring these children to detect possible adverse effects on the long term.[24] Future studies with larger sample size are needed in which subgroups are differentiated with respect to missing data, to study late-effects.

Acknowledgements

This study is a collaborative multicentre investigation which is coordinated by the Paediatric Oncology Centre, Groningen University Hospital. Participating centres: University Medical Centre Utrecht, Utrecht; St Radboud University Medical Centre, Department of Paediatrics, Nijmegen; Free University Medical Centre, Department of Paediatrics, Amsterdam; University Medical Centre, Emma Children's Hospital, Amsterdam; University Medical Centre, Willem-Alexander Children and Youth Centre, Leiden.

The authors would like to thank the children and families for their contribution to this study and Wendy Post (PhD) for her statistical advice.

This study was funded by a grant from the Dutch Cancer Society.

References

1. Waber DP, Tarbell NJ. Toxicity of CNS prophylaxis for childhood leukemia. Oncology 1997;11:259-264
2. Butler RW, Hill JM, Steinherz PG, Meyers PA, Finlay JL. Neuropsychological effects of cranial irradiation, intrathecal methotrexate and systemic methotrexate in childhood cancer. J Clin Oncol 1994;12:2621-2629
3. Freeman A, Weinberg V, Brecher M. Comparison of intermediate-dose methotrexate with cranial irradiation for post-induction treatment of acute lymphoblastic leukemia in children. N Eng J Med 1983;308:477-481
4. Langer T, Martus P, Ottensmeier H, Hertzberg H, Beck JD, Meier W. CNS Late-Effects After ALL Therapy in Childhood. Part III: Neuropsychological Performance in Long-Term Survivors of Childhood ALL: Impairments of Concentration, Attention, and Memory. Med Pediatr Oncol 2002:38:320-328
5. Spiegler BJ, Kennedy K, Maze R, Greenberg ML, Weitzman S, Hitzler JK, Nathan PC. Comparison of Long-Term Neurocognitive Outcomes in Young Children With Acute Lymphoblastic Leukemia Treated With Cranial Radiation or High-Dose or Very High-Dose Intravenenous Methotrexate. J Clin Oncol 2006;24:3858-3864
6. Kingma A, Van Dommelen RI, Mooyaart EL, Wilmink JT, Deelman BG, Kamps WA. No Major Cognitive Impairment in Young Children With Acute Lymphoblastic Leukemia Using Chemotherapy Only: A Prospective Longitudinal Study. J Pediatr Hematol Oncol 2002;24(2):106-113.
7. Kingma A, Van Dommelen RI, Mooyaart EL, Wilmink JT, Deelman BG, Kamps WA. Slight cognitive impairment and magnetic resonance imaging abnormalities but normal school levels in children treated for acute lymphoblastic leukemia with chemotherapy only. J Pediatr 2001;139(3):413-420.
8. Jansen NC, Kingma A, Schuitema A, Bouma, Huisman J, Veerman AJ, Kamps WA. Post-treatment Intellectual Functioning in Children Treated For Acute Lymphoblastic Leukaemia (ALL) with Chemotherapy-only; a prospective, sibling-controlled study. Eur J of Cancer 42:2765-2772,2006
9. Ochs J, Mulhern R, Fairclough D, et al. Comparison of neuropsychologic functioning and clinical indicators of neurotoxicity in long-term survivors of childhood leukaemia given cranial radiation of parental methotrexate: a prospective study. J Clin Oncol 1991;9(1):145-151
10. Carey ME, Hockenberry MJ, Moore IM, et al. Brief report: Effect of intravenous methotrexate dose and infusion rate on neuropsychological function one year after diagnosis of acute lymphoblastic leukemia. J Pediatr Psychol 2006;32(2):189-93
11. Epsy KA, Moore IM, Kaufmann PM, et al. Chemotherapeutic CNS prophylaxis and neuropsychologic change in children with acute lymphoblastic leukemia: a prospective study. J Pediatr Psychol 2001;26(1):1-9
12. Moore IM, Epsy KA, Kaufmann P, et al. Cognitive consequences and central nervous system injury following treatment for childhood leukaemia. Semin Oncol Nurs 2000;16(4):279-290
13. Jansen NC, Kingma A, Tellegen P, et al. Feasibility of Neuropsychological Assessment in Leukemia Patients Shortly After Diagnosis: Directions For Future Prospective Research. Arch Dis Child 2005;90:301-304.
14. Kamps WA, Veerman AJ, van Wering ER, et al: Long-term follow-up of Dutch Childhood Leukemia Study Group (DCLSG) protocols for children with acute lymphoblastic leukemia, 1984-1991. Leukemia 14:2240-2246,2000

15. Veerman AJ, Hahlen K, Kamps WA, et al: High cure rate with a moderately intensive treatment regimen in non-high-risk childhood acute lymphoblastic leukemia. Results of protocol ALL VI from the Dutch Childhood Leukemia Study Group. J Clin Oncol 14:911-918,1996
16. Kamps WA, Bokkerink JP, Hakvoort-Cammel FG, et al: BFM-oriented treatment for children with acute lymphoblastic leukemia without cranial irradiation and treatment reduction for standard risk patients: results ofDCLSG protocol ALL-8 (1991-1996). Leukemia 16:1099-111,2002
17. Jansen NC, Kingma A, Schuitema A, Bouma A, Veerman AJ, Kamps WA. Neuropsychological Outcome in Chemotherapy-only Treated Children with Acute Lymphoblastic Leukemia (ALL). J Clin Oncol 20: , 2008
18. Vander Steene G, Bos A: Wechsler Preschool and Primary Scales of Intelligence-revised, Vlaams-Nederlandse aanpassing (Dutch manual). Lisse, The Netherlands: Swets & Zeitlinger, 1997
19. Vander Steene G, van Haasen PP, de Bruyn EEJ, et al: Wechsler Intelligence Scale for Children-revised (Dutch manual). Lisse, The Netherlands: Swets & Zeitlinger, 1986
20. Stevens, J: Applied Multivariate Statistics for the Social Sciences (3rd ed.), Hillsdale, NJ: Lawrence Erlbaum Associates, 1996
21. Mulhern RK, Kovnar E, Langston J, et al: Long-term survivors of leukemia treated in infancy: factors associated with neuropsychologic status. J Clin Oncol 1:1095-1102,1992
22. Goldstein H, Rasbash J, Plewis I et al. A User's Guide to MlwiN. London: Institute of Education; 1998
23. MlwiN [program]. London: Institute of Education, University of London http://www.ioe.ac.uk/multilevel/index.html;2001.
24. Moleski M. Neuropsychological, neuroanatomical, and neurophysiological consequences of CNS chemotherapy for acute lymphoblastic leukemia. Arch Clin Neuropsychol 2000;15(7):603-30
25. Reinders-Messelink HA, Schoemaker MM, Hofte M, et al. Fine motor and handwriting problems after treatment for childhood acute lymphoblastic leukemia. Med Pediatr Oncol 1996;27:551-5
26. Tuma JM, Appelbaum AS. Reliability and practice effects of WISC-R IQ estimates in a normal population. Educ Psychol Meas 1980;40:671-678
27. Flynn JR. Massive IQ gains in 14 nations: what IQ tests really measure. Psychol Bull 1984;101:171-191.

CHAPTER 5

Neuropsychological outcome in chemotherapy-only treated children with acute lymphoblastic leukemia

Nathalie C.A.J. Jansen, Annette Kingma, Arnout Schuitema, Anke Bouma, Anjo Veerman, Willem Kamps

Journal of Clinical Oncology 2008;26:3025-3030

Abstract

Purpose: To evaluate neuropsychological functioning over time in children treated for acute lymphoblastic leukemia (ALL) with chemotherapy-only.

Patients and Methods: Forty-nine consecutive patients (median age at first assessment, 6.8 years; range, 4.0 to 11.8 years) treated with intrathecal and systemic chemotherapy were included in a nationwide, prospective-longitudinal, sibling-controlled study. Patients and siblings completed three extensive neuropsychological assessments: at diagnosis, 3 to 6 months after completion of (2-year) treatment and 4½ years after diagnosis. Assessments included measures of learning, memory, attention, speed, executive functioning, visual-constructive and fine-motor functioning. Multilevel analyses were applied to evaluate patients' performance over time and to compare patients to 29 siblings (median age of siblings at first assessment, 8.2 years; range, 4.5 to 12.6) and to normative data.

Results: No major differences were found in neuropsychological performance between patients and siblings, with both groups performing mainly in the normal range. The patient group as a whole, however, scored significantly lower than siblings on complex fine-motor functioning at the last evaluation. Large practice effects were found for both patients and siblings in four of 11 tasks. Patients who uttered physical complaints (i.e. pain and/or tiredness) at the first pretreatment assessment scored significantly lower than siblings on attention and speed at the last two evaluations.

Conclusion: Despite intensive and potentially neurotoxic treatment, no evident negative, neuropsychological, late effects were found 4½ years after diagnosis, except for effects on complex fine-motor functioning. Both the large practice effects observed and the poorer performances on sustained attention for patients with physical complaints should be reckoned with in prospective, longitudinal neuropsychological research in children.

Introduction

Since the introduction of elective treatment of the CNS, the prognosis of children with acute lymphoblastic leukemia (ALL) has improved dramatically, and almost 80% will survive.[1] There are, however, questions about whether biological cure is accompanied by unaffected mental development and a good quality of life.

Early treatment protocols, including cranial irradiation (CI) in combination with intrathecal chemotherapy, have been associated with persistent cognitive impairment that results from structural brain damage.[2] In the mid-1980s, chemotherapy-only protocols, therefore, were introduced as standard treatment. To date, it is still debatable whether children with ALL who are treated with chemotherapy-only experience adverse neuropsychological late effects, but deficits generally seem mild in comparison to protocols that involve CI.[3-7] Intelligence, memory, and attention have been most studied. However, findings are conflicting, probably because of significant methodological differences among studies. A recent study showed no major intelligence quotient differences between patients and siblings over time after cessation of treatment; both groups performed mainly in the normal range.[8]

This is the first investigation to apply a prospective, longitudinal study design that assesses a wide variety of cognitive functions in patients with ALL compared with healthy sibling controls. Moreover, by applying multilevel analyses, performances of patients with missing values can be included to reduce the risk of biased outcome.

Patients and methods

Patients and healthy siblings

The present study included 49 consecutive pediatric patients with ALL. Inclusion criteria were newly diagnosed children with high- or standard-risk ALL who were between 4.0 and 12.3 years old and who spoke Dutch as the primary language. Patients with initial CNS leukemia or pre-existent disorders that could interfere with normal cognitive development (e.g. developmental or psychiatric disorders or Down syndrome) were excluded. Written informed consent was obtained according to each hospital's rules.

Patients were treated with chemotherapy only according to the national Dutch Childhood Oncology Group-ALL-9 protocol, which included vincristine, dexamethasone, daunorubicine with high-dose methotrexate (MTX), leucovorin rescue, and triple intrathecal therapy as CNS prophylaxis.[9-11] The total duration of treatment was 108 weeks (cumulative doses of cytostatic drugs and leucovorin rescue are given in the Appendix Table A1).

The control group consisted of 29 healthy siblings who had the same inclusion criteria as the patients with respect to age and normal cognitive development. If there was more than one sibling, the child closest in age to the patient was chosen. Of the initial 29 siblings, two refused to participate at neuropsychological assessment (NPA)-II. One sibling agreed to participate again at NPA-III; hence, a total of 28 could be assessed. Demographic variables for patients and siblings are listed in Table 1. Full details of the group of children included at NPA-I are described in Jansen et al.[12]

Table 1: Characteristics of Patients and Siblings

	Patients	Siblings
Baseline NPA-I [12]		
No.	49	29
Age at diagnosis, years		
Median	6.4	8.2
Range	4.0-11.8	4.5-12.6
Female, %	41	62
Attrition from NPA-I to NPA-II, No.	6	2
NPA-II		
No.	43	27
Age at follow-up, years		
Median	9.0	10.5
Range	6.4-14.1	6.9-15.2
Female, %	40	60
Attrition from NPA-II to NPA-III, No.	3	+1
NPA-III		
No.	40	28
Age at follow-up, years		
Median	11.2	12.7
Range	8.2-16.3	9.1-17.0
Female, %	40	61

Abbreviation: NPA, neuropsychological assessment.

Base-line evaluation (i.e. NPA-I) took place shortly after diagnosis and within 2 weeks of the start of chemotherapeutic treatment.[12] The second evaluation (i.e. NPA-II) was carried out 3 to 6 months after cessation of therapy (i.e. 2.3 to 2.6 years [median, 2.4 years] after NPA-I).[8] The last follow-up (i.e. NPA-III) was performed 1.8 to 2.3 years (median, 2.2 years) after NPA-II. Between NPA-I and NPA-II, the number of patients declined because of relapse of ALL or death in three patients, refusal of continued participation in two patients, and a switch to another treatment protocol in one patient. Between NPA-II and NPA-III, two patients relapsed, and one patient refused continued participation. Consequently, 40 of 49 patients could be assessed three times. Siblings were assessed within 4 weeks of the patients' evaluations.

To optimize standardization, all participants nationwide were tested by one qualified child neuropsychologist. Patients and siblings were individually tested either at home or at the hospital. No difference in scores was found between these two sites.

Measures

Children's neuropsychological skills were individually evaluated with a broad battery of test instruments, including tests for learning and memory, sustained attention and speed, executive functioning, and visual-motor and fine-motor function. Children aged 4 to 6 years could not complete the full test battery because of the age restrictions of most tests. Neuropsychological tests generally apply to functions in various cognitive domains, but, for convenience, we grouped them according to their main function (Table 2). These particular tests measure key cognitive functions, are standardized, have adequate Dutch norms, and have the widest possible age range to avoid unnecessary test-shift. Most tests are widely used in pediatric neuropsychology. The exception is the Dutch Bourdon-Vos (BV) self-paced cancellation task that assesses sustained attention and speed of visual scanning. This test consists of rows of figures (i.e. groups of dots) with a designated target figure to be crossed out by the patient as accurately and quickly as possible.

Although the NPA included intelligence tests, these results are beyond the scope of this article, and results have been published previously.[8] No major intelligence quotient differences were found between patients and siblings over time, both groups performed mainly in the normal range.

Practice effects (i.e. an improvement in performance as a result of repeated assessments) affect the interpretation of memory tests, in particular[19,20]; therefore, two parallel versions were used.[13]

Fitting procedures were used to convert published normative data from discrete into continuous normative data (mean, 50; standard deviation, 10). Raw scores were standardized by means of these continuous normal scores to enable comparisons of standardized scores among patients of any specific age.[21,22]

Table 2: Neuropsychological measures

Neuropsychological domain	Age (years)*	Measure
Learning and Memory	6	Dutch version of Rey's Auditory-Verbal Learning Test: learning and recall[13]
	6	Rey-Osterrieth Complex Figure Test: delayed-recall[14]
Sustained attention and speed	6	Bourdon-Vos, self-paced, continuous-performance cancellation task: speed and accuracy[15]
Executive functioning	6	Wisconsin Card Sorting Test: errors and perseverations[16]
Visual-motor and fine-motor function	4	Beery Developmental Test of Visual-Motor Integration[17]
	6	Rey-Osterrieth Complex Figure Test: copy[14]
	5	Purdue Pegboard (PP): simple fine motor functioning (dominant) and higher-order fine motor functioning (assembly)[18]

* Minimum age of testing

Statistics

As outlined earlier, not all of the initial 49 patients and 29 siblings could complete all tests at every NPA. Most missing values occurred within the patient group at NPA-I as a consequence of pain, sickness or tiredness (n = 13). Also, more patients (n = 16) than siblings (n = 6) were younger than 6 years old at NPA-I and could not be tested completely because of the minimum age of 6 years for most tests.

