MATHEMATICS IN HEALTH CARE WITH APPLICATIONS

By

Luca Paulon

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
MATHEMATICAL MODELS AND METHODS FOR TECHNOLOGY AND SOCIETY
AT
BASIC AND APPLIED SCIENCES FOR ENGINEERING (SBAI),
SAPIENZA UNIVERSITY
ROME, ITALY
OCTOBER 2012

SAPIENZA UNIVERSITY

DEPARTMENT OF

BASIC AND APPLIED SCIENCES FOR ENGINEERING (SBAI)

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "**Mathematics in Health Care with Applications**" by **Luca Paulon** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy** in **Mathematical Models and Methods for Technology and Society**.

Dated: October 2012

External Examiner: _____
Prof. Bernadette Bouchon-Meunier

Research Supervisor: _____
Prof. Francesca Pitolli

Examining Committee: _____
Prof. Daniela Giachetti

_____
Prof. Giulianella Coletti

_____
Prof. Piero Negrini and Prof. Gerardo Toraldo

SAPIENZA UNIVERSITY

Date: **October 2012**

Author: **Luca Paulon**

Title: **Mathematics in Health Care with Applications**

Department: **Basic and Applied Sciences for Engineering (SBAI)**

Degree: **Ph.D.** Convocation: **December-February** Year: **2012-2013**

Signature of Author

To all Families

and in Memory and Honor of

Ruggero Volterra (1927-2010),

Maria Rosaria Garisto in Gazzetta (1974-2012),

Patrizia Ragnoli in Laudi (1964-2013).

Table of Contents

Abstract

The Author aims to show how mathematics can be useful in supporting key activities in a hospital, including: noninvasive measurement of a patient's status (see chapter 1), evaluation of quality of services (see chapter 2), business and clinical administration (see chapter 3), and diagnosis and prognosis (see chapter 4). Such applications suggest the development of innovative projects to improve health care processes, services and systems. In this way, mathematics can be a very important tool for technological and societal development.

Acknowledgements

My gratitude goes to my supervisor Prof. Francesca Pitolli and professors Angelo Gilio, Patrizio Pasqualetti, Romano Scozzafava, Stefano Serra Capizzano, Barbara Vantaggi, for their many suggestions during this research.

Then I should mention that my graduate studies were supported in part by AFaR, that is an Italian national association for biomedical and health care research owned by the Brothers of Saint John of God (known in Italy as *Fatebenefratelli*).

I wish to thank also the following (and they know why): Elisabetta Berrettoni, Daniela Calvetti, Alessandra Campolongo, Carlo Maria Cellucci, Mauro Cenerelli, Carlo Ciampa, Monica Cupani, Federico D'Alessio, Giovanna D'Ari, Michele De Lucia, Maurizio Ferrante, Cecilia Filardo, Elisabetta Franco, Luca Frangipane, Massimo Frangipane, Monica Frangipane, Salvatore Frangipane, Maria Rosaria Garisto and Maurizio Gazzetta, Giovanni Laffi, Francesco Lagana, Lorena Lagana, Jasmine Laudi, Paolo Laudi, Sara Lanfredi, Dario Manfellotto, Massimo Maurici, Mauro Messedaglia, Daniele Milazzo, Marco Muzzi, Patrizia Ragnoli, Francesco Rinaldi, Paolo Maria Rossini, Mario Paulon, Zuliema Paulon, Franco Piredda, SiliconDev srl, Erkki Somersalo, Alfonso Teodori, Addolorata Vassallo.

A special thanks goes to Alessandro Giordani for his big effort in supporting my work. I would say he is more a friend than a colleague.

Last, but non least, I am so much grateful to my wife Bruna Frangipane, who I love so much "*...more than yesterday and less than tomorrow*", and all my family for their patience and *love*, i.e., my mom Elda Zuppante, my father Domenico Paulon, my sister Lorena Paulon and her husband Fabrizio Gentile, my godfather Ruggero Volterra, my wife's parents, particularly Teresa Misuraca and Giuseppe Frangipane. Without them this work would never have come into existence.

Rome, Italy
October, 2012

Luca Paulon

Introduction

This thesis is a collection of health care applications of mathematical methods and models from various disciplines of mathematics. In this way, the Author intends to show the necessity and benefits of multidisciplinary knowledge as well as the importance of an attitude to drive innovation by means of mathematical tools, using probability calculus, numerical analysis, operations research, and other methods.

The integration of mathematics and real-life problems can be very effective. Advanced mathematics tools allow problems to be solved in innovative ways. From a theoretical point of view, each application may suggest if and how (in what direction) the theory on which it is based can be extended. This effort can result not only in the publication of scientific papers but also in the awarding of grants to continue such research activities. This thesis is a proof of these claims. Scientific publications were submitted by the Author in collaboration with others. Furthermore, a grant of 300 thousand euros has been obtained from the Lazio Region of Italy for an innovation project called Business Simulation for Health Care (BuS-4H) that was written by the Author on the subject of business and clinical administration of a hospital.

This document is structured as stand-alone chapters that share a common structure: an introduction, background and algorithms sections and a *target* application.

In chapter 1, some results relevant to a fast solution of large ill-conditioned linear systems are illustrated and then used in a target application, i.e., a preliminary simulation case study for the neuroimaging of a patient's brain activity. Neuroimaging aims at localizing neuronal sources responsible for brain activity. Among neuroimaging techniques, magnetoencephalography (MEG) is particularly attractive because it is completely noninvasive and has high temporal resolution. MEG systems measure the magnetic field generated outside the skull by the synchronous activation of thousands of neurons. Thus, the localization of the neuronal sources of brain activity can be obtained by reconstructing the current flow image underlying the measured magnetic field. This process results in a highly ill-posed and ill-conditioned inverse problem that requires sophisticated numerical methods to be solved.

In chapter 2, some results relevant to the theory of coherent conditional probability are illustrated and then used in a target application, i.e., the evaluation of the quality of network vaccination centers (ASL1, ASL2, ASL3) in the Lazio Region of Italy. Thirteen vaccination centers offering pediatric vaccinations for children under three years of age and the anti-human papillomavirus vaccination for adolescent women were evaluated; these include: two centers in two of the four districts of the ASL1, two centers in one of the five districts of the ASL2 and nine centers in the four districts of the ASL3.

In chapter 3, some results related to the topic of multi-objective optimization are illustrated and then used in a target application, i.e., a case-mix analysis of an obstetrics and gynecology ward. Without requiring drastic changes in the ward, it is shown that it would be possible to reduce the length of hospitalization by one day for many patients who have no complications and whose current average length of

hospitalization is three days. It is also shown that the caesarean section rate could be decreased from 45% to 41%. This represents a first step toward the gold standard value of 20% requested by the Italian National Health System. At the same time, the ward's profit would increase by 11.6% with respect to the actual.

In chapter 4, some results related to survival analysis with competing risks are illustrated and then used in a target application, i.e., the identification of markers of increased risk of stroke in asymptomatic subjects with severe internal carotid artery (ICA) stenosis. In this example, 621 subjects with unilateral asymptomatic severe ICA stenosis were included and prospectively evaluated with a median follow-up of 27 months (min=6, max=68). In all patients, demographic and vascular risk profile, plaque characteristics and progression and common carotid artery intima-media thickness (IMT) were investigated. The outcome measures were subsequent occurrence of ischemic stroke ipsilateral to ICA stenosis and vascular death, while myocardial infarction, contralateral strokes and transient ischemic attack (TIA) were considered competing events. Accordingly, the variables potentially able to predict a stroke event were assessed by means of a survival model for competing risks.