Because of missing values, analyses were performed with multivariate multilevel analyses (MlwiN4; Institute of Education, University of London, London, England), which enable the use of all outcome measures for an individual patient over time (Table 2).[23-24] The following analyses were performed:

Multilevel models were constructed, in which controlling variables were tested stepwise, analogous to hierarchical regression analysis. Controlling variables were as follows: general practice effects as a consequence of repeated assessments, age and sex, age younger than 6 versus older than 6 years at NPA-I as a dichotomous variable, and the presence of physical complaints uttered at NPA-I. If the estimation of the regression weights of one of these factors in a specific step of testing was significantly greater than zero, the factor was kept in the model.

After analysis was controlled for these factors, differences between patients and siblings were tested at every NPA. After this, the interaction between the variables of patients age younger than 6 years at diagnosis was tested.

For the descriptive analyses, patients and siblings were divided into five subgroups, and these were defined according to age and the analyses of the missing values.

The young patient group (< 6 years at NPA-I; n = 16) was established for two reasons: young age has been considered as a risk factor for harmful neuropsychological sequelae[8,25] and young children could not complete the full test battery because of age restrictions of most tests. They could not, therefore, profit equally from practice effects at NPA-II and NPA-III compared with older children. A variable for age younger than 6 years was included in the model, thus precluding the inability to interpret correctly the possible effects of chemotherapy at NPA-II and NPA-III.

Patients older than 6 years at NPA-I who missed more than two tests because of physical complaints (including pain, sickness, and tiredness) or restrictions (such as an infusion in the dominant hand) were defined as the sick group (n = 13). It appeared that sick patients scored lower at NPA-I as a consequence of being sick. Therefore, being sick was included as a variable in the model, to preclude the inability to interpret correctly the possible effects of chemotherapy at NPA-II and NPA-III. The patient group younger than 6 years was too small to apply subanalyses for patients who uttered physical complaints at NPA-I.

Patients older than 6 years at NPA-I who, because of their ages, could perform all tests, were defined as the older group (n = 20).

Healthy siblings at NPA-I (n = 6) were defined as young if aged younger than 6; like the young patients, they could not complete all neuropsychological tests because of age restrictions.

The older sibling group (age > 6 at NPA-I; n = 23) could be assessed fully at every NPA and had no test shift.

Significance levels were established at $P < .05$. Data were analyzed with MlwiN (Institute of Education, University of London, London, England).[26] In multilevel model analysis and patient-sibling comparisons, only main effects and significant confounding variables are presented.

Extensive descriptions of separate multilevel analyses in the results section are beyond the scope of this article; therefore, analyses will be summarized.

Results

Results of assessments will be presented by testing possible confounding variables (practice effects, age, sex, and sickness at NPA-I) using MLwiN-analyses, additional MLwiN-analyses to test possible significant differences between patients and siblings with respect to significant controlling variables and to patients younger than 6 years × time interactions, and additional descriptive statistics of differences (not tested) between the five subgroups according to line charts of mean test-scores.

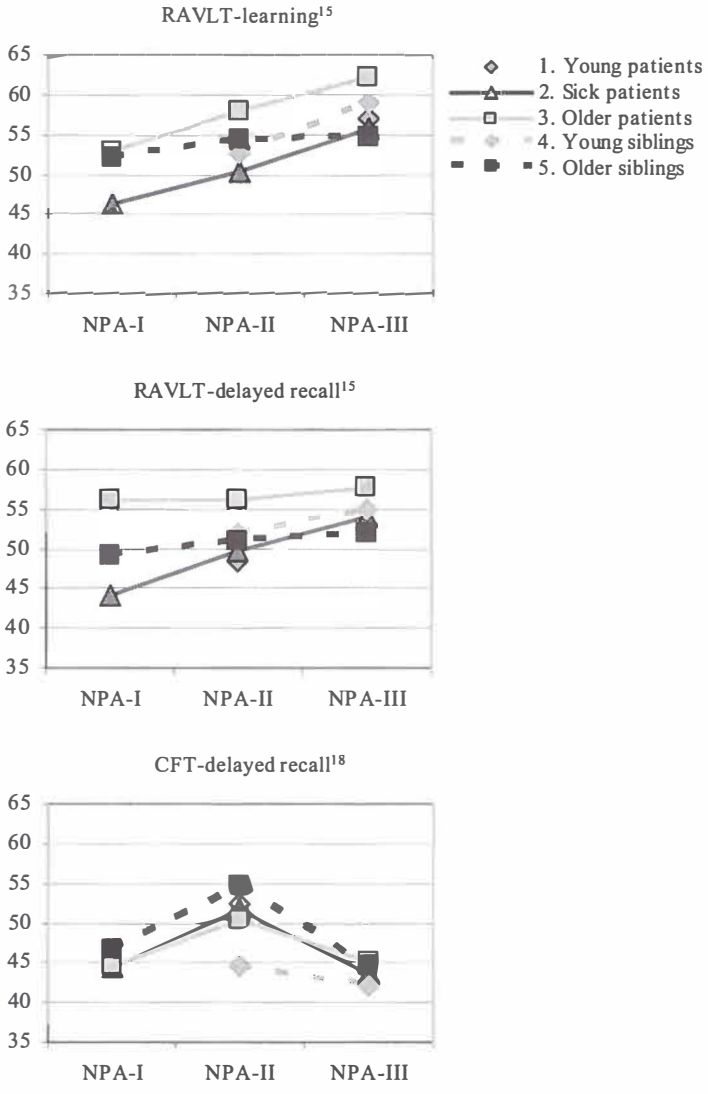
Learning and memory

Multilevel analyses of Rey Auditory-Verbal Learning Test (RAVLT) indicated that test scores of patients and siblings increased significantly over time, probably as a consequence of general practice effects. At NPA-II, mean scores of Rey-Osterrieth Complex Figure Test (CFT) delayed recall were significantly higher compared with both other assessments (Fig. 1c), which can be explained retrospectively by the use of an easier parallel test-version at NPA-II.[14] For RAVLT-learning, sex (i.e. girls scoring higher) and sickness at NPA-I (i.e. sick patients scoring lower) were found to be significant controlling variables.

No significant differences between patients and siblings were found after analysis was controlled for possible confounding effects. Multilevel analyses indicated no significant effects of the interaction of patients younger than 6 years × time.

In the three learning and memory tasks, the pattern between NPA-I and NPA-II was analogous for all subgroups (Fig. 1a to 1c). For RAVLT learning and recall, sick patients showed lower mean scores at NPA-I compared to older siblings, but their scores normalized at NPA-III. As Figure 1c suggests, CFT recall scores at NPA-III were equal for all subgroups.

Figure 1 a-b-c: Descriptive mean scores of five subgroups for learning and memory at three neuropsychological assessments (NPAs). RAVLT, Rey Auditory-Verbal Learning Test; CFT, Rey-Osterrieth Complex Figure Test.



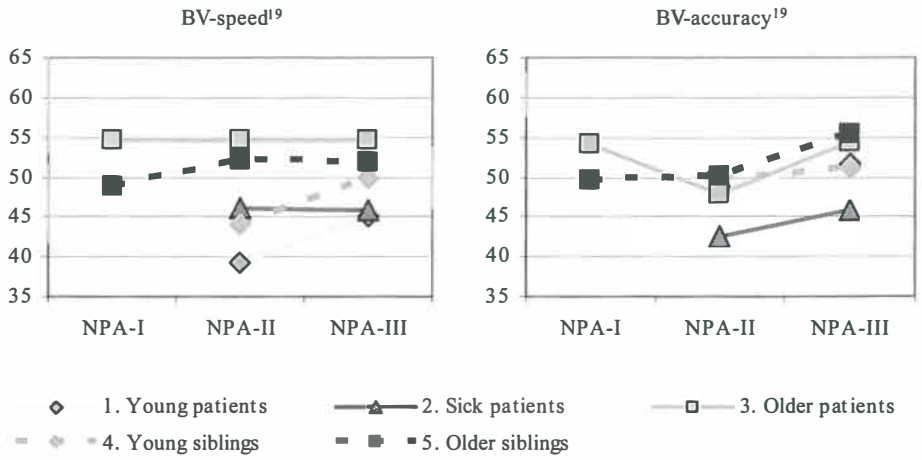
Sustained attention and speed

Multilevel analyses of the speed score of the Bourdon-Vos (BV) test did not indicate increase or decrease with time. Accuracy scores were significantly higher at NPA-III compared with NPA-II. Especially at NPA-II, but also at NPA-III, sick patients scored significantly lower compared with other subgroups on accuracy. Moreover, at NPA-II and NPA-III, young children scored significantly lower compared with other subgroups on speed.

For both speed and accuracy, after analysis was controlled for possible confounds, no significant differences between patients and siblings were found. Multilevel analyses did not indicate significant interactions for patients younger than 6 years × time.

Figures 2 a-b show descriptive mean scores for all subgroups at all three NPAs. Results of sick patients were not given at NPA-I, as there were only three patients who could complete the test at that time. In tests of both speed and accuracy, scores of sick patients were lower at NPA-II and NPA-III compared with older siblings. For speed, young patients scored lower compared with older siblings, but scores increased at NPA-III. Older patients showed a temporary decline in accuracy from NPA-I to NPA-II, but scores increased at NPA-III.

Figure 2 a-b: Descriptive mean scores of five subgroups for sustained attention and speed at three neuropsychological assessments (NPAs). BV, Bourdon-Vos.



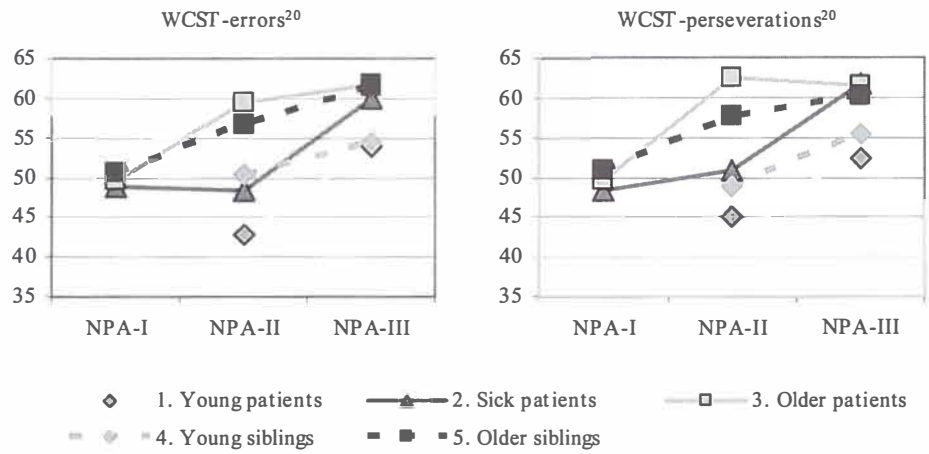
Executive functioning

Multilevel analyses of Wisconsin Card Sorting Test (WCST) errors and perseverations indicated a significant practice effect, as scores increased over time for all subgroups (Fig. 3a and 3b). For both measures, age was a significant confounding variable: older children scored higher. At NPA-II, young children scored significantly lower but performances normalized at NPA-III. Only in the WCST errors, sick patients scored significantly lower at NPA-II, but scores normalized at NPA-III.

For both executive functioning measures, after analysis was controlled for possible confounds, no significant differences between patients and siblings and no interactions of patients younger than 6 years × time were found.

Change in pattern of WCST errors between NPA-I and NPA-III is analogous for all subgroups; the lowest (i.e. unfavorable outcome) mean scores were at NPA-I, and the highest (i.e. favorable outcome) mean scores were at NPA-III. Sick patients scored lower at NPA-II compared with other subgroups, but scores increased strongly at NPA-III (Fig. 3a). Score obtained by young patients increase strongly from NPA-II to NPA-III.

Figure 3 a-b: Descriptive mean scores of five subgroups for executive functioning at three neuropsychological assessments (NPAs). WCST, Wisconsin Card Sorting Test.



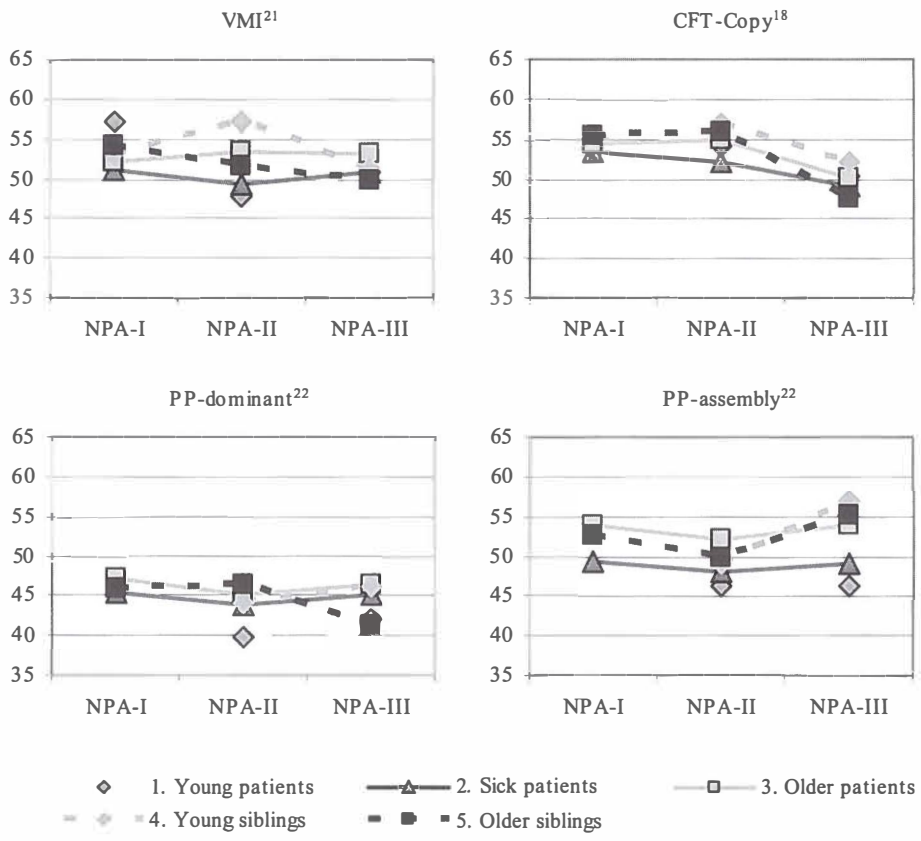
Visual-motor and fine-motor functioning

Multilevel analyses of visual-motor and fine-motor functioning tasks indicated a significant decrease of Beery Developmental Test of Visual-Motor Integration (VMI) scores at NPA-II. CFT copy scores decreased significantly with time at NPA-III. For both VMI and CFT copy, age was a significant confounding variable: younger children scored higher. For CFT copy and Purdue Pegboard (PP) dominant-hand tests, sex was a significant confounding variable: girls scored higher.

For both Rey copy and PP dominant hand tests, after analysis was controlled for possible confounding variables, no significant differences between patients and siblings or significant interactions between patient groups and age younger than 6 years at diagnosis were found. With PP assembly, there were significant differences between patients and siblings at NPA-III: patients scored lower. With VMI, there is a significant interaction effect between age and VMI: older patients scored lower at NPA-I than younger patients.

Except for VMI, pattern between NPA-I and NPA-II was analogous for all groups (Fig. 4a to 4d). No general trend could be detected. As mentioned for CFT delayed recall, CFT copy scores were higher at NPA-II because of the use of an easier parallel version of the test. With CFT copy, a decrease in score was seen; mean scores for all groups at NPA-III were lower compared with NPA-I. Older siblings had the lowest mean scores at NPA-III. For PP assembly, all patient subgroups showed lower mean scores at NPA-II and NPA-III compared with both sibling subgroups.

Figure 4 a-b-c-d: Descriptive mean scores of five subgroups for visual-motor and fine-motor functioning at three neuropsychological assessments (NPAs). VMI, Beery Developmental Test of Visual-Motor Integration; CFT, Rey-Osterrieth Complex Figure Test; PP, Purdue Pegboard.



Discussion

In this prospective, longitudinal study, a relatively large cohort of children with ALL who were treated with chemotherapy only was followed neuropsychologically until 4½ years after diagnosis.

In general, no major differences in neuropsychological performance were found between patients and siblings 2 years after cessation of therapy. Overall, our results are consistent with the few other published studies, which did not find evident, negative, neuropsychological late effects of chemotherapy.[27-29] The patient group as a whole, however, scored significantly lower than siblings on complex fine-motor functioning at the last evaluation. This relative deterioration can be explained by a minor deficit in processing speed[30], or it can be a result of vincristine neuropathy.[31] The latter explanation would, however, seem less likely, as patients scored well on a simple fine-motor task (dominant hand on the PP).

Some additional results need special attention. First, the sick group scored significantly lower than siblings on the accuracy scale of sustained attention at the second and third evaluation. It is noteworthy that patients who uttered physical complaints shortly after diagnosis still performed less well on a sustained attention task 4½ years later. We hypothesize that these patients have reduced cognitive reserve capacity. If true, the speed and attention task should be attempted at baseline; those patients who cannot complete it because of physical complaints are those who should be considered at risk for mild impairment. This finding then may apply beyond children with cancer and may be used for chronic, severely ill pediatric patients. Generally, speed and attention tasks are among the most sensitive for detecting acquired brain damage.[18] The prefrontal cortex is important in the development of attention, and this brain area is not fully developed until adulthood. It might, therefore, be more sensitive for the detection of neuropsychological sequelae in children. Moreover, the applied BV test could be a sensitive task for children with less cognitive reserve capacity because of its tedious nature. Other studies also found (subtle) late effects within the domain of attention and speed of information processing[30,32] or minor deficits on fine-motor functioning[31]; both apply to the performance on the BV test.

Secondly, large practice effects were found for both patients and siblings on certain tasks, which means that patients and siblings scored higher at the third evaluation

compared with the first two evaluations.[33] A control for these practice effects in longitudinal research by including a healthy control group is essential. Also, two parallel forms were used for the RAVLT to minimize practice effects. As potentially negative effects of chemotherapeutic treatment on neuropsychological functioning seem to be small, differences between patients and siblings may manifest only as differences in practice effects.[19,20] Also, practice effects should be taken into account in the interpretation of repeated individual assessments in clinical practice.

Sick patients showed a temporary, relative decline in executive functioning shortly after cessation of therapy, but scores normalized at the last evaluation. It is likely that some patients need time to recover; therefore, conclusions should not been drawn too soon after completing treatment.

Our study underscores the merit of analyzing missing values at the first pretreatment and at any other assessment. Unlike the more traditional repeated measurement models, data of patients with missing values can also be included in multilevel techniques. The excluding of these patients would have biased our results, as they were found to be at risk for mild attentional impairment.

Despite the application of strict methodology, our study has (insurmountable) limitations.[34] First, patients with incomplete data on some tests at the first pretreatment assessment (sick patients) generally had lower baseline scores on tests that they did complete than patients who had a complete data set. This could lead to underestimation of the deterioration of scores over time, had it been possible to test all patients fully. Second, numbers of patients included were too small for subset analysis of age. One cannot preclude a possible age effect if larger numbers of patient cases had been evaluated. Although no significant differences were found, there was a general trend, and young patients consistently scored less well.

In summary, patients who undergo potentially neurotoxic treatment generally do not show evident, negative, late effects of treatment on neuropsychological functioning 4½ years after diagnosis. We cannot, however, preclude that changes can still occur several years after cessation of treatment. Patients showed a slight decrease on complex fine-motor functioning compared with healthy siblings. Moreover, patients who could not been tested fully at diagnosis are at risk of slightly lower attention functioning. Additional research

should address the pathogenesis of these attention deficits in specific patients who are prone to negative sequelae.

Supplemental Table AI

DCLSG-ALL-9 protocol for children with acute lymphoblastic leukemia: cumulative drug doses per m^2 body surface area in 108 weeks' total treatment duration; all administrations i.v. except oral dexamethasone, 6-mercaptopurine, methotrexate (in non-high risk patients) and triple intrathecal therapy (TIT).

Cytostatic drug	Non-high risk (n = 32)	High risk (n = 17)
Vincristine	68 (max. 85) mg	62 (max. 77) mg
Dexamethasone	1365 mg	1238 mg
l-asparaginase	24000 IU	114000 IU
6-mercaptopurine	17500 mg	24850 mg
Methotrexate	2100 mg (po)	1650 mg (iv)
HD-methotrexate	3 (weekly) x 2000 mg	4 (biweekly) x 3000 mg
(Leucovorin rescue)	3 x 15 mg per course	3 x 15 mg per course
TIT age-adjusted	13 x	15 x
Daunorubicin		175 mg
Cytosine-arabinoside		1920 mg
Cyclophosphamide		1920 mg

Acknowledgment

This study is a collaborative multicenter investigation which is coordinated by the Pediatric Oncology Center, Groningen University Hospital. Participating centers: University Medical Center Utrecht, Utrecht; St Radboud University Medical Center, Department of Pediatrics, Nijmegen; Vrije Universiteit University Medical Center, Department of Pediatrics, Amsterdam; University Medical Center, Emma Children's Hospital, Amsterdam; University Medical Center, Willem-Alexander Children and Youth Center, Leiden.

We thank the children and their families for their contribution to this study. We are grateful to Dr. Wendy Post for her statistical advice and to G. Sinnema, PhD, for his cooperation while completing this report.

References

1. Pui CH, Evans WE: Treatment of acute lymphoblastic leukemia. N Engl J Med 354:166-178,2006
2. Pederson H, Clausen N: The development of cerebral CT changes during treatment of acute lymphocyctic leukemia in childhood. Neuroradiology 22:79-84,1981
3. Langer T, Martus P, Ottensmeier H, et al: CNS late-effects after ALL therapy in childhood. Part III: neuropsychological performance in long-term survivors of childhood ALL: impairments of concentration, attention, and memory. Med Pediatr Oncol 38:320-328,2002
4. Butler RW, Hill JM, Steinherz PG, et al: Neuropsychologic effects of cranial irradiation, intrathecal methotrexate, and systemic methotrexate in childhood cancer. J Clin Oncol 12:2621-2629,1994
5. Mulhern RK, Fairclough D, Ochs J: A prospective comparison of neuropsychologic nnperformance of children surviving leukemia who received 18-Gy, 24-Gy, or no cranial irradiation. J Clin Oncol 1:1348-1356, 1991
6. Kingma A, Van Dommelen RI, Mooyaart EL, et al: Slight cognitive impairment and MRI abnormalities but normal academic career in children treated for acute lymphoblastic leukemia with chemotherapy only. J Pediatr 3:413-420,2001
7. Waber, DP, Carpentieri SC, Klar N, et al: Cognitive sequelae in children treated for acute lymphoblastic leukemia with dexamethasone or prednisone. J Pediatr Hematol Oncol 22:206-213,2000
8. Jansen NC, Kingma A, Tellegen P, et al: Feasibility of neuropsychological assessment in leukemia patients shortly after diagnosis: Directions for future prospective research. Arch Dis Child 90:301-304,2005
9. Kamps WA, Veerman AJ, van Wering ER, et al: Long-term follow-up of Dutch Childhood Leukemia Study Group (DCLSG) protocols for children with acute lymphoblastic leukemia, 1984-1991. Leukemia 14:2240-2246,2000
10. Veerman AJ, Hahlen K, Kamps WA, et al: High cure rate with a moderately intensive treatment regimen in non-high-risk childhood acute lymphoblastic leukemia. Results of protocol ALL VI from the Dutch Childhood Leukemia Study Group. J Clin Oncol 14:911-918,1996
11. Kamps WA, Bokkerink JP, Hakvoort-Cammel FG, et al: BFM-oriented treatment for children with acute lymphoblastic leukemia without cranial irradiation and treatment reduction for standard risk patients: results of DCLSG protocol ALL-8 (1991-1996). Leukemia 16:1099-111,2002
12. Jansen NC, Kingma A, Schuitema A, et al: Post-treatment intellectual functioning in children treated for acute lymphoblastic leukemia (ALL) with chemotherapy only; a prospective, sibling-controlled study. Eur J Cancer 42:2765-2772,2006
13. Van der Burg W, Kingma A: Performance of 225 Dutch school children on Rey's Auditory Verbal Learning Test (AVLT): Parallel test-retest reliabilities with an interval of 3 months: normative data. Arch Clin Neuropsychol 14:545-559,1999
14. Lezak MD, Howieson DB, Loring DW: Neuropsychological assessment (ed 4). New York, Oxford University Press,2004
15. Bourdon-Vos test manual (ed 3). Lisse, The Netherlands: Swets & Zeitlinger, 1998

16. Chelune GJ, Baer RA: Developmental norms for the Wisconsin Card Sorting test. J Clin Exp Neuropsychol 8:219-28,1986
17. Beery KE: Revised administration, scoring and searching manual for the development test of visual-motor integration. Cleveland, Modern Curriculum Press, 1982.
18. Gardner RA, Broman M: The Purdue Pegboard; normative data on 1334 schoolchildren. J Clin Child Psychol 8:156-162,1979
19. Benedict RHB & Zgaljardic DJ: Practice effects during repeated administrations of memory tests with and without alternative forms. J Clin Exp Neuropsychol 20:339-352,1998
20. Spikman JM, Timmerman ME, Zomeren AH van, et al: Recovery versus retest effects in attention after closed head injury. J Clin Exp Neuropsychol 21:585-605,1999
21. Tellegen PJ, Laros JA: The construction and validation of a nonverbal test of intelligence: the revision of the Snijders-Oomen tests. Eur J Psychol Assess 9:147-157,1993
22. Snijders JTh, Tellegen PJ, Laros JA: SON-R 5½-17, Snijders-Oomen nonverbal intelligence test. Manual and research report. Groningen: Wolters-Noordhoff, 1989
23. Maas CJM, Snijders TAB: The multilevel approach to repeated measures for complete and incomplete data, Quality & Quantity 37:71-89,2003
24. Snijders TAB, Bosker, RJ: Multilevel Analysis, An introduction to basic and advanced multilevel modelling, London, Sage publications,1999
25. Mulhern RK, Kovnar E, Langston J, et al: Long-term survivors of leukemia treated in infancy: factors associated with neuropsychologic status. J Clin Oncol 1:1095-1102,1992
26. Goldstein H, Rasbash J, Plewis I, et al: A User's Guide to MLwiN. London, Institute of Education, 1998
27. Raymond-Speden E, Tripp G, Lawrence B, et al: Intellectual, neuropsychological, and academic functioning in long-term survivors of leukemia. J Pediatr Psychol 25:59-68,2000
28. Rodgers J, Marckus R, Kearns P, et al: Attentional ability among survivors of leukaemia treated without cranial irradiation. Arch Dis Child 88:147-50,2003
29. Spiegler BJ, Kennedy K, Maze R et al: Comparison of long-term neurocognitive outcomes in young children with acute lymphoblastic leukemia treated with cranial radiation or high-dose or very high-dose intravenous methrotrexate. J Clin Oncol 24:3858-3864,2006
30. Mennes M, Stiers P, Vandenbussche E, et al: Attention and information processing in survivors of childhood acute lymphoblastic leukemia treated with chemotherapy only. Pediatr Blood Cancer 43:1-8,2004
31. Gidding CE, Kellie SJ, Kamps WA et al: Vincristine revised. Critical reviews in oncology/hematology 29:267-287,1999
32. Buizer AI, De Sonneville LMJ, Van den Heuvel-Eibrink MM et al. Chemotherapy and attentional dysfunction in survivors of childhood acute lymphoblastic leukemia: effect of treatment intensity. Pediatr Blood Cancer 1-9,2005
33. Basso MR, Lowery N, Ghormley C, et al: Practice effects on the wisconsin card sorting test-64 card version across 12 months. Clin Neuropsychol 15:471-8,2001
34. Butler RW, Copeland DR: Neuropsychological effects of central nervous system prophylactic treatment in childhood leukemia: methodological considerations. J Pediatr Psychol 18:319-338,1993

CHAPTER 6

Cognitive functioning after chemotherapy-only in children with acute lymphoblastic leukaemia (ALL): a review, discussion and methodological issues.

Nathalie C. Jansen, Willem A. Kamps, Anke Bouma,
Aag Jennekens-Schinkel, Annette Kingma

Abstract

Purpose: To review the literature on cognitive functioning in children with ALL treated with chemotherapy-only (ChO), in relation to methodological variability and limitations.

Methods: A literature search and selection according to strict criteria yielded 26 original articles including 805 patients.

Results: Methods of studies and applied psychometric instruments widely varied. Results were categorized and discussed according to one or more cognitive domain(s), i.e. intelligence, attention, learning and memory, perceptual-spatial functioning, executive functioning, language, and fine motor functioning. Seven studies had a prospective and/or longitudinal design, with matched healthy controls in 3 of these 7 studies. These three reports failed to find evident adverse effects of ChO on cognition, apart from deterioration of complex motor functioning. The 4 remaining longitudinal reports generally showed slight decline for a variety of functions. Nineteen cross-sectional studies found a spectrum of neuropsychological deficits but had important methodological limitations. Most distinctive results were found for below average arithmetic skills. Risk factors for cognitive impairment, i.e. young age, female sex and higher MTX dose were partly investigated but no clear consensus emerged about an adverse effect on neuropsychological performance.

Conclusion: The great variety in instruments and methodology preclude definite and firm conclusions but ChO treatment was generally associated with good cognitive outcome with most patients functioning within normal limits.

Introduction

Modern multi chemotherapy has changed childhood acute lymphoblastic leukaemia (ALL) into a curable disease with survival rates up to 80%.[1,2] Prophylactic central nervous system (CNS) treatment has been a decisive factor for this strongly improved survival. In the past, CNS prophylaxis included cranial irradiation (CI) and intrathecal (IT) therapy, unfortunately resulting in cognitive impairment. Many reports showed significant decline of intelligence and specific neuropsychological deficits and academic problems.[3] As also other side-effects appeared and an alternative treatment with an equal survival-rate was developed, CI was eliminated from most treatment protocols and replaced by IT and high-dose (HD) systemic chemotherapy only (ChO).[4] Although ChO has undoubtedly less adverse effects than CI, the controversy stills remains whether children treated on ChO protocols have a cognitive development equal to healthy peers.