In sum, the Author aims to show how mathematics can be useful in supporting key activities in a hospital, including: noninvasive measurement of a patient's status (see chapter 1), evaluation of quality of services (see chapter 2), business and clinical administration (see chapter 3), and diagnosis and prognosis (see chapter 4). Such applications suggest the development of innovative projects to improve health care processes, services and systems. In this way, mathematics can be a very important tool for technological and societal development.

Chapter 1

Neuroelectric source localization using fast solution of large linear systems

1.1 Introduction

Many research studies aim to develop techniques for successful diagnosis of patients with multiple sclerosis, Alzheimer's disease and other pathologies using magnetoencephalography (MEG). The goal of such studies is to support doctors, by means of noninvasive techniques, in making early diagnoses, i.e., in distinguishing patients from healthy control subjects. MEG complements other brain activity measurement techniques such as electroencephalography (EEG), positron emission tomography (PET) and functional magnetic resonance imaging (fMRI). This fact explains why an increasing number of hospitals are currently using MEG.

At a more detailed level, MEG systems measure the magnetic field generated outside the skull by the synchronous activation of thousands of neurons (i.e. brain areas). The localization of such areas and their dynamic activation can be used to characterize brain functionalities or dysfunctionalities; in this respect, MEG is a

fundamental tool for brain imaging.

From a mathematical point of view, the localization of neuronal sources underlying MEG measures is a highly ill-posed and ill-conditioned inverse problem that requires sophisticated numerical methods for solution.

Thus, in this chapter we face the problem of solving large ill-conditioned linear systems ($Ax = b$) that may be presented by real applications, for instance an MEG linear inverse problem. This kind of system can be practically solved using iterative methods. In this context, the conjugate gradient (CG) method is an iterative method that is particularly attractive because it can converge to an approximated solution $\hat{x}$ in a finite number of iterations regardless of the dimension of matrix A and because the error estimate, $\|x - \hat{x}\|$, is limited. Furthermore, the CG method is very efficient if the matrix A is sparse or if it can be compressed without introducing too much distortion in the starting problem. This can be done in a wavelet domain using the discrete wavelet transform (DWT) and suitable thresholding techniques.

The target application at the end of this chapter is a first step toward a concrete innovation goal, which is the tuning of an efficient and effective algorithm for neuronal source localization that could be implemented and embedded in an MEG recording system. This real-time brain imaging functionality can support researchers or clinicians during the monitoring of a subject's brain activity.

To realize this work the Author worked at the "Dipartimento di Scienze di Base e Applicate per l'Ingegneria " (SBAI) of the University of Rome "Sapienza", in collaboration with, as listed in a publication in preparation, Francesca Pitolli, Marco Muzi and Cecilia Filardo.

1.2 Background

1.2.1 Preliminaries

This section and the two following sections contain basic facts that are primarily based on book [1]. Further details can be found in [2].

Definition 1.2.1. The Euclidean scalar product

$$(x, y) = \sum_{i=1}^{n} x_i \, y_i$$

is a scalar product on the finite dimensional vector space $\mathbb{R}^n$, being $x = (x_1, ..., x_n) \in \mathbb{R}^n$ and $y = (y_1, ..., y_n) \in \mathbb{R}^n$.

Definition 1.2.2. The vector p-norm

$$\|x\|_p = \Big(\sum_{i=1}^{n} |x_i|^p \Big)^{1/p} \text{ , for } 1 \leq p < \infty$$

is a vector norm on $\mathbb{R}^n$.

Remark 1.2.1. The pair $(\mathbb{R}^n, \|\cdot\|_p)$ is called a normed space.

Definition 1.2.3. The Euclidian vector norm

$$\|x\|_2 = (x, x)^{1/2} = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2}$$

is a vector p-norm, with $p = 2$, on $\mathbb{R}^n$.

Remark 1.2.2. The Cauchy-Schwarz inequality

$$|(x, y)| \leq \|x\|_2 \|y\|_2$$

holds for any pair $x, y \in \mathbb{R}^n$.

Definition 1.2.4. The maximum norm

$$\|x\|_\infty = \max_{1\le i\le n} |x_i|$$

is a vector norm on $\mathbb{R}^n$.

Definition 1.2.5. Let $A \in \mathbb{R}^{n\times n}$ be a symmetric and positive definite matrix. The A-norm

$$\|x\|_A^2 = x^*\, A\, x$$

is a vector norm on $\mathbb{R}^n$.

Remark 1.2.3. From now on, x^* denotes the transpose of vector x and A^* the transpose of matrix A.

Definition 1.2.6. The Frobenius norm

$$\|A\|_F = \left(\sum_{i,j=1}^{n} |a_{i,j}|^2 \right)^{1/2} = \left(tr(A\, A^*) \right)^{1/2}$$

is a matrix norm compatible (or consistent) with the Euclidean vector norm, i.e. $\|Ax\|_2 \le \|A\|_F \|x\|_2$, $\forall x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n\times n}$.

Remark 1.2.4. Let I be the identity matrix of dimension n. Then $\|I\|_F = \sqrt{n}$.

Definition 1.2.7. Let $\|\cdot\|$ be a vector norm. The function

$$\|A\| = \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|}$$

is a matrix norm called the induced matrix norm or natural matrix norm. It is consistent with $\|\cdot\|$, i.e. $\|Ax\| \le \|A\|\|x\|$.

Definition 1.2.8. Let A be a square matrix and A^{-1} the inverse matrix. The condition number of A is

$$K(A) = \|A^{-1}\|\, \|A\| \,,$$

with $\|\cdot\|$ being an induced matrix norm. Note that $K(A) \ge 1$, $K(A) = K(A^{-1})$, and its definition depends on the choice of the norm $\|\cdot\|$.

1.2.2 Iterative methods

Let $(\mathbb{R}^n, \|\cdot\|)$ be a normed space and

$$Ax = b \tag{1.2.1}$$

a linear system with $A \in \mathbb{R}^{n\times m}$ and $b \in \mathbb{R}^n$.

The basic idea of iterative methods involves construction of a sequence of vectors $x^{(k)}$ that enjoy the following property of *convergence*

$$x = \lim_{k\to+\infty} x^{(k)} \Leftrightarrow \lim_{k\to+\infty} \|x - x^{(k)}\| = 0$$

where x is the solution to (1.2.1) and k a positive number called iteration number of the method. In general, an iterative method can be stated as

$$x^{(0)} = f_0(A, b)$$

$$x^{(k+1)} = f_{k+1}\ (\ x^{(k)}, x^{(k-1)}, \dots\ , x^{(k-m)},\ A\ , b\ )$$

with $m + 1 > 0$ the so–called 'order' of the method. If each f_i is independent of i, i.e. $\forall i = 1, 2, ..., k + 1$ it is $f_i = f$, the iterative method is said to be stationary. If $\forall i = 1, 2, ..., k + 1$ f_i is a linear function of $x^{(k)}, \dots\ , x^{(k-m)}$, the method is said to be linear.