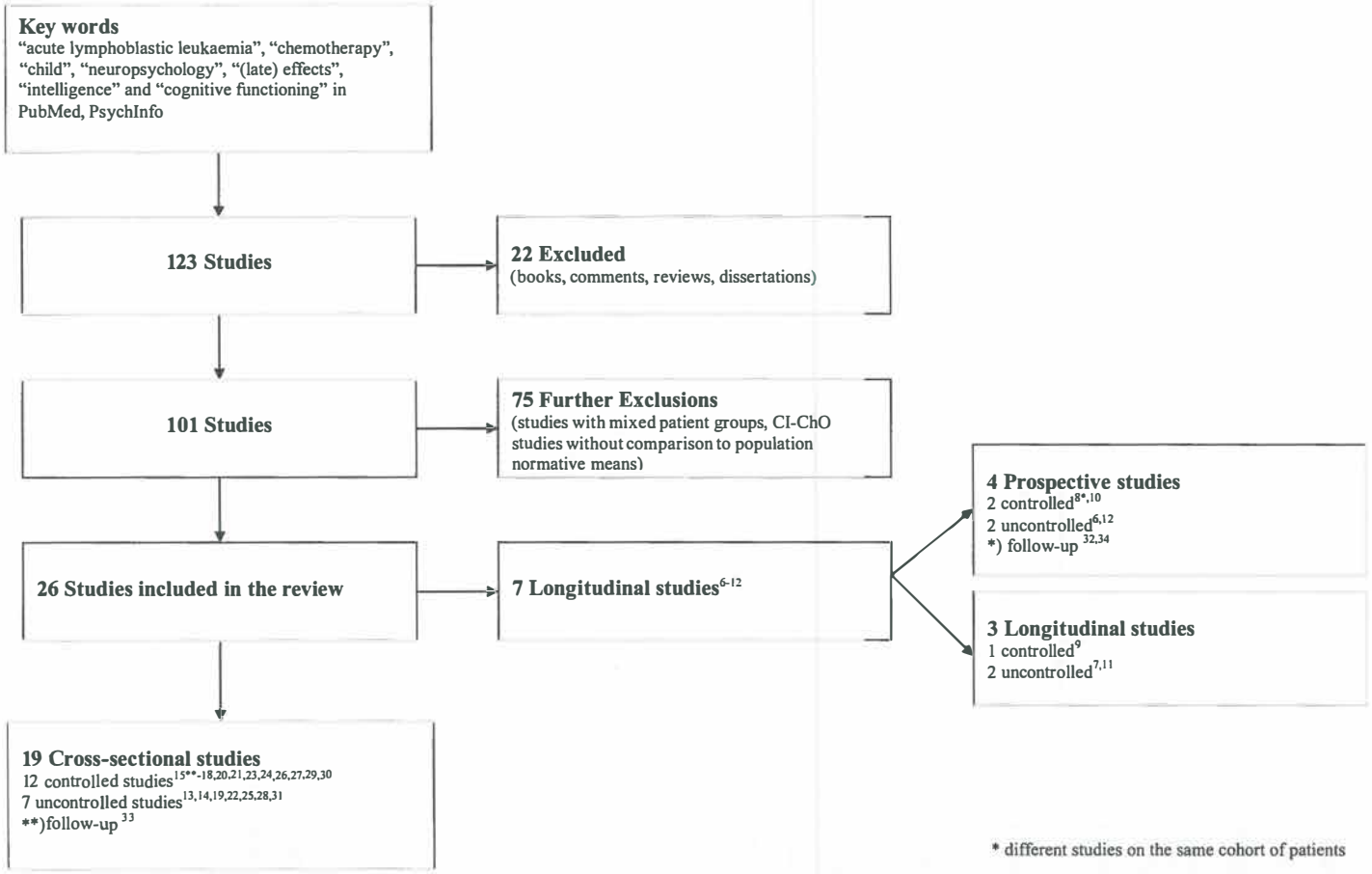
Therefore, the literature on cognitive functioning in children with ALL treated with ChO was reviewed. In relation to the reviewed literature, methodological issues and an optimal design to assess potential neuropsychological sequelae of ChO will be discussed.

Methods

Studies were retrieved from PubMed (National Library of Medicine, Bethesda, MD) and PsychInfo (American Psychological Association, Washington, DC) using (combinations of) the keywords: "acute lymphoblastic leukaemia", "chemotherapy", "child", "neuropsychology", "(late) effects", "intelligence" and "cognitive functioning". Additionally, reference lists of the retrieved studies were examined. Studies were selected according to the following criteria: (1) regarding ALL treatment with ChO, (2) using formal assessment of cognitive function or psychometric data on academic achievement, (3) including comparison with population norms (= uncontrolled) or healthy peers (= controlled), (4) offering original data, and (5) published after 1985, (6) in English, French, or German. Studies comparing patients treated with ChO to protocols with CI were included only if also including a comparison to normative data or healthy controls.

Excluded were studies with mixed diagnoses, books, comments, reviews, dissertations and case reports. In case of multiple reports on (one) the same cohort or patient group, the one published first was used for the overview of demographic characteristics.

Figure 1: Literature search.



Each study was read and coded by one author (NJ); results were controlled through independent reading and coding by one co-author (AK). The included studies varied widely with respect to design and psychometric measures. To enhance comparability, we grouped them along two dimensions: longitudinal versus cross-sectional and controlled (i.e. ALL patients to (healthy) controls) versus uncontrolled (i.e. comparison to normative data only). Studies were considered prospective if the first assessment was completed within 1 month after diagnosis.

Percentages of eligibility and attrition were given as these may affect the outcome.

Results of studies, i.e. neuropsychological tests were categorized into 7 cognitive domains as reported by the original authors. If no domain was given, results were attributed to a domain by the authors of this review, in accordance with international neuropsychological standards.[5] Hence, the following domains were established: intelligence or general mental development (in very young children); executive functioning; attention; language; perceptual-spatial functioning; learning and memory; and fine motor functioning (Table 1). School achievement was recorded separately.

Results were judged with respect to significant group differences and, if present, effect sizes or confidence intervals. These values characterize the magnitude of an effect or the strength of a relationship between two variables. Additionally, we analyzed whether the authors' interpretation of the outcome was based on (solely) statistical analysis of group comparisons (e.g. significant p-value) and/or other definitions of deficit or impairment.

Results

Study designs

From 123 studies obtained through the keywords, 26 could be included in this review according to the inclusion criteria (Fig. 1). These 26 original studies referred to 805 ALL-patients[6-31], ranging from 4 - 132 patients per study. Mean or median age at first evaluation as reported in 22/26 studies was 9 years (range 1.7-15.8). Based on 23 studies, 63% girls and 37% boys participated. Nineteen studies (73%) were cross-sectional[13-31] and seven out of 26 studies (27%) were longitudinal.[6-12] Only three out of these seven longitudinal studies were controlled[8-10] and four of seven were prospective.[6,8,10,12] Two studies referred to the

same cohort of children, yielding 5 articles on different functions.[8,15,32-34] Longitudinal studies yield the most powerful evidence and were analyzed first with special attention for the controlled and/ or prospective designs.

Table 1: Neuropsychological tests applied in reviewed studies, categorized per domain.

Domain	Tests / subtests (with aim to measure the domain)	*Used by ref*
Mental development or Intelligence	Denver developmental scales	**8**
	McCarthy scales of children's abilities	**7,11**
	Abbreviated Wechsler intelligence scales (3 subtests) (Wechsler intelligence scale for children -III, Wechsler adult intelligence scale -III)	25
	Culture fair test (not specified)	29
	Stanford-Binet intelligence test (4th ed)	31
	Stanford-Binet vocabulary (4th ed)	26
	Wechsler preschool and primary scale of intelligence-(revised)	**6,8-10,12**,22, 31
	Wechsler intelligence scale for children-(revised)/ III	**7-12**,14,16-19,22-24, 28-32
	Wechsler abbreviated scale of intelligence	**6**
	Wechsler adult intelligence scale-(revised)/ III	18-19,24,28-30
Executive Function	Amsterdam neuropsychological tasks	21,33
	Controlled oral word association test	7, 16, 24
	McCarthy scales of children's abilities, numerical, sorting	7
	Sorting task (not specified)	16
	Trail making test	**10**,20
	Wisconsin card sorting test	**7,8,34**
	Wechsler intelligence scale for children -revised, digit backward subtest	**7**
Attention	Amsterdam neuropsychological tasks	15,21,33
	Bourdon-Wiersma / Bourdon-Vos dot cancellation test	**8-10,34**
	Conner's continuous performance test	25
	Detroit test of learning aptitude (attentional domain)	13
	Double mark test	16
	d2 test	29
	Gordon diagnostic system	28
	Matching familiar figures test	13
	Trail making test	**10**,20,24
	VIGIL (not specified)	27
	Speed of information processing subtest of the British ability scales	27
	Wechsler preschool and primary scale of intelligence, sentence repetition	**9,10**
	Wisconsin card sorting test	27
	Wechsler intelligence scale for children -revised, symbol search, digit span, digit symbol, arithmetic, coding	**9,10**,16, 18,20, 26,27
Language	Peabody picture vocabulary test	**6,7**,31
	Token test of language comprehension	24
	Verbal fluency	**8**,18
	Wechsler, vocabulary subtest (Wechsler abbreviated scale of intelligence, Wechsler preschool and primary scale of intelligence-revised)	**6**
Perceptual-spatial	Benton line orientation	31
	Developmental test of visual-motor integration	**6-10**,13,14,18, 26,31,**34**
	Hooper visual organisation test	26
	Rey complex figure test (copy)	**8**,20,26,**34**
	Trail making test	31
	Wechsler, block design subtest (Wechsler abbreviated scale of intelligence, Wechsler preschool and primary scale of intelligence-revised)	**6**
Learning and memory	Benton visual retention test	24,31
	Children's California verbal learning test	**7**,23
	Children's memory scale	28
	Continuous recognition memory (hits)	**7**
	McCarthy scales of children's abilities, picture memory subtest	**7**
	Porteus labyrinths test	16
	Recurring figures test	29
	Rey auditory verbal learning test	**8-10**,16,31,**34**
	Rey complex figure test	**8**,16,20,26,**34**
	Stanford-Binet intelligence scale-(4th ed), bead memory, memory for objects, memory for sentences	**6**
	Taylor/Babcock-Levy story recall Test	13
	Verbal selective reminding test	24
	Wechsler memory scale-(III)	28
	Wide range assessment of memory and learning	17,18,22
Fine motor	Detroit test of learning aptitude (motor domain)	
	Developmental test of visual-motor integration	13,14
	Finger tapping	**10**
	Lafayette grooved pegboard	31
	Purdue pegboard	**8-10**,18,26,**34**

Bold = longitudinal studies
Not bold = Cross-sectional studies

Longitudinal studies

Characteristics of participating children

Seven longitudinal studies included 207 patients with 2 or 3 assessments, with follow-up time varying from 4½[34] to 7 years.[6]

Four out of seven studies reported the percentage of eligible patients who participated; mean number of missing patients in these studies was 18% (range 6-44).[7-10] Two of the longitudinal studies reported non-significant differences between included and excluded patients with regard to potential risk factors for poorer cognitive performance such as sex, age at diagnosis, and socioeconomic status.[8-9] Three studies offered the attrition percentage during the study with a mean loss for follow-up of approximately 21% (range 0-52).[8-9,12] Reasons for attrition were predominantly relapse of ALL or death. Six out of 7 longitudinal studies reported on socioeconomic status; the majority of subjects was from middle-class backgrounds.[6-7,9-12]

Four studies applied a prospective study design according to our definition[6,8,10,12]; in one study the first evaluation was done 8 months after diagnosis.[7] A healthy control group with the same inclusion criteria as the patients with respect to age, language and normal development was added in 3 studies.[8-10]

In three studies, children were treated according to different ChO protocols. Two studies differentiated between protocols, whereas one study pooled the data of all children irrespective of ChO protocol. Intrathecal methotrexate [IT MTX] combined with other chemotherapeutic agents was given in varying doses.[6] Patients in all studies received HD IV MTX in doses varying between 1 g/m^2 and 5 g/m^2; in most studies, the dose on was not reported.

Cognitive domains and psychometric measures (Table I)

Intelligence or mental development was studied in all 7 longitudinal studies, using 5 different scales. As Table 1 shows, a wide diversity of tests was used to assess specific cognitive functions. Subtests of intelligence scales were often used to evaluate both intelligence and other cognitive functions. One study used a comprehensive cognitive assessment battery covering all 7 domains.[8]

Academic achievement was evaluated in three studies, mostly using the wide range achievement test – revised (WRAT-R).[7,11-12]

Results of longitudinal studies (Table 2)

In 3 of the 7 longitudinal, controlled studies yielding 4 reports, no evident negative neuropsychological late effects of ChO were found on cognition, apart from deterioration of complex fine-motor functioning.[9-10,32,34] However, slight impairment in verbal IQ and attention, not resulting in poor school performance, was found in 2 studies.[9,10] In the study of Jansen et al, it was noted that patients who uttered physical complaints (i.e. pain and/or tiredness) at the first pre-treatment assessment scored significantly lower than siblings on attention and speed at follow-up, although this could not been attributed to treatment effects.[34] The authors hypothesized that these patients have reduced cognitive reserve capacity as they scored already lower at the start of therapy. In the other 4 uncontrolled longitudinal studies, modest decline was reported for a variety of functions at the last evaluation although performances continued to be in the average range.[6,7,11,12] Two studies found adverse sequelae of ChO within the domain of perceptual-spatial[6,7] and one study described decrease in verbal IQ scores.[12] Another study reported subtle deficits within the domain of language[7] and learning and memory.[6] Modest decline in academic arithmetic was found in 3 studies.[7,11,12]

Four studies based the outcome on statistically significant differences between patients and controls and additionally used some measure for defining "deficit(s)".[9,10,12,32] One prospective study defined "clinically important decreases" as a negative change in individual test score of ≥ 15 IQ points[12], whereas another study used a cut-off of more than 10 IQ points below the mean norm scores.[32] In both studies of Kingma et al, impaired performance was also calculated per test measure and per patient at the last evaluation, defined as a standard score of 1.64 SD below the normative mean.[9,10]

Five studies reported on effect sizes or confidence intervals, interpretation of graphics, or a combination of these.[7-10,12] One study created composite scores to minimize statistical comparisons based on Pearson product correlations.[6]

Table 2: Treatment effects as reported in the 26 studies

Domain		*Significant differences*	*No significant differences*	*Not applicable*	*Reported treatment effects*					
						VIQ		*PIQ*		*TIQ*	
						Pt	*Co*	*Pt*	*Co*	*Pt*	*Co*
Intelligence	Controlled	**10**,16,17,22,24,25	**9**,19,23,26,28,29,31,**32**,**34**	15,20,33	**10**	100	109*	114	112	107	112
					16	100	112*	98	114*	99	113*
					17	87	109	80	99	82	104*
					22	87	100*	85	100*	nr	nr
					24	98	114*	97	112*	98	114*
					25	nr	nr	nr	nr	94	100*
	Uncontrolled	**12**,14	**6-7**,**11**,18,27,30	13,21	**12**	97	104*	nr	nr	100	104*
					14	98	100	92	100*	95	100
Executive function	Controlled	33	**10**,16,20,24,**34**	**9**,17,19,22,23,25,26,28,29,31,**32**	33	One subtest of the ANT					
	Uncontrolled		**7**,21	**6**,**11**,**12**,13,14,18,27,30,33							
Attention	Controlled	**10**,15,16,20,25,28,**34**	**9**,24,26,29	17,19,22,23,31,**32**	**10**	Trail making test					
					15	Four subtests of the ANT					
					16	One subtest of the WISC					
					20	Trail making test, subtest WISC					
					25	Connor (one subtest?)					
					28	One subtest Gordon diagnostic system					
					34	In subgroup 1 subtest Bourdon-Vos					
	Uncontrolled	13,21	18,27	**6**,**7**,**11**,**12**,14,30,33	13	DTLA, attentional domain					
					21	Three subtests of the ANT					
Language	Controlled		24,31	**9**,**10**,15-17,19,20,22,23,25,26,28,29,**32**,**34**							
	Uncontrolled	**7**	**6**,18	**11-14**,21,27,30,33	**7**	Verbal fluency					
Fine motor	Controlled	**9**,**34**	**10**,26,31	15-17,19,20,22-25,28,29,32	**9**	One subtest Purdue pegboard					
					34	One subtest Purdue pegboard					
	Uncontrolled	13,18	14	**6**,**7**,**11**,**12**,21,27,30,33	13	DTLA: motor domain					
					18	Two subtests Purdue pegboard					
Perceptual-spatial function	Controlled		**9-10**,20,26,31,**34**	15-17,19,22-25,28,29,**32**							
	Uncontrolled	**6**,**7**,13,14,18		**11**,**12**,21,27,30,33	**6**	Visuomotor integration, subtest of WISC					
					7	Visuomotor integration,					
					13	Visuomotor integration					
					14	Visuomotor integration					
					18	Visuomotor integration					
Learning and memory	Controlled	**9**,17,20	**10**,16,22-24,26,28,29,31,**34**	15, 19,5,**32**	**9**	RAVLT, delayed recall					
					17	WRAML					
					20	One subtest complex figure Rey					
	Uncontrolled	**6**,13	**7**,18	**11**,**12**,14,21,27,30,33	**6**	Two subtests of SBI					
					13	Delayed recall (Taylor/Babcock story recall)					

Bold = longitudinal studies, ANT = Amsterdam neuropsychological tasks, SBI = Stanford Binet intelligence test, WISC = Wechsler Intelligence Scales

Cross-sectional studies

Characteristics of children participating in the studies

Nineteen studies with a cross-sectional design included 602 patients with a reported age at evaluation in 15 studies (mean or median 11 years, range 1-33.7 years).[13-31] The average proportion of participating eligible patients, reported in eleven studies, was 81%.

Eight studies reported on socioeconomic status of the participating subjects.[14-17,20,24,26,30]

Three studies included both a healthy control group and a clinical control group, which consisted either of cancer patients who had received local therapy only[15,16] or patients with chronic asthma.[24] One study included control patients surviving from diverse non CNS solid tumors.[30] Eight other studies had collected control data from healthy peers, mostly matched for age and gender.[17,18,20,21,23,26,27,29] Three studies did not include a control group.[13,14,27]

The reported mean time since treatment varied between 1 to 7 years.

In five studies, patients were treated on different protocols and three studies controlled for this confounding variable.[15,21,23] IT MTX was given in all studies with dose varying between 6 g/m^2 and 15 g/m^2. High dose intravenous methotrexate [IV MTX] was given in all but 5 studies[16,17,20,24,26] with doses varying between 1 g/m^2 and 8 g/m^2 except one study applying a very high dose of 33.6 g/m^2.[28]

Cognitive domains and psychometric measures used (Table 1)

Similar to the longitudinal studies, intelligence and cognitive functioning were evaluated using a wide variety of neuropsychological instruments. Intelligence was most often assessed in 13 studies, using 7 different scales. To examine cognitive abilities, 35 different (sub)tests were applied in 15 studies. Academic achievement was evaluated in 8 studies, most of them using the WRAT(R) and the Woodcock-Johnson-revised (WJ-R).

Results of cross sectional studies (Table 2)

In thirteen of the 19 cross-sectional studies, patient groups showed statistically significant lower scores in comparisons to controls and/ or norm groups[13-18,20-22,24-25,28,33], most frequently within the domain of attention[13,15,16,20,21,25,28] and intelligence.[14,16,17,22,24,25] Other studies reported deficits within the memory domain[17,20,13], perceptual-spatial functioning[13,14,18] or within the domain of fine motor functioning.[13,18]

Although in three other studies the group differences were not statistically significant, these studies reported on elevated proportions of abnormal scores, i.e. an IQ below 1 SD below the normative mean.[19,26,30]

Six out of 8 studies found a modest decline in academic achievement, predominantly with regard to arithmetic.[13,18,23-25,31]

Eight studies based the outcome on some kind of definition of a deficit.[16,19,20,22,25,26,30,31] For intelligence, the clinical importance of a change in mean test results was mostly set at ≥ 15 IQ point's decrease. For determining the meaningfulness of a result, seven studies reported on effect sizes or confidence intervals, or a combination of these.[13-15,19-21,27]

Risk factors

Younger age, female sex, higher dose of chemotherapeutic agents such as MTX, lower SES and longer time since treatment interval have been reported as "risk factors", i.e. as factors related to cognitive impairment.

In 6 studies, a younger age at diagnosis (patients younger than 5 or 6 years) was related to significantly worse cognitive outcome.[15,16,20,30,31,33] The majority of the reviewed studies, however, either did not find significant age effects[7,9,10,14,24,25] or did not report on this risk factor.

On gender effects, inconsistent results were reported; 1 uncontrolled longitudinal and 4 cross-sectional studies found a disadvantage for girls[7,14,15,30,33], whereas 2 controlled longitudinal and 4 cross-sectional studies did not found differences between males and females.[9,10,16,21,22,24]

Five studies commented on MTX dose; in 3 studies cognitive outcome was related to higher IV MTX dose in dosages varying from of 1 to 3 g/m^2[2,6,15,22] but 2 studies reporting a MTX dose of 10 mg/m^2 IV, found however no relationship with IV MTX dose.[16,25] Occasionally, a shorter infusion rate and no leucovorin rescue have been reported for poorer outcome, but have seldom been included in the analyses of the studies of this review.

Recently, it has been hypothesized that differences in outcome may be related to variations in the duration of infusions with a worse outcome for longer duration.[6] Others hypothesized an increased neurotoxicity of MTX if given in combination with Ara-C.[16] Four studies examined the relation between lower SES and poorer cognitive outcome and only one could confirm this relationship.[7,14,16,24] However, 58 % of the included patients had an ethnic minority status, which may have affected their test results. Only 5 cross-sectional studies evaluated time since treatment as risk factor; 2 studies found poorer performance after longer follow-up[13,24] whereas 2 other studies showed no effect.[14,16] Contrary, a short time since end of treatment was found a significant risk factor for poorer performance in 1 controlled cross-sectional study using 3 subtests of the Amsterdam Neuropsychological Tasks (ANT).[33] However, this study included patients that were treated according to 4 different ChO protocols.

Discussion

Reviewed studies showed substantial methodological variability and inconsistent results. Based on the few longitudinal and/or controlled designs, results suggest no or mild cognitive impairment after ChO treatment. Most conclusive findings specifically indicated discrete decrease of complex fine motor functioning. Additionally, impairment was found for a variety of neuropsychological functioning and below average arithmetic skills, but for most patients performance was within normal limits. A minority of studies included analysis of potential risk factors for poorer performance, i.e. younger age at diagnosis, female sex and higher MTX dose, but no consensus emerged with an equal number of studies finding a positive and a negative effect. These unequivocal conclusions are due to often unavoidable, methodological difficulties in this field of research. The following issues have been identified:

1. The cross-sectional design of the majority of the studies impeded a strict comparison between the patient's cognitive functioning prior to and after ChO treatment. Cross-sectional studies may attribute impairment to treatment whereas poorer functioning in fact already was present prior to the disease. Neuropsychological assessment at diagnosis has been considered difficult to perform in children with cancer, but this has been proven to be feasible and reliable and is recommended to discriminate between possible adverse sequelae of disease and treatment and the patients' innate abilities.[8]
2. Inclusion criteria differed among studies or were unspecified. Therefore, outcome differences could have been attributable to pre-existing group differences rather than treatment effects.[34] Another problem could have been the issue of selective attrition. The majority of the longitudinal studies did not report on the reason for decreased number of patients during the study. Possibly, patients with poorer or better functioning were more likely to drop out, resulting in an overrepresentation of patients with better or worse test results, respectively.

   Finally, the percentage of girls in the reviewed studies was higher than boys, whereas the incidence and sex-ratio is about equally in ALL patients.[35] This could potentially result into an overrepresentation of deficits as female sex has been suggested a risk factor for cognitive impairment following ALL treatment[7,14,15,30,33], but further research is needed on this topic.

3. Use and choice of a control group is essential in evaluating cognitive effects of treatment. The majority of the studies in this review compared patients to normative data rather than to matched healthy peers. Unfortunately, quality of norms varies widely across neuropsychological instruments and countries. For example, if the normative sample is small or not similar to the patient group with regard to education, the comparison group might not have been representative for the patient group.

   A control group is also prerequisite to control for practice effects in a repeated measurement design. Some tests show a significant increase in scores by practice alone but most tests have no estimates of gains associated with repeated testing. Thus, stable scores could actually reflect decrease of performance if this test normally is associated with practice effects.

   Finally, norms can be outdated due to the Flynn effect. For intelligence tests, it is recommended to correct for an IQ increase of approximately 3 points every 10 years since collection of test norms.[36] However, this correction is not necessarily correct for any test at any evaluation. Moreover, this correction is not known for most neuropsychological measures.

4. The bewildering number and variety in tests to assess cognitive functions (see Table 1) precludes appropriate comparison of results in this review. For example, to assess attention 13 different instruments were used and five tests that were attributed to a particular cognitive domain by one group of authors were attributed to a different domain by other authors. The lack of consensus with respect to a standard comprehensive test battery in ALL research cries for a solution but this problem is not uniquely associated with ALL research.

   Another significant problem is that subtests of intelligence scales were often used both to assess IQ and a more specific cognitive function, whereas these subtests lack validity for measuring cognitive functions independently. Moreover, if patients perform poorly on an intelligence subtest, this will wrongly result in a poor score for the specific cognitive function too.

   The comprehensiveness of the assessments varied considerably. Intelligence was most studied, as IQ tests are easily available and widely used by general psychologists and educational services. Although IQ is a robust indicator of general cognitive capacities, intelligence is unfortunately highly correlated with education and many leukemia

patients suffer from substantial school absenteeism. Language was least studied, whereas a lag in language development would certainly have implications for school career.

5. A few longitudinal studies encountered the almost insurmountable problem of test shift, *i.e.* the need to change tests with growing age because most tests have a limited age range. Norms for consecutive tests for older age are not necessarily similar and can only be longitudinally compared if such test characteristics are available, or test version should be included as confounding variable in longitudinal analyses. Moreover, some functions can only be assessed at an older age which makes longitudinal follow-up of young patients challenging. The NEPSY-II, a test battery to assess neuropsychological functioning in six functional domains for children ages from 3 to 8 years[37], meets with the aforementioned problems but is not yet available in every country.
6. In the majority of the reviewed studies, patients were compared with normative data obtained from the healthy population. However, group analyses should be completed with establishment of the proportion of children with deficits because a group mean may have declined due to a few patients with major deterioration, whereas most of the patients remained stable. Additionally, the definition and severity of deficits should be offered.
7. The statistical procedures could be ameliorated in future studies. A correction for multiple analyses was seldom reported in studies in this review. Moreover, data of patients with missing values were often excluded from the analyses resulting in attrition and outcome bias. By applying relatively new sophisticated techniques like multilevel analyses, data of patients with missing values can also be included.[38]

Conclusions

Despite the expanding literature on neuropsychological effects of chemotherapy in children treated for ALL, the controversy about specific treatment-related impairment remains. There are many methodological issues that preclude definite conclusions on the possible effects of ChO on cognitive functioning but ChO treatment was generally associated with good neuropsychological outcome. More prospective studies with healthy, matched control groups and an extensive neuropsychological test battery are still needed. Neuropsychological assessment should include tests of intelligence, verbal-auditory and visual memory, visual-motor integration, perceptual-spatial functioning, attention, executive functioning, language and fine motor functioning so that appropriate educational intervention can be developed, if necessary.

In addition, to improve understanding of the mechanisms causing cognitive impairment in specific patients remains an important goal. To identify patients who are most at risk for cognitive sequela and hopefully prevent or cure these adverse effects of chemotherapy is a challenge for future research.

References

1. Pui CH, Evans WE: Treatment of acute lymphoblastic leukemia. N Engl J Med 2006;354:166-178
2. Chessells JM, Harrison G, Richards SM, et al. Failure of a new protocol to improve treatment results in paediatric lymphoblastic leukaemia: lessons from the UK Medical Research Council trials UKALL X and UKALL XI. Br J Haematol 2000;118:445-55
3. Roman DD, Sperduto PW. Neuropsychological effects of cranial radiation: current knowledge and future directions. Int J Radiat Oncol Biol Phys 1995;13(4):983-998
4. Butler RW, Hill JM, Steinherz PG, Meyers PA, Finlay JL. Neuropsychologic effects of cranial irradiation, intrathecal methotrexate, and systemic methotrexate in childhood cancer J Clin Oncol 1994;12:2621-2629
5. Baron IS. Neuropsychological evaluation of the child. Oxford University Press, 2004
6. Carey ME, Hockenberry MJ, Moore IM, et al. Brief report: Effect of intravenous methotrexate dose and infusion rate on neuropsychological function one year after diagnosis of acute lymphoblastic leukemia. J Pediatr Psychol 2006;32(2):189-93
7. Espy KA, Moore IM, Kaufmann PM, et al. Chemotherapeutic CNS prophylaxis and neuropsychologic change in children with acute lymphoblastic leukemia: a prospective study. J Pediatr Psychol 2001; 26(1):1-9
8. Jansen NC, Kingma A, Tellegen P, et al. Feasibility of neuropsychological assessment in leukaemia patients shortly after diagnosis: directions for future prospective research. Arch Dis Child 2005;90(3):301-4
9. Kingma A, van Dommelen RI, Mooyaart EL, et al. Slight cognitive impairment and magnetic resonance imaging abnormalities but normal school levels in children treated for acute lymphoblastic leukemia with chemotherapy only. J Pediatr 2001;139(3):413-420
10. Kingma A, Dommelen RIv, Mooyaart EL, et al. No Major Cognitive Impairment in Young Children With Acute Lymphoblastic Leukemia Using Chemotherapy Only: A Prospective Longitudinal Study. J Pediatr Hematol Oncol 2002;24(2):106-113
11. Moore IM, Espy KA, Kaufmann P, et al. Cognitive consequences and central nervous system injury following treatment for childhood leukaemia. Semin Oncol Nurs 2000; 16(4):279-290
12. Ochs J, Mulhern R, Fairclough D, et al. Comparison of neuropsychologic functioning and clinical indicators of neurotoxicity in long-term survivors of childhood leukaemia given cranial radiation of parental methotrexate: a prospective study. J Clin Oncol 1991;9(1):145-151
13. Brown RT, Madan-Swain A, Pais R, Lambert RG, Sexson S, Ragab A. Chemotherapy for acute lymphocytic leukemia: Cognitive and academic sequelae. J Pediatr 1992;121(6):885-889
14. Brown BT, Madan-Swain A, Walco GA, et al. Cognitive and academic late effects among children previously treated for acute lymphocytic leukemia receiving chemotherapy as CNS prophylaxis. J Pediatr Psychol 1998;23(5):333-340
15. Buizer AI, De Sonneville LMJ, Van den Heuvel-Eibrink MM et al. Chemotherapy and attentional dysfunction in survivors of childhood acute lymphoblastic leukemia: effect of treatment intensity. Ped Blood Cancer 2005;45(3):281-90

16. Giralt J, Ortega JJ, Olive T, Verges R, Forio I, Salvador L. Long-term neuropsychologic sequelae of childhood leukemia: comparison of two CNS prophylactic regimens. Int J Rad Oncol Biol Phys 1992;24:49-53
17. Hill DE, Ciesielski KT, Sethre-Hofstad L, Duncan MH, Lorenzi M. Visual and verbal short-term memory deficits in childhood leukemia survivors after intrathecal chemotherapy. J Pediatr Psychol 1997;22(6):861-870
18. Kaemingk KL, Carey ME, Moore IM, et al. Math weaknesses in survivors of acute lymphoblastic leukemia compared to healthy children. Child Neuropsychol 2004;10(1):14-23
19. Khong PL, Leung LH, Fung AS, et al. White matter anisotropy in post-treatment childhood cancer survivors: preliminary evidence of association with neurocognitive function. J Clin Oncol 2006;20;24(6):884-90
20. Lesnik PG, Ciecielski KT, Hart BL, et al. Evidence for cerebellar-frontal subsystem changes in children treated with intrathecal chemotherapy for leukemia. Arch Neurol 1998;55:1561-1568
21. Mennes M, Stiers P, Vandenbussche E, et al. Attention and information processing in survivors of childhood acute lymphoblastic leukemia treated with chemotherapy only. Pediatr Blood Cancer 2005;44(5):478-86
22. Montour-Proulx I, Kuehn SM, Keene DL et al. Cognitive changes in children treated for acute lymphoblastic leukemia with chemotherapy only according to the Pediatric Oncology Group 9605 protocol. J Child Neurol 2005; 20:129-133
23. Precourt S, Robaey Ph, Lamothe I, et al. Verbal Cognitive functioning and learning in girls treated for acute lymphoblastic leukemia by chemotherapy with or without cranial irradiation. Dev Neuropsychology 2002; 21(2):173-195
24. Raymond-Speden E, Tripp G, Lawrence B, Holdaway D. Intellectual, neuropsychological, and academic functioning in long-term survivors of leukemia. J Pediatr Psychol 2000; 25(2):59-68
25. Reddick WE, Shan ZY, Glass JO et al. Smaller white-matter volumes are associated with larger deficits in attention and learning among long-term survivors of acute lymphoblastic leukemia. Cancer 2006;106(4):941-949
26. Regan JM, Reed RN. Neuropsychological functioning in survivors of childhood leukemia. Child Stud J 1998;28(3):179-200
27. Rodgers J, Marckus R, Kearns P, Windebank K. Attentional ability among survivors of leukaemia treated without cranial irradiation. Arch Dis Child 2003;88(2):147-50
28. Spiegler BJ, Kennedy K, Maze R, et al. Comparison of long-term neurocognitive outcomes in young children with acute lymphoblastic leukemia treated with cranial radiation or high-dose or very high-dose intravenous methotrexate. J Clin Oncol 2006; 24(24):3858-3864
29. Uberall MA, Haupt K, Meier W, et al. P300 abnormalities in long-time survivors of acute lymphoblastic leukemia in childhood-Side effects of CNS prophylaxis? Neuropediatrics1996;27:130-135
30. Von der Weid N, Mosimann I, Hirt A, Wacker P, Nenadov Beck M, Imbach P et al. Intellectual outcome in children and adolescents with acute lymphoblastic leukaemia treated with chemotherapy alone: age- and sex-related differences. Eur J Cancer 2003;39:359-365
31. Wilson DA, Nitschke R, Bowman ME, et al. Transient white matter changes on MR images in children undergoing chemotherapy for acute lymphoblastic leukemia: Correlation with neuropsychologic deficiencies. Radiology 1991;180:205-209