The iterative process is stopped if the norm of the absolute error $\|e^{(k)}\| = \|x^{(k)} - x\|$ is less than or equal to a fixed tolerance $\varepsilon > 0$. If the exact solution is not available, it is necessary to introduce a suitable stopping criterion to monitor the convergence of the iteration. A practical stopping criterion in a program implementation consists of continuing the iteration until the residual $\|r^{(k)}\| = \|b - Ax^{(k)}\|$ or the normalized

residual $\|r^{(k)}\|/\|b\|$ is less or equal to ε, supposed to be greater than the machine epsilon.

Remark 1.2.5. We obtain the following control on the errors:

$$\|e^{(k)}\| = \|x - x^{(k)}\| = \|A^{-1}b - x^{(k)}\| = \|A^{-1}r^{(k)}\| \leq \|A^{-1}\|\varepsilon$$

and

$$\|e_r^{(k)}\| = \frac{\|x - x^{(k)}\|}{\|x\|} \leq \frac{\|A^{-1}\|\,\|r^k\|}{\|x\|} \leq K(A)\frac{\|r^k\|}{\|b\|} \leq \varepsilon\, K(A),$$

the matrix norm $\|\cdot\|$ being consistent with the vector norm $\|\cdot\|$, A a square matrix, $K(A)$ being the condition number of A and $e_r^{(k)}$ the relative error.

1.2.3 Iterative gradient methods

Let

$$Ax = b \tag{1.2.2}$$

be a linear system with a symmetric and positive definite matrix $A \in \mathbb{R}^{n\times n}$.

If A is not symmetric and positive definite, we can derive the so–called system of normal equations

$$A^*Ax = A^*b \tag{1.2.3}$$

that has a real and symmetric positive definite square matrix.

The following linear nonstationary iterative methods having the form

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}\ , \tag{1.2.4}$$

with α_k the length of an increment along a direction $p^{(k)}$ are defined to solve systems such as (1.2.2) or (1.2.3).

Proposition 1.2.1. *If the matrix A is symmetric and positive definite then solving system $Ax = b$ is equivalent to finding the minimizer $x \in \mathbb{R}^n$ of the quadratic form*

$$\Phi(y) = \frac{1}{2}y^*Ay - y^*b$$

which is called the energy of the system $Ax = b$.

Proof. See [1], page 134. □

The problem is thus to determine the minimizer x of Φ starting from a point $x^{(0)} \in \mathbb{R}^n$ and consequently to select suitable directions in which to move to get as close as possible to the solution x. The optimal direction that joins the starting point $x^{(0)}$ to the solution point x is obviously unknown *a priori*. Therefore, we must take a step from $x^{(0)}$ along another direction $p^{(0)}$, and then fix along this latter a new point $x^{(1)}$ from which to iterate the process until convergence. Thus, at the generic step k, $x^{(k+1)}$ is computed as in (1.2.4).

The most natural idea is to take the descent direction of maximum slope $\nabla\Phi(x^{(k)})$, which yields the *gradient method* or *steepest descent method*.
In fact the gradient of Φ is $\nabla\Phi(y) = \frac{1}{2}(A^* + A)y - b = Ay - b$. Then $\nabla\Phi(x^{(k)}) = Ax^{(k)} - b = -r^{(k)}$, i.e. the direction of the gradient of Φ coincides with the residual's direction and can be immediately computed using the current iterate. The gradient method moves at each step k along the direction $p^{(k)} = r^{(k)}$.
To compute the parameter α_k let us write explicitly $\Phi(x^{(k+1)})$ as a function of a parameter α

$$\Phi(x^{(k+1)}) = \frac{1}{2}(x^{(k)} + \alpha r^{(k)})^* A(x^{(k)} + \alpha r^{(k)}) - (x^{(k)} - \alpha r^{(k)})^* b.$$

Differentiating with respect to α and setting the result equal to zero, yields the desired value of α_k

$$\alpha_k = \frac{r^{(k)^*} r^{(k)}}{r^{(k)^*} A r^{(k)}},$$

which depends only on the residual at the k-th step.

A different approach is suggested by the following definition:

Definition 1.2.9. A direction $x^{(k)}$ is said to be *optimal* with respect to a direction $p \neq 0$ if

$$\Phi(x^{(k)}) \leq \Phi(x^{(k)} + \lambda p) \ , \ \forall \lambda \in \mathbb{R}.$$

To preserve optimality between successive iterates, the descent directions must be mutually *A-orthogonal* or *A-conjugate*, i.e. $p^* A q = 0$.

A method employing A-conjugate descent directions is called *conjugate*. The resulting conjugate gradient (CG) method is obtained by choosing the descendent directions

$$p^{(k+1)} = r^{(k+1)} - \beta_k p^{(k)}$$

mutually *A-orthogonal* using the parameter

$$\beta_k = \frac{(Ap^{(k)})^* r^{(k+1)}}{(Ap^{(k)})^* p^{(k)}}$$

and the acceleration parameter

$$\alpha_k = \frac{p^{(k)^*} r^{(k)}}{p^{(k)^*} A p^{(k)}} \, .$$

The complete CG algorithm is shown in section 1.3.

1.2.4 On the convergence of the conjugate gradient method

In the absence of rounding errors, the CG method, which is an iterative method, can be considered as a direct method because it terminates after a finite number of steps. In fact, the following result holds.

Theorem 1.2.2. *Let A be a symmetric and positive definite matrix and let λ_1, λ_n be its maximum and minimum eigenvalues, respectively. The conjugate gradient method for solving $Ax = b$ converges after at most n steps. Moreover, the error $e^{(k)}$ at the $k-$th iteration (with $k < n$) is orthogonal to $p^{(j)}$, for $j = 0, ..., k-1$, and*

$$\|e^{(k)}\|_A \leq \frac{2c^k}{1+c^{2k}}\|e^{(0)}\|_A \quad , \textit{ with } \quad c = \frac{\sqrt{K_2(A)}-1}{\sqrt{K_2(A)}+1} \quad \textit{and} \quad K_2(A) = \frac{\lambda_1}{\lambda_n} \, .$$

Proof See [1], p. 154.

In the CG method, unlike the gradient method, the convergence is influenced by the *whole* spectrum of A and not only by its extreme eigenvalues. Indeed, the CG method typically converges in three phases: an initial phase of rapid convergence but short duration, which depends essentially on the initial error; a fairly linearly convergent phase, which depends on the condition number; and finally a superlinearly convergent phase, which depends on how the smallest eigenvalues are distributed [3].

The termination property of the CG method stated in theorem 1.2.2 is rigorously valid only in exact arithmetic. For matrices A of large size, the iterations are stopped when the error falls below a fixed tolerance. The accumulating rounding errors prevent the descent directions from being A-conjugate and can even generate null denominators in the computation of coefficients α_k and β_k. The latter phenomenon, known as *breakdown*, can be avoided by introducing suitable stabilization procedures; in such an event, we employ stabilized gradient methods. Despite the use of these strategies, it may happen that the CG method fails to converge (in finite arithmetic). In such a case, the only reasonable possibility is to restart the iterative process, taking as residual the last computed value. By so doing, the *cyclic CG method* or *CG method with restart* is obtained, for which, however, the convergence properties of the original CG method are no longer valid.

1.2.5 Preconditioning

If P is a symmetric and positive definite matrix, the preconditioned conjugate gradient (PCG) method consists of applying the CG method to the so–called preconditioned system

$$P^{-1/2}AP^{-1/2}y = P^{-1/2}b,$$

with $y = P^{1/2}x$ and P the so–called preconditioner.

In practice, the PCG method is implemented without explicitly requiring the computation of $P^{1/2}$ or $P^{-1/2}$. The complete algorithm is shown in section 1.3.

Particularly interesting is the Tony Chans preconditioner. The following lemma is based on Lemma 1 in [4] (see also [5] for a generalization and other convergence properties).

Lemma 1.2.3. *Let be defined*

$$\mathcal{M}_U \equiv \{\tilde{A} = U^*\Lambda U \mid \Lambda \text{ is any } n \times n \text{ diagonal orthogonal matrix}\}.$$

For any arbitrary $A \in \mathbb{R}^{n\times n}$, let $\tilde{A}$ be the minimizer of $\|\tilde{A} - A\|_F$ over all $\tilde{A} \in \mathcal{M}_U$. The Tony Chan's preconditioner P is uniquely determined by A and is given by

$$P = U^*\delta(UAU^*)U \ , \tag{1.2.5}$$

with $\delta(UAU^*)$ being the diagonal matrix whose diagonal is equal to the diagonal of UAU^*.

Proof. See [4]. □

It is worth noticing that we can use as U an orthogonal discrete wavelet transform matrix W (see next section) and verify that the resulting preconditioner P, as stated in (1.2.5), is symmetric and positive definite so as to use it directly in the PCG method.

1.2.6 Matrix compression using the wavelet transform

The wavelet transform is a widely used tool in signal processing and image analysis [7][8]. There are also interesting applications of wavelet transforms that can be used in linear algebra.

In particular, we are interested in the property that a linear system $Ax = b$ can become a sparse system in a suitable wavelet basis [9]. To this end, we first wavelet-transform A by

$$A_W = WAW^* \,, \tag{1.2.6}$$

and the right-hand side b by

$$b_W = Wb \,, \tag{1.2.7}$$

where W is an orthogonal discrete wavelet transform matrix (i.e., $W^{-1} = W^*$).
We then compress A_W with thresholding and solve the system

$$A_{W_c} x_{W_c} = b_W \,, \tag{1.2.8}$$

in which $A_{W_c} \simeq A_W$ is the compressed version of A_W. As the thresholding method, we can use, for instance, hard-thresholding [10], which consists of setting to zero entries with values less than a given threshold.
Finally, we transform back the solution by the inverse wavelet transform, that is

$$x \approx W^* x_{W_c}. \tag{1.2.9}$$

1.3 Algorithms

1.3.1 Conjugate gradient method

Input

- an initial vector solution $x^{(0)}$;
- a symmetric and positive definite matrix A;
- a vector b;
- a tolerance ε, i.e., a positive number greater than the machine epsilon;
- a maximum number of iterations $k_{max} \geq 1$.

Algorithm

- set $k = 0$;
- compute $r^{(0)} = b - Ax^{(0)}$;
- set $p^{(0)} = r^{(0)}$;
- while $\big(k <= k_{max}$ or $\|r^{(k)}\| >= \varepsilon\ \big)$;

  {

  - compute $\alpha_k = \dfrac{p^{(k)^T} r^{(k)}}{p^{(k)^*} A p^{(k)}}$;
  - compute $x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}$;
  - compute $r^{(k+1)} = r^{(k)} - \alpha_k A p^{(k)}$;

– compute $\beta_k = \dfrac{(Ap^{(k)})^T r^{(k+1)}}{(Ap^{(k)})^T p^{(k)}}$;

– compute $p^{(k+1)} = r^{(k+1)} - \beta_k p^{(k)}$;

– increment k.

}

1.3.2 Preconditioned conjugate gradient method

Input

- an initial vector solution $x^{(0)}$;
- a square symmetric and positive definite matrix A_n;
- a vector b_n;
- a symmetric and positive definite preconditioner P;
- a tolerance ε, i.e., a positive number greater than the machine epsilon;
- a maximum number of iterations $k_{max} \geq 1$.

Algorithm

- set k=0;
- compute $r^{(0)} = b - Ax^{(0)}$;
- compute $z^{(0)} = P^{-1} r^{(0)}$;
- set $p^{(0)} = z^{(0)}$;

- while $\left(k <= k_{max} \text{ or } \|r^{(k)}\| >= \varepsilon\right)$

  {

  - compute $\alpha_k = \dfrac{z^{(k)^T} r^{(k)}}{p^{(k)^T} A p^{(k)}}$;

  - compute $x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}$;

  - compute $r^{(k+1)} = r^{(k)} - \alpha_k A p^{(k)}$;

  - compute $z^{(k+1)}$, solving the linear system $P z^{(k+1)} = r^{(k+1)}$;

  - compute $\beta_k = \dfrac{z^{(k+1)^T} r^{(k+1)}}{z^{(k)^T} r^{(k)}}$;

  - compute $p^{(k+1)} = z^{(k+1)} - \beta_k p^{(k)}$;

  - increment k;

  }

1.3.3 Discrete wavelet transform

Input

- a set of wavelet coefficients [8];
- a positive number l representing the so–called multi-resolution levels [8];
- a square matrix A or a vector b ;

Algorithm

- compute A_W or b_W, that is the discrete wavelet transform (DWT) of A or b, using the so called *pyramidal algorithm* [8], which uses the Fast Fourier Transform.

1.4 Applications

1.4.1 Neuroelectric source localization from magnetoencephalography data: a preliminary case study

In this section, we use the algorithms of section 1.3 to efficiently obtain the solution of a neuroimaging inverse problem, i.e. the localization of neuronal sources responsible for brain activity from magnetoencephalography (MEG) data [11].

By means of special sensors called single superconducting quantum interference devices (SQUIDs), which are positioned near the head of a subject, MEG systems measure the magnetic field generated outside the skull by the synchronous activation of thousands of neurons. MEG complements other brain activity measurement techniques such as electroencephalography (EEG), positron emission tomography (PET), and functional magnetic resonance (fMRI). Its strengths consist of independence of head geometry and better spatial resolution compared to EEG and non-invasiveness as opposed to PET. Many studies have reported successful diagnosis of cases of multiple sclerosis, Alzheimer's disease and other pathologies using the MEG perspective in combination with other methods. The goal is to support doctors in making an early diagnosis, i.e., in distinguishing patients from healthy control subjects. This explains why an increasing number of hospitals are currently using MEG in combination with EEG, fMRI or PET.

The localization of the neuronal sources underlying the measured magnetic field is a highly ill-posed and ill-conditioned inverse problem that requires sophisticated numerical methods to be solved [12] [13] [14].

The mathematical problem

From now on, we shall use the following relevant assumptions: only the radial component of the magnetic field is recorded by our MEG measurement system; the magnetic measurements made close to the scalp are due to currents in the entire gray matter of the brain, even if they occur mainly whitin the cerebral cortex; the brain is a spherical symmetric homogeneous conductor G of radius R; and the corresponding mathematical formulation of the forward problem and its discretization are the same as in [15].