32. Jansen NC, Kingma A, Schuitema A, et al. Post-treatment intellectual functioning in children treated for acute lymphoblastic leukaemia (ALL) with ChO: a prospective, sibling-controlled study. Eur J Cancer 2006;42(16):2765-72
33. Buizer AI, De Sonneville LMJ, Van den Heuvel-Eibrink MM et al. Visuomotor control in survivors of childhood acute lymphoblastic leukaemia treated with chemotherapy only. J Int Neuropsychol 2005;11(5):554-56
34. Jansen NCAJ, Kingma A, Schuitema A, et al. Neuropsychological outcome in chemotherapy-only treated children with acute lymphoblastic leukemia. J Clin Oncol 2008;26(18):3025-30
35. Pizzo PA, Poplack DG: Principles and Practice of Pediatric Oncology, fourth ed. Philadelphia, PA: Lippincott Williams & Wilkins, 2001
36. Flynn JR. Massive IQ gains in 14 nations: what IQ tests really measure. Psychol Bull 1984;101:171-191
37. Korkman M, Kirk U, Kemp S: NEPSY-Second Edition (NEPSY-II). A developmental neuropsychological assessment, Harcourt Assessment, 2007
38. Snijders TAB, Bosker, RJ: Multilevel Analysis, An introduction to basic and advanced multilevel modelling, London, Sage publications, 1999

CHAPTER 7

General discussion

Acute lymphoblastic leukaemia (ALL) is the most common malignancy in childhood. It is, nevertheless, a rare disease in terms of absolute numbers, with ± 120 new cases diagnosed in The Netherlands each year. The incidence of ALL peaks in the pre-school age (3-4 years of age), a highly dynamic developmental period when the central nervous system (CNS) is presumed to be extremely vulnerable to damage. Before the 1970s, ALL was virtually always fatal since early treatments with anti-neoplastic drugs could not prevent relapse in the CNS, a 'sanctuary for leukaemia cells'. A breakthrough in the treatment of children with ALL came with the introduction of prophylactic CNS therapy, which considerably prolonged life expectancy. Traditionally, CNS treatment included cranial radiation therapy (CRT) and intra-thecal chemotherapy, the latter with methotrexate with or without additional agents.[1] When long-term follow-up data on large numbers of children treated with CNS therapy including CRT became available, late adverse effects, including severe neuropsychological sequelae, were incontestable.[2] Alternative CNS-directed strategies were designed, in which CRT was abandoned and risk groups were defined to which treatment protocols could be tailored. Treatment protocols in the Netherlands were among the first to abandon CRT and use chemotherapy only. These chemotherapy only protocols (ChO) usually consist of three phases, i.e. remission induction and prophylactic CNS therapy, delayed intensification, and maintenance treatment with systemic and intrathecal (IT) chemotherapy. The question that we are presently faced with is whether chemotherapy-only has negative effects on cognitive development in children who are treated for ALL. Findings of previous studies on this topic are inconclusive and the main aim of this thesis is to address this question.

Earlier research

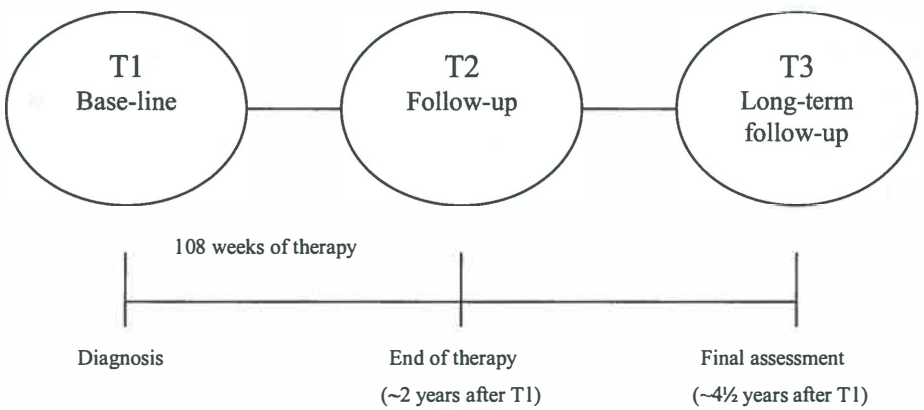
In 2001 Kingma et al addressed the neuropsychological late effects of leukaemia treatment in children younger than seven years at diagnosis. In short, Kingma could not confirm that children performed more poorly than controls after treatment with ChO and concluded that further research on attention and memory skills was required. In 1999, a collaborative nation-wide investigation was initiated by the Paediatric Oncology Centre, University

Medical Centre Groningen, using a prospective, longitudinal, controlled study design of which the present thesis offers the results.

Strengths of the present study

The obvious methodology for outcome studies is a prospective rather than a cross-sectional study design and therefore this design was chosen in our study. We assessed cognitive functioning within 2 weeks after start of ChO-treatment as a baseline for the evaluation of cognitive outcome after treatment. The second assessment (T2) was scheduled about 2 years later, i.e. two to six months after cessation of therapy. This timeframe enabled patients to recover from the intensive treatment and warranted the clearance of most cytostatic drugs by the body. To evaluate potentially negative, adverse, late cognitive effects, a third assessment (T3) was performed two years after T2, about 4½ years after diagnosis (see Fig. 1).

Figure 1: Time-frame of the study



A multi-centre study was essential to obtain a sample of patients that was sufficiently large for adequate data analysis, which also would merit generalization of findings. Between January 1999 to July 2001, 50 consecutive patients, a relatively large sample compared to previous studies, were included according to strict inclusion criteria (see review, chapter 6).

All ALL patients were treated according to protocol 9 of the Dutch Childhood Leukaemia Study Group (DCLSG), later the Dutch Childhood Oncology Group (DCOG). During follow-up loss of patients was minor.

A single qualified neuropsychologist did all the assessments to prevent inter-examiner effects, a potential threat to the internal validity of the study.

The selection of a control group is a matter of debate. We considered a consecutive sibling control group as the best option, because it enabled us to control for differences in social economic status and for normal learning effects of repeated assessment. A potential disadvantage of sibling controls may be that the emotional impact of having a sibling with a life threatening disease affects neuropsychological test scores.[3] However, at baseline, and thus in a period of strong emotional impact of the illness, mean scores on intelligence scales (mean IQs in both groups were high average) did not differ significantly between patients and siblings and/ or normative population. Another disadvantage of this type of control group is the difference in age between patients and controls. Comparison of characteristics of the groups of patients and siblings in this study showed slight, statistically insignificant differences with respect to gender and age at testing; median age of siblings (Md = 8.2) was 1.8 years higher than of patients (Md = 6.4). Yet, we accounted for a potential bias of age differences between the patient and control group by applying age as a controlling variable in the multilevel analyses.

Further, a comprehensive rather than a narrow assessment of cognitive functions was chosen to prevent methodological deficiencies of previous studies. Specific measures of cognitive functioning were assessed besides a general measure, such as intelligence quotient (IQ). Specific measures were derived from tests of verbal-auditory and visual memory, visual-motor integration, attention, cognitive flexibility and fine motor functioning. Former research suggested impairments in one or more of these domains after treatment of ALL. Furthermore, such functions are potentially related to problems in school and daily live.

To analyse such a complex data set, a fully multivariate model was used which is roughly comparable to a multivariate linear repeated measures model, but the former model enabled us to deal with missing values. In contrast to traditional models that delete missing cases from the case list, which is a major loss, we were not obliged to remove incomplete cases (*e.g.* patients with missing data because of death or relapsed disease).

Weaknesses of the study

Although we did every effort to include all new patients, the number of patients missing at baseline was high; 35 out of 79 eligible patients. It is, however, unlikely that these missing patients have affected the results, as the missing patients did not differ significantly from the included patients regarding demographic features and initial disease characteristics.

The limited age range of 4 to 12 years inclusion was chosen because the duration of follow-up was 4½ to 5 years. Most psychometric tests for children and youngsters can be used up to the age of 17; hence, the upper age threshold at the time of inclusion could not exceed 12 years. We wanted to include patients younger than four years at diagnosis to assess the potential risk factor 'young age', but as the minimum age for Wechsler intelligence scales is four years, they could not be included. Moreover, most psychometric tests for children can only be used by the age of 6.

Research questions and discussion

1. Is neuropsychological assessment to establish base-line functioning feasible in seriously ill, newly and recently diagnosed patients with ALL?

Renouncing a baseline assessment is frequently justified by the argument that ALL patients are too ill for reliable assessment of cognition. However, to enable interpretation of any change in neurocognitive development after treatment, a baseline profile of cognitive functioning is essential. We were challenged to determine whether or not a neuropsychological assessment was feasible in children in the acute phase of ALL and early hospitalisation.

We assessed cognitive functioning in ALL patients as early after diagnosis as possible. The results demonstrated that most children who were very recently diagnosed with the life-threatening disease ALL can reliably cooperate in an extensive standardized neuropsychological assessment. We could also establish that ALL and psychological factors have no adverse effect on IQ or neuropsychological function in the acute phase prior

to ChO. The majority of the patients enjoyed the assessment and found it a rather welcome distraction among numerous medical procedures.

2. Is change in cognitive functioning during and after 2 years of treatment the same in patients and sibling controls?

Intelligence was preserved during and after 2 years treatment (as shown by group-wise VIQ, PIQ, and FSIQ), both comparing ALL children to their healthy siblings and to the normative population. Although patients scored significantly lower compared to siblings on performance IQ at the end of maintenance therapy, their score had normalized 2 years later. Overall, significant practice effects (i.e. an improvement of performances as a result of repeated assessments) were found for both patients and siblings on performance IQ and full scale IQ.

For specific domains covered by the present study, i.e. learning, memory, attention, speed, executive functioning, and visual-constructive functioning, no major differences were found between patients and siblings, except for complex fine-motor functioning after 2 years treatment with patients scoring significantly lower than siblings. Both patients and siblings showed practice effect on the Rey's Auditory-Verbal Learning Test and the Wisconsin Card Sorting Test.

3. Is long-term cognitive functioning in children treated for ALL with ChO equal to healthy controls and normative data, 4½ to 5 years after diagnosis?

Two years after cessation of therapy, no negative effects of ChO were demonstrated on intelligence and the specific domains of learning, memory, attention, speed, executive functioning, visual-constructive and fine-motor functioning. However, patients still scored significantly worse compared to controls on the complex fine-motor task of the Purdue Pegboard in which pegs need to be positioned in holes with both hands simultaneously. The lower scores of patients on this task may be explained by a minor deficit in processing speed[4] or as a result of vincristine neuropathy.[5] The latter explanation seems, however, less likely as patients had scores similar to controls on a simple fine-motor task (placing pegs in their holes with the dominant hand) of the Purdue Pegboard. Results of our series of studies partly confirmed earlier findings.

Previous, longitudinal studies found no or minor influence of ChO treatment on intelligence. Studies on the remaining cognitive functions show that subtle deficits may exist within a variety of domains.

The pathogenesis of late CNS damage induced by ChO is not yet fully understood. Different mechanisms have been postulated to explain the underlying neurological basis of cognitive dysfunction potentially induced by ChO. Damage to cortical matter and subcortical white matter has received most attention, and MTX has been mentioned most frequently in this context.[6] MTX may cause direct damage to endothelial cells resulting in demyelisation and possibly loss of endothelial cells.

The fact that oppositely to our study some previous studies found negative effects of ChO may be the result of a retrospective study design, lack of a proper control group, or heterogeneous patient groups.

4. Is possibly poorer performance or cognitive deterioration related to specific patients' characteristics?

At baseline we recorded complaints about pain, fatigue or other physical discomfort in patients. We observed that the 'sick' patients (26.5%) still perform more poorly than their controls in tests of sustained attention 4½ years later, although performances were still in the normal range. They also performed poorer on full-scale IQ and performance IQ than their siblings. We hypothesize that these patients (n = 13, mean age at baseline 7.7 years) have reduced cognitive reserve capacity. This hypothesis means that an individual's cognitive reserve determines his or her threshold to give up on neuropsychological tasks. Those patients with relatively fewer cognitive resources are those who may be more reluctant to complete neuropsychological tasks and may utter more easily physical complaints because of frustration. The group of 'sick' patients consisted of 7 girls and 6 boys. The percentage of high risk patients within the 'sick' group (38.5%) did not differ from the whole patient group (34.7%).

The human male is more vulnerable than the female ('the fragile male') regarding neurocognitive development: developmental disorders such as specific reading delay, hyperactivity, autism and related disorders, clumsiness, stammering, and Tourette's syndrome occur 3 to 4 times more often in boys than in girls.[7] Interestingly, in children treated for ALL with CRT, the female sex was often found to be a risk factor for subnormal

cognitive development. It has been suggested that hormonal factors may interact with radiotherapy in the development of brain structures. Female hormones may be associated with a higher susceptibility of the CNS for adverse cognitive sequalae. However, this finding has not been observed for ChO. Analyses in our study did neither reveal differences between boys and girls nor differences between high risk and non-high risk patients.

Practical implications

Our results are reassuring for both patients with ALL and their parents, professional caregivers and oncologists: General and specific cognitive functioning continue to develop normally over time in patients treated with ChO. However, the patient group as a whole scored significantly worse than the healthy control group on complex fine motor skills at the last evaluation.

A surprising observation was that patients, who uttered pain, fatigue or other physical discomfort at baseline, still perform poorer than their controls on PIQ, FS-IQ, and on a sustained attention task. Four and half to 5 years later, mental vulnerability of these children may be associated with a decreased cognitive reserve, and future studies should attempt to trace these children early in the treatment. By applying a systematic inventory such as the Child Health Questionnaire (CHQ), interventions may be started, if necessary, on time. Deficits in fundamental areas such as attention and speed of information processing have implications for school career, occupational functioning, social relationship, emotional regulation, coping skills, and general quality of life.

Individual versus group analysis

The developmental process in children is, by definition, dynamic. Furthermore, there is wide normal inter-personal variability of cognitive functioning. Are we able to evaluate these individual differences in this thesis, in which 'abnormality' or 'deficit' was defined by statistical definitions of deviations from group scores? Additionally, when should a deficit

be considered clinically relevant and when should it be considered as directly caused by the illness or its treatment, in other words, to organic factors? In general, the group approach stresses statistical significance between group differences, whereas a more individual approach may reveal children who are affected by a, perhaps small, negative change in neuropsychological functioning. Also, definitions of an individual deficit and/or impairment are not clear. Should one define a 'deficit' as a deviation of more than 1 standard deviation below a norm score or the mean score of the control group, or should one apply stricter definitions? We need criteria to define deficits and impairment in individual cases.

Lack of consistency in the definition of cognitive impairment makes comparisons between studies difficult. The publication of the International Classification of Human Functioning for Children and Youngsters (ICF-CH) will be of great help in terminological cleaning up.[10]

Remaining questions

Despite intensive and potentially neurotoxic treatment, no evident, negative, late effects were found concerning general or specific neuropsychological measures immediately after cessation of therapy and 2 years thereafter (4½ years after diagnosis of ALL). However, very late damage presenting beyond the follow up of 4½ years after diagnosis cannot be excluded by the present study. Furthermore, we could not trace risk factors for poorer outcome such as gender or disease characteristics; our nation-wide patient group was too small for adequate analysis.