Let e_ρ, e_θ, e_φ be the unit vectors associated with the spherical coordinates ρ, θ, φ, respectively, with $0 \leq \rho \leq R$, $0 \leq \theta \leq \pi$, $0 \leq \varphi \leq 2\pi$.

The radial component of the magnetic field at point r is

$$B_\rho(r) = B(r) \cdot e_\rho(r) = \frac{\mu_0}{4\pi} \int_G \frac{J(r') \times (r - r') \cdot e_\rho(r)}{\|r - r'\|^3} \, dv' \ , \tag{1.4.1}$$

with $\mu_0 = 4\pi 10^{-7}$ being the magnetic permeability of the free space.

$B_\rho(r)$ can be evaluated exactly only for the particular current distribution $J(r)$. For instance, if currents are represented by a single dipole then the magnetic field is

$$B_\rho(r) = \frac{\mu_0}{4\pi} \frac{Q \times (r - r_Q) \cdot e_\rho(r)}{\|r - r_Q\|^3} \, , \tag{1.4.2}$$

with r_Q the dipole position and Q the dipole moment. In the more general case, the integrals (1.4.1) have to be approximated.

Let

$$J(r) = J_\rho(\rho, \theta, \varphi) e_\rho + J_\theta(\rho, \theta, \varphi) e_\theta + J_\varphi(\rho, \theta, \varphi) e_\varphi \tag{1.4.3}$$

be the vector current distribution. Then, the radial magnetic field is determined by

only J_θ and J_φ, that is

$$B_\rho(r) = \frac{\mu_0}{4\pi} \int_G \left[J_\theta \left(\frac{y\cos\varphi - x\sin\varphi}{a^{3/2}} \right) - J_\varphi \left(\frac{x\cos\varphi\cos\theta + y\sin\varphi\cos\theta - z\sin\theta}{a^{3/2}} \right) \right] dv' \ , \tag{1.4.4}$$

with $a = (x - \rho\sin\theta\cos\varphi)^2 + (y - \rho\sin\theta\sin\varphi)^2 + (z - \rho\cos\theta)^2$.

From equation (1.4.4) evaluated at the magnetometer sites r_i $(i = 1, ..., M)$ the so–called forward (discretized linear) model is derived by means of a quadrature method [15], that is

$$Tj = [T_\theta \ T_\varphi][j_\theta \ j_\varphi]^* \, , \tag{1.4.5}$$

where j is the discretized current density vector and j_θ and j_φ are the discretizations of J_θ and J_φ, respectively.

Now let x be the unknown current density vector and d a given MEG data vector; the linear system

$$A\,x = b = T^* d \tag{1.4.6}$$

represents a well–defined linear inverse problem, with T^* the transpose of T and $A = T^*T$ a symmetric positive definite matrix.

Remark 1.4.1. The linearity of the problem is a consequence of the relevant assumptions regarding the neuronal source and volume conductor representations.

Even if matrix A is an almost-full matrix, a sparse representation of A can be obtained by applying a discrete wavelet transform (DWT) [8]. The DWT of A is given by

$$A_W = T_W^* T_W \, , \tag{1.4.7}$$

being T_W the DWT of T and T_W^* the DWT of T^* obtained using the pyramidal algorithm [8].

Then, T_W is compressed by hard-thresholding [10], i.e., by setting to zero the entries under a given threshold. Thus, the linear system to be solved is

$$A_{W_c} x_{W_c} = b_W \,, \tag{1.4.8}$$

with $A_{W_c} = T_{W_c}^* T_{W_c}$, T_{W_c} the compressed version of T_W and b_W the DWT of b obtained with the aforementioned pyramidal algorithm.

Because the inverse problem (1.4.6) is to reconstruct the current density vector x once the data d are given, the recovered current density in the physical space can be obtained by the inverse wavelet transform of $x_{W_c}^{(\nu)}$, that is, the approximation in wavelet coordinates obtained after ν iterations of the CG method.

Numerical tests

To test the CG method, we solve a particular linear inverse problem, i.e. the localization of a current dipole in a homogeneous sphere G of radius $R = 10$ cm from MEG data.

The operator (1.4.4) was approximated by a trapezoidal quadrature rule [15]. The resulting matrix T was an almost-full matrix with dimensions 1024×65536.

Two sets of synthetic data were generated, one containing noiseless data and the other containing noisy data obtained by adding white Gaussian noise with linear signal-to-noise ratio equal to 1, sampling equation (1.4.2) in 1024 sites distributed on a hemisphere Σ concentric to the sphere G with $dist(\Sigma, \partial G) = 1$ cm, and supposing a current dipole located in $P_D = (x_D, y_D, z_D) = (3, 5, 8)$ cm.

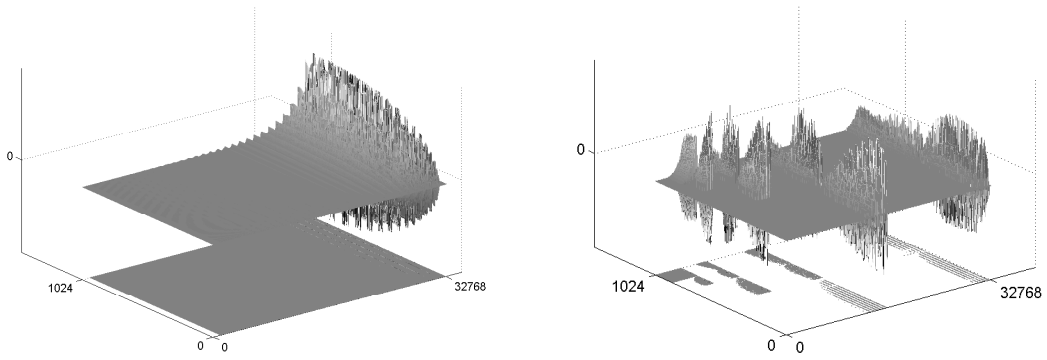
To transform T, we use a DWT having the following wavelet coefficients

$$c_0 = (1+\sqrt{3})/4\sqrt{2}\ ,\ c_1 = (3+\sqrt{3})/4\sqrt{2}\,,$$

$$c_2 = (3-\sqrt{3})/4\sqrt{2}\ ,\ c_3 = (1-\sqrt{3})/4\sqrt{2}\,,$$

which correspond to the Daubechies orthogonal wavelet filters of length 4 (see [7]). The resulting matrix T_W was compressed by hard-thresholding, retaining only the entries over a given level-dependent threshold. In our tests, 3 multi-resolution levels were used, and we retained only the entries above 10% of the maximum of the wavelet coefficients at any level.

In Fig. 2.1, some entries of the matrix T (exactly T_θ, see Fig. 2.1a) and of the compressed matrix T_{W_c} (exactly $T_{\theta_{W_c}}$, see Fig. 2.1b) are displayed.



(a) (b)

Figure 2.1. Entries of T_θ (a) and $T_{\theta_{W_c}}$ (b). The projection on the horizontal plane shows that T_θ is a full matrix, while $T_{\theta_{W_c}}$ has just few nonzero entries.

The matrix T_{W_c} is a highly sparse matrix; the resulting compression factor was approximately 165. Obviously, compression allows us to save not only memory storage but also computation time.

As we stated at the beginning of this section, we performed a variety of tests to study the performance of CG with respect to MEG-ISTA using the soft thresholding (LWSOFT) method and Landweber iterations (LW)[15]. Each test was done for both noiseless and noisy data.

Computations were performed using Matlab R2009a on a computer with a 64-bit Windows Vista Ultimate OS, an Intel Core i7-920 (quad core, 2.81Ghz, 256Kb L2, 8Mb L3) and 12Gb ram (6x2 Gb, 1333 Mhz, DDR3).

The decrease in the residue (Figures 2.2a and 2.2b) and the resulting localization error (Table 2.1 and Figure 2.3), as a function of the number of iterations performed by each method show that the CG method performed better than the LW and LWSOFT methods for the specific fixed problem, even in the presence of noise. In fact, after only 10 iterations, the CG method gives the same results given by the LW and LWSOFT methods after approximately 50 iterations.

Method	Iterations ν	LocErr (noiseless data)	LocErr (noisy data)	Time (sec)
CG	10	2.15	2.15	0.76
LWSOFT	10	2.15	5.67	7.44
LW	10	2.15	12.98	7.26

Table 2.1. Localization errors (LocErr) in millimeters after 10 iterations of the CG, LWSOFT and LW methods.

1.5 Conclusions and future work

A magnetoencephalography linear inverse problem, i.e., a large ill-conditioned linear system, can be efficiently solved using the conjugate gradient method. The results obtained from a preliminary case study suggest that it would be worthwhile to invest more time and effort in this direction, introducing suitable preconditioners and testing the spectral properties of the resulting preconditioned linear systems. A Tony Chan's preconditioner [6] could be used in combination with the discrete wavelet transform. Taking into account the fact that the tuning of an optimal preconditioner is a difficult task [16], the definition of the linear inverse problem to be solved assumes a paramount relevance. For example, we could consider the using of realistic head models, such as those used in Brainstorm (an open source software application dedicated to magnetoencephalography and electroencephalography data analysis) and build special preconditioners for this problem. Furthermore, a Bayesian probabilistic approach, justified by partial knowledge of brain functioning, noisy measures and modeling errors, should also be specifically tested [14].

The final goal is the tuning of an efficient and effective algorithm that could be implemented and embedded in a magnetoencephalogram recording system. This real-time neuronal localization functionality could support researchers and clinicians during the monitoring of subjects' brain activity.

The Author intends also to investigate how to reuse this work for the analysis of hospitalization data sets resulting from the standard information system of a hospital. In fact, another relevant goal would be the embedding of some yet to be identified data analysis resulting from the use of such tools (i.e. fast solution of large linear systems) into a decision support system for the clinical governance of a hospital, such

as the Business Simulator for Health Care (BuS-4H) system, which the Author, who is the creator of the corresponding research project, wishes to develop in collaboration with an Italian company called “SiliconDev srl” and several clinical, mathematics and engineering departments of the University of Rome “Sapienza”, the University of Rome “Tor Vergata”, the “Fatebenefratelli” Association for Biomedical Research (AFaR) and the “San Giovanni Calibita Fatebenefratelli Isola Tiberina” Hospital in Rome.

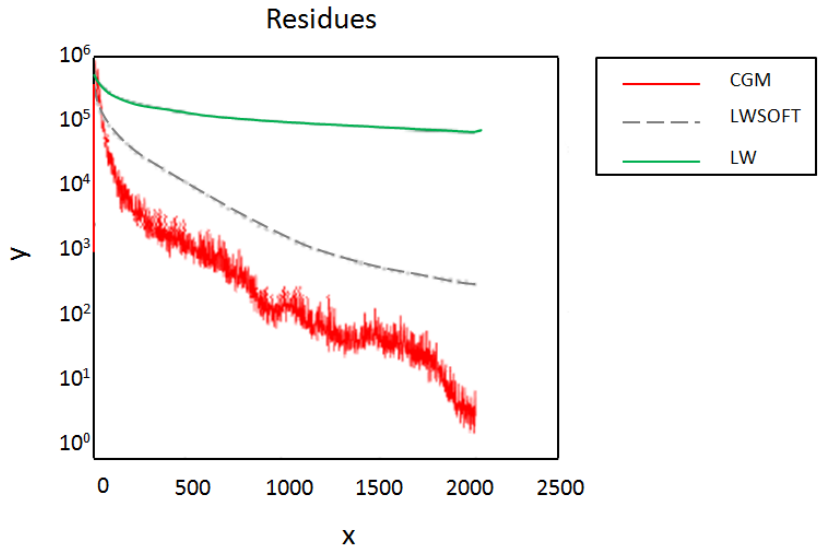


Figure 2.2a. The residue as a function of number of iterations of the CG, LW and LWSOFT methods, supposing a noiseless data vector d.

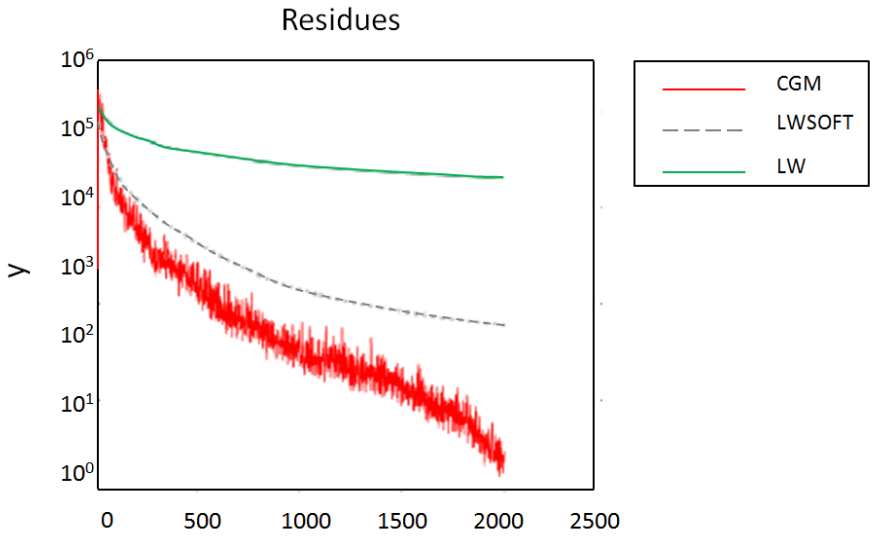


Figure 2.2b. Behavior of the residue as a function of number of iterations of the CG, LW and LWSOFT methods, supposing a noisy data vector d.