Future prospective, longitudinal studies should include a larger proportion of very young (< 4 years at diagnosis) patients to measure influences of factors such as age of onset of leukaemia. As claimed in the literature, very young children are particularly prone to deleterious effects of prophylactic CNS treatment.[8,9] Other risk factors, such as gender and dose of chemotherapeutic agents need further elucidation as the sample size of the present thesis was too small to analyse these factors in their mutual dependency. The current study suggested that patients who complained of fatigue at diagnosis remain at risk of slightly weaker attentiveness and performance IQ. We hypothesized that these patients have

reduced cognitive reserve capacity. The speed and attention task should be attempted at baseline; those patients who cannot complete it due to physical complaints are those who should be considered at risk for mild impairment. Further research should address the possible pathogenesis of these attentional deficits in specific patients.

Future research should embody uniform methods of data-analysis in which inevitable phenomena in longitudinal research such as missing values and practice effects can be accounted for. Multilevel analyses enable the use of all outcome measures for an individual patient over time. Guidelines should be developed addressing these specific issues, aimed at improving the quality of future studies.

Closing remark

Our nationwide, prospective, longitudinal, sibling-controlled study demonstrates that the use of chemotherapy-only in children with ALL has no evident effects on cognition. Our results are reassuring for patients, parents and professional caregivers. However, minor deficits can not be precluded in certain risk groups and future research with similar designs employing larger samples should further address this question. Prospective-longitudinal study designs like our study should become the standard for evaluating possible treatment effects as they allow us to examine real cognitive change.

References

1. Pui CH, Evans WE: Treatment of acute lymphoblastic leukemia. N Engl J Med 354:166-178,2006
2. Roman DD, Sperduto PW. Neuropsychological effects of cranial radiation: current knowledge and future directions. Int J Radiat Oncol Biol Phys 13(4):983-998,1995
3. Waber DP, Mullenix PJ. Neuropsychological sequelae in childhood acute lymphoblastic leukemia (ALL). In: Yeates KO, Ris MD, Taylor HG. eds. Pediatric Neuropsychology: Research, Theory and Practice. Guilford Publications, 1999
4. Mennes M, Stiers P, Vandenbussche E, et al: Attention and information processing in survivors of childhood acute lymphoblastic leukemia treated with chemotherapy only. Pediatr Blood Cancer 43:1-8,2004
5. Gidding CE, Kellie SJ, Kamps WA et al: Vincristine revised. Crit Rev Oncol Hematol 29:267-287,1999
6. Khong PL, Leung LH, Fung AS et al: White matter anisotropy in post-treatment childhood cancer survivors: preliminary evidence of association with neurocognitive function. J Clin Oncol 24 (6):88-90,2006
7. Kraemer S. The fragile male. Brit Med J 321:1609-12,2000
8. Waber DP, Urion DK, Tarbell NJ, et al. Late effects of central nervous system treatment of acute lymphoblastic leukemia in childhood are sex-dependent. Dev Med Child Neurol 1990;32:238–248.
9. Von derWeid N, Mosimann I, Hirt A, et al. Intellectual outcome in children and adolescents with acute lymphoblastic leukaemia treated with chemotherapy alone: Age- and sex-related differences. Eur J Cancer 39:359–365,2003
10. ICF-CY, Nederlandse vertaling van de International Classification of Functioning, Disability and Health, versie voor kinderen en jongeren (2008). WHO-FIC Collaborating Centre. Bilthoven: RIVM

SUMMARY

Summary

Acute lymphoblastic leukaemia (ALL) is the most common form of malignancy in children. With current survival rates of approximately 85%, increasing attention is paid to the potentially adverse effects of treatment on quality of life in children with ALL. Until the mid-eighties, prophylactic central nervous system irradiation was an essential element of leukaemia treatment but this irradiation has been associated with deleterious effects on cognitive abilities. With the introduction of 'chemotherapy-only' protocols which replaced irradiation by intrathecal and high dose systemic chemotherapy, the number of patients with cognitive deterioration has dramatically decreased with stable or even improved survival rates. However, despite a more favourable neuropsychological outcome after chemotherapy-only protocols compared to protocols including central nervous irradiation, there is no consensus concerning possible more subtle negative effects of chemotherapy on intellectual development and specific neuropsychological functions such as attention, memory, and visual motor abilities. Prior studies concerning neuropsychological outcome after chemotherapy-only in children with ALL had major methodological limitations.

In 1999, the nationwide study presented in this thesis was therefore initiated by the Paediatric Oncology Centre of the University Medical Centre Groningen. Main objective was to assess potential early and late negative late effects of chemotherapy-only on cognitive functioning in children treated for ALL. In this prospective longitudinal study, 50 newly diagnosed consecutive ALL patients aged 4 to 12 years treated according to Dutch Childhood Oncology Group (DCOG) ALL-9 protocol were included. Neuropsychological functioning of patients (median age at baseline testing 6.6 years) and 29 sibling-controls (median age at baseline testing 8.2 years) was assessed by an extensive test battery shortly after diagnosis, a few months after cessation of therapy and two years later. Repeated evaluations included measures of intelligence, memory, attention, visual-constructive function and fine-motor abilities. We applied hierarchical regression analyses and multilevel analyses to assess possible changes and differences over time compared to siblings. Multilevel technique allowed for the inclusion of patients with missing data in the analyses. Patients' results were additionally compared to population norms to assess whether test performance was in the normal range.

In **Chapter 1**, the rational of this study with a short introduction to ALL treatment was given.

In **Chapter 2**, baseline assessment of neuropsychological functioning in newly diagnosed patients at a median age of 6.6 years was described. To determine late potentially negative effects of chemotherapy on neurocognitive functioning, adequate and reliable assessment of the child's cognition prior to treatment is essential. However, it had been generally assumed that neuropsychological assessment in seriously sick children with a life threatening disease was not feasible and prospective studies are therefore rare. We performed an assessment in newly diagnosed patients with ALL within two weeks after start of chemotherapeutic treatment and demonstrated that neuropsychological assessment of patients during early hospitalisation is feasible and reliable. Mean scores on intelligence (verbal, performance and full-scale IQ) and specific neuropsychological tasks were not significantly different between patients and siblings. Both groups scored average compared to population norms, indicating no adverse effect of illness and psychological factors on cognitive functioning in patients with recently diagnosed ALL.

In **Chapter 3**, the early effects of chemotherapy-only on intelligence were described comparing results shortly after diagnosis to test results after cessation of two years of therapy. At both assessments, patients showed average performance on intelligence tests compared to population norms. Longitudinal analysis and cross-sectional comparisons revealed no significant differences between patients and controls. We concluded that no evident negative effects on intelligence were found, despite intensive and potentially neurotoxic treatment. Although young patients (< 6 years of age) showed stable performances on performance-IQ whereas siblings' scores increased, later analyses (Chapter 4) showed no differences.

Chapter 4 presented the long-term effects of chemotherapy-only on intellectual functioning. Verbal, performance and full-scale IQ were assessed three times; shortly after diagnosis, a few months after cessation of therapy and 2 years later (4½ year after diagnosis). Longitudinal analyses showed no major differences between patients and controls in scores over time. Patients and controls showed average scores on intelligence

tests compared to population norms. Remarkably, higher scores on performance IQ and full-scale IQ were observed for both patients and siblings after multiple assessments, a phenomenon which has been attributed to practice effects. A subgroup of patients (n=13) who uttered physical complaints at the baseline assessment scored significantly lower than siblings on full-scale IQ and performance IQ after treatment (both at the second and the third assessment).

Chapter 5 described the findings of the effects of chemotherapy-only on specific neuropsychological functions shortly after diagnosis, a few months after cessation of therapy and two years later. Multilevel analyses were applied to evaluate patients' performance over time and to compare patients to siblings and to normative data. No major differences were found for attention, memory, language, visual-constructive function and fine-motor abilities between patients and siblings, with both groups performing mainly in the normal range. The patient group as a whole, however, scored significantly lower than siblings on complex fine-motor functioning at the last evaluation, although this was still in the normal range. We hypothesized that the difference in fine-motor development was a result of peripheral vincristine neuropathy.
Large practice effects were found for both patients and siblings on four of eleven measures, including Rey´s Auditory-Verbal Learning Test (learning and delayed recall) and the Wisconsin Card Sorting Test (errors and perseverations). Contrary to earlier studies, gender and young age at diagnosis appeared to be no risk factor for a poorer neuropsychological performance. However, a possible new risk factor emerged from our study: Patients who uttered physical complaints (pain and/or tiredness) at the baseline assessment scored significantly lower than siblings on a sustained attention task at the last two evaluations.

Chapter 6 gave a review of the literature on cognitive functioning in children with ALL treated with chemotherapy-only, in relation to methodological variability and limitations. A literature search and selection according to strict criteria yielded 26 original studies. Like mentioned before, earlier studies had important methodological limitations, including no prospective-longitudinal design, the absence of a proper control group and incomplete assessments. Furthermore, methods of studies and applied psychometric instruments widely varied. Seven studies had a prospective and/or longitudinal design, with matched healthy

controls in three of these seven studies. These three reports (one concerned the study described in this thesis) with the most optimal design failed to find evident adverse effects of chemotherapy-only on cognition, apart from deterioration of complex fine-motor functioning. The 4 remaining longitudinal reports generally showed modest decline for a variety of functions. Nineteen cross-sectional studies found a spectrum of neuropsychological deficits but had important methodological limitations.

Chapter 7 offered a general discussion, conclusions and prospective for future research. First, we discussed the strengths and weaknesses of our study. Strengths include a multi-centre, prospective, longitudinal sibling-controlled study design with a comprehensive assessment of cognitive functions and a low percentage of attrition. A weakness is the possibly limited generalisability of our results to patients younger than four years of age. However, it is hard to account for this limitation since most neuropsychological tests have age restrictions. Another limitation was the relatively high number of missing patients at inclusion of the study. However, this did not concern selective attrition with regards to patient or disease characteristics, therefore, a bias in outcome does not seem likely.

It was shown that neuropsychological assessment is feasible in children in the acute phase of ALL. Despite intensive and potentially neurotoxic treatment, no evident, negative effects of chemotherapy-only were demonstrated on intelligence and specific neuropsychological functions (like learning, memory, attention, speed, executive functioning, visual-constructive and fine-motor functioning). However, 4½ year after diagnosis, patients still scored significantly worse compared to controls on a complex fine-motor task. We could not identify gender and young age as risk factors for poorer performance. Nevertheless, we did demonstrate that patients who uttered physical complaints at the first assessment and could not perform all tests, still performed more poorly than their controls at a test of sustained attention 4½ years later. They also performed poorer on full-scale IQ and performance IQ than their siblings.

Our results are reassuring for both patients with ALL and their parents, and oncologists and other healthcare professionals. General and specific cognitive functioning continues to develop normally up to 4½ year after diagnosis in patients treated with chemotherapy-only. However, scores on a complex fine-motor task at the last evaluation in the patient group were significantly worse than those of the healthy controls, but they were still in the

average range. Criteria are needed to define clinically relevant changes in performances, since no consensus exists about when a certain difference should be considered as clinically relevant.

Future prospective, longitudinal studies should include a larger proportion of very young patients to determine whether this group of patients is more vulnerable for negative effects of the treatment. Other risk factors, such as dose of chemotherapeutic agents need further elucidation. Theoretically, subtle problems on the long-term can not be fully excluded. Future research should employ uniform methods of data-analysis in which phenomena of longitudinal research such as missing values and practice effects can be optimally accounted for. Multilevel analyses particularly appear suitable for this type of research.

Conclusions were: 1. No evident, negative, neuropsychological, late effects of chemotherapy-only were found 4½ years after diagnosis, apart from complex fine-motor functioning. 2. The large practice effects should be reckoned with in future prospective longitudinal neuropsychological research, which emphasizes the need for a proper control group. 3. Patients who uttered physical complaints at the baseline neuropsychological assessment appear to be at risk for poorer performance of sustained attention and performance IQ. This may allow for timely detection of patients at risk, and may give opportunities for early intervention.

SAMENVATTING

Samenvatting

Acute lymfatische leukemie (ALL) is de meest voorkomende vorm van kanker bij kinderen. Met een huidig overlevingspercentage van ongeveer 85% wordt er in toenemende mate aandacht geschonken aan de mogelijke negatieve effecten van behandeling op de kwaliteit van leven van kinderen met ALL. Profylactische behandeling van het centraal zenuwstelsel (CZS) in de vorm van schedelbestraling was tot midden jaren '80 een essentieel element van de leukemiebehandeling, maar deze behandeling is geassocieerd met schadelijke effecten op de cognitieve capaciteiten. Met de introductie van 'alleen chemotherapie' protocollen waarin de bestraling is vervangen door intrathecale en hooggedoseerde systemische chemotherapie, is het aantal patiënten met cognitieve schade aanzienlijk afgenomen bij gelijk blijvende of zelfs verbeterde overleving. Echter, ondanks een gunstigere uitkomst van het neuropsychologisch functioneren na behandeling met alleen chemotherapie in vergelijking met protocollen met schedelbestraling, bestaat er geen consensus over de mogelijke meer subtiele negatieve effecten van chemotherapie op de intellectuele ontwikkeling en specifieke neuropsychologische functies. Eerdere studies naar dit onderwerp laten belangrijke methodologische tekortkomingen zien.

In 1999 is dan ook de landelijke studie zoals gepresenteerd in deze thesis geïnitieerd door het Kinderoncologisch Centrum Groningen (KOCG) van het Universitair Medisch Centrum Groningen. Het belangrijkste doel was om potentiële vroege en late negatieve effecten van alleen chemotherapie op het cognitief functioneren van kinderen met ALL te onderzoeken. In deze prospectieve longitudinale studie werden 50 nieuw gediagnosticeerde patiënten geïncludeerd met ALL tussen de leeftijd van 4 en 12 jaar die werden behandeld volgens het de Stichting KinderOncologie Nederland (SKION) protocol 9. Het neuropsychologisch functioneren van de patiënten (mediane leeftijd bij de eerste meting 6.6 jaar) en 29 broertjes en zusjes als controle (mediane leeftijd bij de eerste meting 8.2 jaar) werd gemeten met behulp van een uitgebreide testbatterij vlak na de diagnose, een aantal maanden na beëindiging van de therapie en twee jaar later. Herhaald neuropsychologisch onderzoek werd gedaan op het gebied van intelligentie, geheugen, aandacht, visueel-constructief functioneren en fijne motoriek. Mogelijke veranderingen in prestaties door de tijd heen werden onderzocht middels hiërarchische regressie analyse en multilevel analyse en vergeleken met de prestaties van controles. Multilevel methoden stonden toe dat ook

patiënten met ontbrekende data geïncludeerd konden worden in de analyses. Resultaten van patiënten werden daarnaast vergeleken met bestaande normgegevens om te onderzoeken of de testprestaties binnen het normale bereik vielen.

In **Hoofdstuk 1** werd de aanleiding voor het opzetten van het onderzoek en een korte introductie van de behandeling van ALL gegeven.

In **Hoofdstuk 2** werden de resultaten van het eerste neuropsychologisch onderzoek van nieuw gediagnosticeerde patiënten bij een mediane leeftijd van 6.6. jaar beschreven. Om eventuele late, negatieve effecten van chemotherapie op het cognitief functioneren te bepalen is een adequate en betrouwbare meting van het cognitief functioneren voor de behandeling essentieel. Echter, er wordt vaak aangenomen dat neuropsychologisch onderzoek bij ernstig zieke kinderen met een levensbedreigende ziekte in de acute fase niet haalbaar is en daarom zijn prospectieve studies zeldzaam. Wij onderzochten nieuw gediagnosticeerde kinderen met ALL binnen 2 weken na start van de chemotherapeutische behandeling en toonden aan dat neuropsychologisch onderzoek van patiënten gedurende de eerste ziekenhuisopname haalbaar en betrouwbaar is. Gemiddelde scores op zowel intelligentie (verbaal, performaal en totaal IQ) als specifieke neuropsychologische taken verschilden niet significant tussen patiënten en controles. Beide groepen scoorden gemiddeld ten opzichte van de bestaande normgegevens, wat erop wijst dat er geen negatieve effecten van ziekte en psychologische factoren zijn op het cognitief functioneren van patiënten met recent gediagnosticeerde ALL.

In **Hoofdstuk 3** werden de vroege effecten van chemotherapie op de intelligentie beschreven, waarbij resultaten van intelligentieonderzoek vlak na diagnose werden vergeleken met die kort na staken van de twee jaar durende behandeling. Op beide metingen lieten patiënten gemiddelde scores zien op intelligentietesten in vergelijking met normgegevens. Zowel in de longitudinale analyse als de cross-sectionele vergelijkingen, werden geen significante verschillen gevonden tussen de prestaties van patiënten en controles. Ondanks de intensieve en potentieel toxische behandeling werden geen evidente negatieve effecten op de intelligentie gevonden. Alhoewel jonge patiënten (< 6 jaar) een

gelijkblijvend performaal IQ lieten zien terwijl de scores van de controles toenamen, lieten latere analyses (Hoofdstuk 4) geen verschillen meer zien.

Hoofdstuk 4 presenteerde de lange-termijn effecten van alleen chemotherapie op het intellectueel functioneren. Het verbaal, performaal en totaal IQ werd drie keer gemeten, te weten kort na diagnose, na staken van de behandeling en twee jaar later (4½ jaar na diagnose). In de longitudinale analyses werden geen grote verschillen gevonden tussen patiënten en controles door de tijd heen. Zowel de patiënten als controles lieten gemiddelde scores zien op intelligentietesten in vergelijking met normgegevens. Opvallend was dat in zowel de groep patiënten als de groep controles hogere scores bij herhaalde metingen werden gevonden voor zowel het performaal IQ als het totaal IQ, een fenomeen dat werd toegeschreven aan leereffecten. Een subgroep van patiënten (n=13) die bij de eerste meting fysieke klachten uitten, lieten in vergelijking met controles op de twee laatste metingen significant lagere scores zien op het performaal en totaal IQ.

In **Hoofdstuk 5** werden de bevindingen van het onderzoek naar de effecten van alleen chemotherapie op het specifiek neuropsychologisch functioneren beschreven vlak na diagnose, een aantal maanden na staken van de twee jaar durende behandeling en twee jaar later. Multilevel analyses werden gebruikt om de prestaties van patiënten op de verschillende meetmomenten te evalueren en te vergelijken met de prestaties van controles en normgegevens. Er werden geen belangrijke verschillen gevonden in prestaties van patiënten en controles wat betreft aandacht, geheugen, taal, visueel-constructief functioneren en fijne motoriek. Beide groepen scoorden binnen het normale bereik. Echter, op de laatste meting scoorde de gehele patiëntengroep significant lager dan de controles op een complexe fijn-motorische taak, hoewel de score nog wel gemiddeld was in vergelijking met normgegevens. Onze hypothese was dat het verschil in fijn-motorische ontwikkeling het gevolg zou kunnen zijn van een perifere vincristine neuropathie.
Bij vier van de elf afgenomen maten werden grote leereffecten gevonden bij zowel de patiënten als de controles, namelijk de Rey auditieve verbale leertaak (inprenting en uitgestelde reproductie) en de Wisconsin Card Sorting Test (fouten en perseveraties). In tegenstelling tot eerdere studies, bleken geslacht en een jonge leeftijd bij diagnose geen risicofactoren voor een slechtere neuropsychologische prestatie. Echter, een mogelijk

nieuwe risicofactor werd wel gevonden in onze studie: Patiënten die bij de eerste meting fysieke klachten uitten (pijn en/ of vermoeidheid) scoorden op de laatste twee metingen significant lager dan controles op een volgehouden aandachtstaak.

Hoofdstuk 6 gaf een overzicht van de literatuur over het cognitief functioneren van kinderen met ALL die behandeld zijn met alleen chemotherapie in relatie tot methodologische variatie en beperkingen. Een literatuuronderzoek en –selectie volgens strikte criteria leverde 26 studies op. Zoals eerder genoemd, hadden eerdere studies belangrijke methodologische beperkingen, zoals geen prospectief-longitudinale studie-opzet, het ontbreken van een geschikte controlegroep en incomplete metingen. Bovendien varieerden de methoden en toegepaste psychometrische instrumenten aanzienlijk. Zeven studies hadden een prospectieve en/of longitudinale studieopzet; drie hiervan hadden een controlegroep met gematchede gezonde controles. Deze drie studies met het meeste optimale studie-ontwerp vonden geen belangrijke negatieve effecten van alleen chemotherapie op de cognitie behalve een achteruitgang van het complex fijn-motorisch functioneren. De 4 overige longitudinale studies lieten over het algemeen een matige achteruitgang zien voor een verscheidenheid aan functies. Negentien cross-sectionele studies vonden een spectrum aan neuropsychologische defecten, maar hadden belangrijke methodologische tekortkomingen.

Hoofdstuk 7 bood een algemene discussie, conclusies en aanbevelingen voor toekomstig onderzoek. Ten eerste werden de sterke en zwakke punten van ons onderzoek besproken. Sterke punten zijn een multi-center, prospectieve, longitudinale gecontroleerde studie opzet, het gebruik van een uitgebreide testbatterij en zeer weinig verlies van patiënten gedurende de follow-up. Een zwakte is de mogelijke beperking in de generaliseerbaarheid van de resultaten van onze studie voor patiënten jonger dan vier jaar. Echter, als gevolg van de leeftijdsrestrictie van de meeste testen is deze beperking moeilijk te ondervangen. Het relatief hoge aantal gemiste patiënten bij de inclusie was een andere beperking, maar dit betrof geen selectieve uitval ten aanzien van patiënt of ziektekenmerken.
Er kon worden aangetoond dat neuropsychologisch onderzoek haalbaar en betrouwbaar is af te nemen gedurende de acute fase van ALL. Ondanks intensieve en potentieel neurotoxische behandeling werden er 4½ jaar na diagnose geen evident negatieve effecten

van alleen chemotherapie op intelligentie en specifieke neuropsychologische functies (zoals leren, geheugen, aandacht, snelheid, executief functioneren, visueel-constructief functioneren en fijne motoriek) gevonden. Echter, 4½ jaar na diagnose scoren patiënten nog steeds significant lager dan controles op een complex fijn-motorische taak. Wij konden geslacht en jonge leeftijd niet als risicofactoren identificeren. Niettemin werd aangetoond dat patiënten die bij de eerste meting fysieke klachten uitten en daardoor niet alle testen konden doen, in vergelijking met controles nog steeds zwakkere scores lieten zien op een volgehouden aandachtstaak 4 ½ jaar later. Deze patiënten scoorden ook slechter op het totaal en performaal IQ in vergelijking met controles.

Onze resultaten zijn geruststellend voor zowel patiënten met ALL en hun ouders, alsmede oncologen en andere professionele zorgverleners. Patiënten die zijn behandeld met alleen chemotherapie blijken zich tot 4½ jaar na diagnose normaal te ontwikkelen op het vlak van het algemeen en specifiek cognitief functioneren. Echter, bij de laatste meting waren de scores op een complex fijn-motorische taak significant slechter dan die van gezonde controles, alhoewel de patiëntengroep nog steeds binnen het normale bereik scoorde. Er zijn criteria nodig om klinisch relevante veranderingen in prestaties te definiëren, aangezien er geen consensus bestaat over wanneer een bepaald verschil moet worden beschouwd als klinisch relevant.

In toekomstig prospectief, longitudinaal onderzoek zouden meer jonge kinderen moeten worden geïncludeerd om te bepalen of deze groep patiënten kwetsbaarder is voor negatieve effecten van de behandeling. Andere risicofactoren zoals dosering van chemotherapie moeten verder worden onderzocht. In theorie kunnen subtiele problemen op de lange termijn niet geheel worden uitgesloten. Toekomstig onderzoek zou uniforme methodes van data-analyse moeten toepassen waarin voor problemen van longitudinaal onderzoek zoals ontbrekende waardes en leereffecten, zo optimaal mogelijk gecorrigeerd kan worden. Multilevel analyses lijken bij uitstek geschikt voor dit soort onderzoek.

De conclusies luidden: 1. 4½ jaar na diagnose werden geen evident, negatieve neuropsychologische late effecten van alleen chemotherapie gevonden, met uitzondering van het complex fijn-motorisch functioneren. 2. In toekomstig prospectief, longitudinaal onderzoek zou rekening moeten worden gehouden met de grote leereffecten, wat de behoefte aan een geschikte controlegroep onderstreept. 3. Patiënten die bij de eerste meting

fysieke klachten uitten, lopen een mogelijk risico op een slechtere prestatie voor wat betreft volgehouden aandacht en het performale IQ. Dit gegeven zou gebruikt kunnen worden om patiënten die mogelijk een risico lopen, vroeg op te sporen zodat op tijd een interventie kan worden toegepast.

Dankwoord

Een decennium wordt met de totstandkoming van dit proefschrift afgesloten. Velen hebben de afgelopen jaren een belangrijke bijdrage hieraan geleverd. Graag wil ik hen op deze plaats bedanken.

Allereerst natuurlijk alle kinderen met leukemie, hun broertjes en zusjes, en hun ouders die hebben deelgenomen aan deze studie. Het was bijzonder om jullie in de acute fase van het ziek zijn te ontmoeten en jullie een aantal jaren te mogen volgen zodat belangrijke data kon worden verzameld. Zonder jullie was dit proefschrift er niet geweest.

Mijn eerste promotor prof. dr. Kamps, beste Willem, dank voor je begeleiding. Vanaf de start van het project straalde jij rust en vertrouwen uit. Jouw korte, bondige commentaar hielpen mij altijd weer op weg.

Prof. dr. Bouma, beste Anke, van jou kreeg ik altijd snel en uitgebreid commentaar. Bedankt voor het nauwkeurig lezen van de artikelen en je suggesties.

Co-promotor, dr. Kingma, beste Annet, jij stond aan de wieg van dit project. Jouw immer kritische en scherpe geest hebben mij veel inzichten verschaft. Ik ben blij dat ik zoveel van je heb mogen leren. Bedankt voor je vertrouwen en intensieve betrokkenheid. Lang leve het internet en de telefoon. Ik kijk met plezier terug op het naborrelen en het bezoeken van de vele Groningse eetgelegenheden.

De beoordelingscommissie bestaande uit prof. dr. R. Pieters van het Erasmus MC en prof. dr. O. Brouwer en prof. dr. W.H. Brouwer van het UMCG, wil ik graag bedanken voor hun bereidheid om plaats te nemen in de commissie.

Dankzij de financiële steun van KWF Kankerbestrijding kon dit onderzoek worden uitgevoerd.

Drs. Arnout Schuitema, zeer veel dank voor het met raad en daad bijstaan bij het bedenken, uitvoeren en interpreteren van allerlei ingewikkelde statistische analyses. In de samenwerking met jou is de statistiek voor mij gaan leven. Bedankt!

De co-auteurs van de artikelen wil ik hartelijk danken voor hun deskundige bijdrage. Prof. dr. Anjo Veerman, dr. Jaap Huisman, dank voor jullie waardevolle adviezen. Speciale dank gaat uit naar drs. Rieneke van Dommelen die de pilot van mijn onderzoek heeft gedaan en mij ingewerkt heeft. Dank je wel!

De psychologen van de afdeling kinderoncologie van de academische ziekenhuizen met wie ik ten tijde van het onderzoek te maken heb gehad, Jaap Huisman, Martha Grootenhuis, Ansje van Veldhuizen, Esther Meijer en Jantien Vrijmoet-Wiersma, dank voor de introducties aan patiënten en jullie belangstelling voor het onderzoek.

Mijn (oud) collega's van de afdeling Medische psychologie & Maatschappelijk werk wil ik hartelijk danken voor de collegialiteit en de mentale steun. Ik kan het niet nalaten om een aantal in het bijzonder te noemen. Allereerst mijn oud kamergenoot Meike, wat jammer dat je nu ergens anders werkt, Aletta, Huub, Barbara, Maaike, Marielle, Annemieke en Karijn, bedankt voor jullie gezelligheid, betrokkenheid en goede werksfeer. Prof. dr. Gerben Sinnema bedankt voor je interesse en de ruimte die je me gegeven hebt om het proefschrift in goede orde af te ronden.

Veel dank ook aan mijn collega's van de sector Neuropsychologie. Harmke, Joost, Lili, Monique, Olga en Stella, bedankt voor jullie interesse en de gezelligheid op de werkvloer. Dr. Aag Jennekens-Schinkel, jouw passie voor het vak is bewonderenswaardig. Bedankt voor je steun, inspiratie en de leerzame discussies. Heel veel dank!

Lieve Ceciel, naast een verre collega, was je ook een goede vriendin. Na een moeilijk overleg kon ik bij jou stoom afblazen. Ik wil je bedanken voor je vriendschap en support.

Lieve familie, schoonfamilie en vrienden, bedankt voor de steun en afleiding in de afgelopen jaren en jullie begrip dat ik bij momenten iets minder sociaal was. Maurice, dank voor de snelle hulp bij het bewerken van de omslag.

Mijn zussen en paranimfen Daniëlle en Angelique, ik ben blij dat jullie mij bijstaan. De drie gezusters zijn nu eindelijk allen gepromoveerd! Lieve Daniëlle, in de beginfase heb ik vaak bij jou in mijn 'eigen' kamertje geslapen en werd ik gewekt door mijn neefje of nichtje, om vroeg in het UMCG patiëntjes te kunnen zien. Bedankt! Lieve Liek, naast een analytische en kritische geest, ben jij steunend en creatief. Juist op de momenten dat ik het onderzoek niet meer zag zitten, kon jij me weer motiveren.

Lieve mam, bedankt, zonder jou was ik niet zover gekomen. Jij kon zaken voor mij relativeren en zorgde er vaak weer voor dat ik met goede moed verder kon en bleef volhouden. Ik ben heel blij met jouw nuchterheid, steun en vertrouwen.

Lieve Vincent, ik ben heel blij met jou! Bedankt voor je liefde en onvoorwaardelijke steun. We gaan samen weer leuke dingen doen!

Curriculum vitae

Nathalie Jansen werd op 22 maart 1972 in de Noordoostpolder geboren. Na het behalen van haar VWO diploma in 1991, behaalde zij de propedeuse HBO-verpleegkunde om vervolgens in 1992 te starten met de opleiding Psychologie aan de Rijks Universiteit Leiden. Bij het kiezen van haar afstudeerrichting, stapte zij over naar de Universiteit van Amsterdam, om zich aan deze universiteit verder te scholen in de klinische psychologie met als studieroute klinische psychobiologie en neuropsychologie. Na het behalen van haar diploma in 1997 startte zij in 1998 met een werkervaringsplaats als psycholoog op de afdeling ontwikkelingsneurologie van het Academisch Ziekenhuis Kinderen van Vrije Universiteit Brussel. In 1999 begon zij aan het landelijk onderzoek dat gecoördineerd werd vanuit het universitair medisch centrum Groningen onder leiding van prof. dr. W. Kamps, prof. dr. A. Bouma en dr. A. Kingma en beschreven is in dit proefschrift. Deze werkzaamheden heeft zij enkele jaren gecombineerd met een freelance baan als psycholoog voor een psychologisch en psychiatrisch adviesbureau en een baan als psycholoog binnen Curium Academisch Centrum kinder- en jeugdpsychiatrie te Oegstgeest, afdeling psychodiagnostiek. In mei 2002 deed de kans zich voor om haar onderzoekswerkzaamheden te combineren met een baan binnen de afdeling medische psychologie en maatschappelijk werk en de sector neuropsychologie van het Wilhelmina Kinderziekenhuis Utrecht, waar zij tot op heden nog steeds werkt. In 2002 behaalde zij haar registratie als gezondheidszorgpsycholoog en in 2007 begon zij aan de opleiding tot klinisch psycholoog kinder- en jeugd bij de Regionale Instelling voor Nascholing en Opleiding (RINO) te Utrecht die zij eind 2010 verwacht af te ronden.