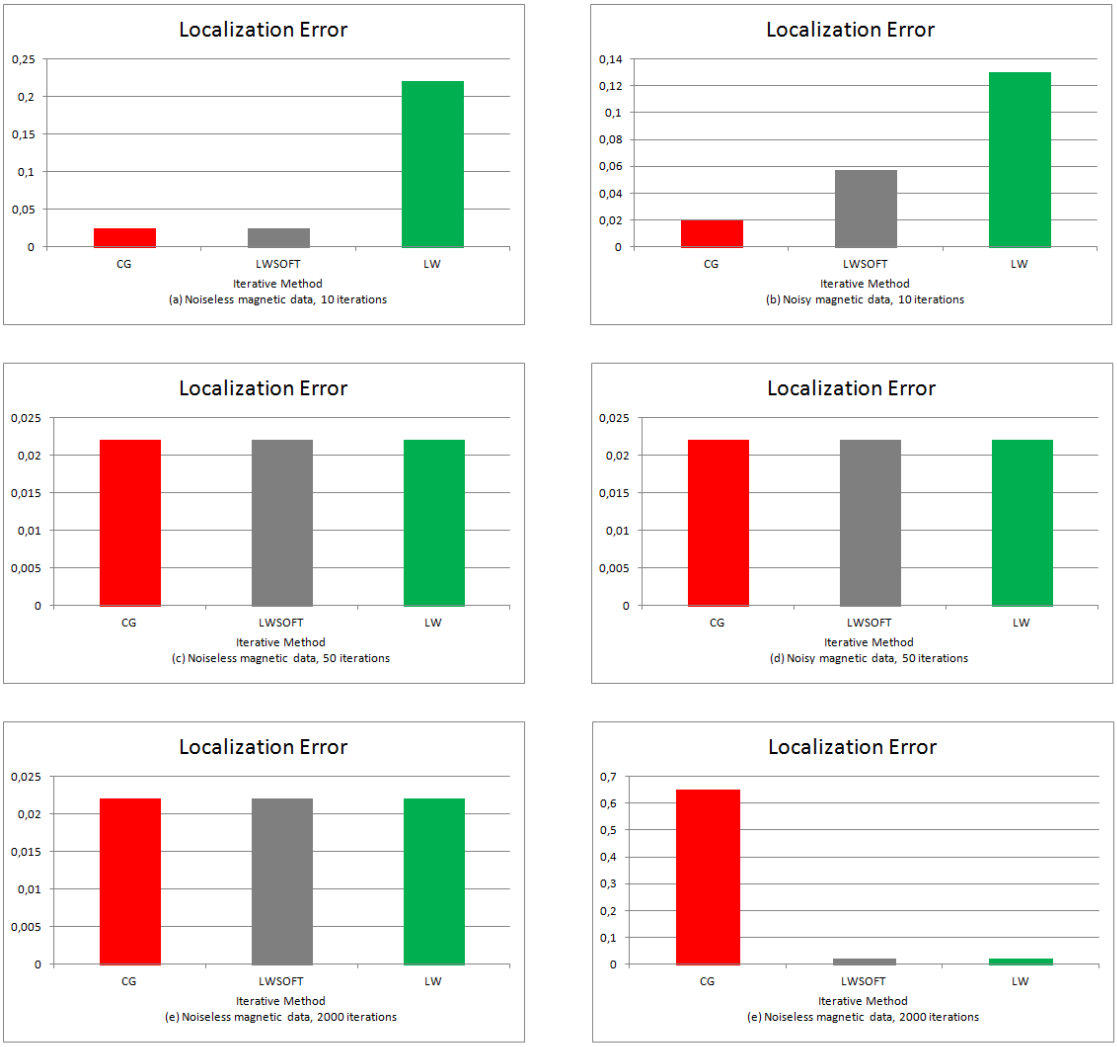


Figure 2.3. The localization error after exactly 10, 50 and 2000 iterations of the CG, LW and LWSOFT methods. The effects of Gaussian noise are shown on the right, and the noiseless condition is shown on the left.

Bibliography

[1] Quarteroni A, Sacco R, Saleri F, *Numerical Mathematics*, Springer (2000)

[2] Golub GH, Van Loan CF, *Matrix Computations*, The Johns Hopkins University Press, Baltimore (1983).

[3] Axelsson O, *Iteration number for the conjugate gradient method*, Mathematics and Computers in Simulation *61(3-6)* (2003).

[4] Jin XQ, Wei YM, Xu W, *A Stability Property Of T. Chan's Preconditioner*, SIAM J. Matrix Anal. Appl. 25(3) (2003).

[5] Serra Capizzano S, *A Korovkin-type Theory for finite Toeplitz operators via matrix algebras*, Numer. Math. 82(1) (1999).

[6] Chan TF, *An optimal circulant preconditioner for Toeplitz systems*, SIAM Journal on Scientific and Statistical Computing *9(4)* (1988).

[7] Daubechies I, *Ten Lectures on Wavelets*, Philadelphia, PA: SIAM (1992).

[8] Strang G, Nguyen T, *Wavelets and Filter Banks*, Wellesley, MA: Wellesley-Cambridge Press (1996).

[9] Dahmen W, Harbrecht H, Schneider R, *Compression techniques for boundary integral equations-asymptotically optimal complexity estimates*, SIAM J. Numer. Anal., *43(6)*, (2006).

[10] Donoho DL, *De-Noising by Soft Thresholding*, IEEE Trans. Info. Theory, *43* (1993).

[11] Del Gratta C, Pizzella V, Tecchio F, Romani GL, *Magnetoencephalography - A Noninvasive Brain Imaging Method with 1ms Time Resolution*, Rep. Prog. Phys. *64* (2001).

[12] Bertero M, Boccacci P, *Introduction to Inverse Problems in Imaging*, Philadelphia, PA: Institute of Physics (2002).

[13] Engl HW, Hanke M, Neubauer A, *Regularization of Inverse Problems*, Dordrecht, Netherlands: Kluwer (2000).

[14] Calvetti D, Hakula H, Pursiainen S, Somersalo E, *Conditional gaussian hypermodels for cerebral source localization*, SIAM Journal On Imaging Sciences , *2(3)* (2009).

[15] Filardo C, *A fast iterative algorithm to solve the magnetoencephalography inverse problem*, University of Rome "La Sapienza" (2012).

[16] Chen K, *Matrix Preconditioning Techniques and Application*, Cambridge Monographs on Applied and Computational Mathematics (2005).

[*This page was intentionally left blank*].

Chapter 2

Measuring quality of services using coherent conditional probability

2.1 Introduction

The concept of quality, embraced during the second half of the last century, has undergone an exponential spread in many areas. The need to ensure quality standards in industry (aviation, military, manufacturing, and other areas) led to a continuous search for quality-assurance methods that can satisfy the needs of quite different fields. This search included the health care world, which for years has been engaged in a search for useful systems for measuring and improving performance from the point of view of both quality and quantity.

Thus, the aim of this chapter is to realize, by focusing on some relevant aspects, a suitable model for an effective evaluation of health care providers such as a hospitals or networks of hospitals.

To evaluate the quality we need to refer to some linguistic notions such as "very efficient", which is a fuzzy concept. Fuzzy sets have been introduced to manage linguistic and vague information [1]. In this chapter, we manage vague information

in probabilistic terms by using a probabilistic reinterpretation of fuzzy sets [2]. Some applications of this approach already exist [3], even in the health care context [4].

The target application at the end of this chapter is a first step towards an effective innovation goal, which is the realization of a quality measure system for an effective evaluation of all vaccination centers in the Lazio region of Italy.

To realize this work, the Author worked at the Fatebenefratelli Association for Biomedical Research (AFaR), University of Rome "Sapienza" and University of Rome "Tor Vergata", in collaboration with, as listed in a related publication under review and among others, Massimo Maurici, Alessandra Campolongo, Alessandro Giordani, Maurizio Ferrante and Elisabetta Franco.

2.2 Background

2.2.1 Coherent conditional probability

Our approach to probability [5] [6] [7] is based on *coherence* (a concept that goes back to Bruno de Finetti [8]).

A peculiarity (which entails a large flexibility in the management of any kind of uncertainty) of this approach to conditional probability is that, due to its *direct* assignment as a whole, the knowledge (or the assessment) of the "joint" and "marginal" unconditional probabilities $P(E \wedge H)$ and $P(H)$ is not required; moreover, the *conditioning* event H (which *must* be a *possible* event) may have *zero probability*.

A conditional probability $P(\cdot|\cdot)$ is defined on a set of $\mathcal{G} \times \mathcal{B}^o$ conditional events $E|H$ such that $\mathcal{G}$ is a Boolean algebra and $\mathcal{B} \subseteq \mathcal{G}$ is closed with respect to (finite) logical sums. More in details, it holds the following definition:

Definition 2.2.1. A conditional probability $P : \mathcal{C} \to [0, 1]$ is such that

(i) $P(H|H) = 1$, for every $H \in \mathcal{B}^o = \mathcal{B} \setminus \{\emptyset\}$,

(ii) $P(\cdot|H)$ is a (finitely additive) probability on $\mathcal{G}$ for any given $H \in \mathcal{B}^o$,

(iii) $P\big(E \wedge A|H\big) = P(E|H) \cdot P\big(A|E \wedge H\big)$, for every E, $A \in \mathcal{G}$ and E, $E \wedge H \in \mathcal{B}^o$,

being $\emptyset$ the impossible event.

It is possible, through the concept of *coherence*, to handle those situations (that come to the fore in real applications) where we need to assess P on an *arbitrary* set of conditional events $\mathcal{C} = \{E_1|H_1, \ldots, E_n|H_n\} \subseteq \mathcal{G} \times \mathcal{B}^o$. Thus the role of coherence is that of ruling an extension process, starting from the classic *axioms for a conditional probability*.

Definition 2.2.2. The assessment $P(\cdot|\cdot)$ on $\mathcal{C}$ is *coherent* if there exists $\mathcal{C}' \supset \mathcal{C}$, with $\mathcal{C}' = \mathcal{G} \times \mathcal{B}^o$ such that P can be extended from $\mathcal{C}$ to $\mathcal{C}'$ as a *conditional probability*.

Concerning coherence, a fundamental result is the following [8].

Theorem 2.2.1. *Let $\mathcal{K}$ be any family of conditional events, and take an* arbitrary *family $\mathcal{C} \subseteq \mathcal{K}$. Let P be an assessment on $\mathcal{C}$; then there exists a (possibly not unique) coherent extension of P to $\mathcal{K}$* if and only if *P is coherent on $\mathcal{C}$.*

Coherence of conditional assessments can be ruled by a fundamental *characterization theorem* [8], which is based on checking the compatibility of a suitable sequence of linear systems.

From this characterization theorem the following results derive [8].

Proposition 2.2.2. *Let $\mathcal{C} = \{E|H_i\}_{i=1,...,n}$ be such that the events H_i's are a partition of Ω. Then any function $P : \mathcal{C} \to [0,1]$ such that*

$$P(E|H_i) = 0 \ \text{ if } \ E \wedge H_i = \emptyset \ \text{ and } \ P(E|H_i) = 1 \ \text{ if } \ H_i \subseteq E$$

is a coherent conditional probability.

Moreover if the only coherent conditional probability takes values in $\{0,1\}$, then it is $H_i \wedge E = \emptyset$ for every H_i, such that $P(E|H_i) = 0$ and it is $H_i \subseteq E$ for every H_i such that $P(E|H_i) = 1$.

Proposition 2.2.3. *Let $\mathcal{C} = \{E|H_i\}_{i=1,...,n}$ be a set of conditional events such that $\mathcal{H}^o = \{H_1, ..., H_n\}$ is a partition of Ω and let $P(\cdot)$ be a probability distribution on $\mathcal{H}^o$. Then, for every function $p : \mathcal{C} \to [0,1]$ such that $p(E|H_i) = 0$ if $E \wedge H_i = \emptyset$ and $p(E|H_i) = 1$ if $H_i \subseteq E$, the global assessment*

$$\{p(E|H_i), P(H_i)\}_{i=1,...,n}$$

is a coherent conditional probability assessment.

By Proposition 2.2.3, when a probability P is given on $\{H_1, ..., H_n\}$, this probability is coherent also together with any likelihood function $p(E|\cdot)$ on the same events H_i. Moreover its extension, by the total probability rule, is uniquely defined on all the events of $\mathcal{H}$, the algebra generated by $\mathcal{C} = \{E|H_i\}_{i=1,...,n}$, with positive probability .

The global coherence of P and p is also preserved when P is defined on a subset $\mathcal{D}$ of $\mathcal{H}$, as the next result shows.

Theorem 2.2.4. *Let p be a coherent conditional probability assessment on $\mathcal{C} = \{E|H_i\}_{i=1,...,n}$ with $\mathcal{H}^o = \{H_1, ..., H_n\}$ a partition of Ω and let $\mathcal{H}$ be the algebra generated by $\mathcal{C}$. Consider a coherent probability P on a set $\mathcal{D} \subseteq \mathcal{H}$, then the assessment $\{p, P\}$ is globally coherent.*

Theorem 2.2.4 implies that the check of global coherence of $\{p(E|\cdot), P(\cdot)\}$, when $\mathcal{D} \subseteq \mathcal{H}$, simply reduces, from a computational point of view, to the check of coherence of P on $\mathcal{D}$. In fact, by proposition 2.2.2, the check of coherence of p has linear complexity. We recall that indeed the problem of checking coherence of P is NP-complete [10].

Remark 2.2.1. We need to stress the relevance of condition $\mathcal{D} \subseteq \mathcal{H}$. In fact, for example, when $\mathcal{H}$ is included in the algebra generated by $\mathcal{D}$, even if the events of $\mathcal{D}$ are mutually exclusive, the coherence of the assessments P and p do not imply the global coherence of $\{p(E|\cdot), P(\cdot)\}$ (for more details see example 3 in [3]).

Concerning the coherence of a set of a likelihood functions the following result holds [8].

Theorem 2.2.5. *Let $\mathcal{C} = \{E_j|H_{j_i}\}_{i=1,...,n_j\,;j=1,...,n}$ be a set of conditional events such that the set $\mathcal{H}_j^o = \{H_{j_1}, ..., H_{j_{n_j}}\}$ is a partition of Ω, for every j, and the events of set $\mathcal{E} = \{E_j\}_{j=1,...,n}$ are logically independent. For every j, let $P_j : \mathcal{H}_j \mapsto [0,1]$ be a probability distribution and $p(E_j|\cdot) : \mathcal{H}_j^o \mapsto [0,1]$ a coherent conditional probability. If the probability distributions P_j's are globally coherent on $\mathcal{H}^* = \bigcup_j \mathcal{H}_j^o$, then the assessment $\{P_j, p(E_j|\cdot)\}_{j=1,...,n}$ is globally coherent in $\mathcal{E} \times \mathcal{H}^*$.*

Remark 2.2.2. Note that the proposition 2.2.3 is a particular case of theorem 2.2.5.

2.2.2 Fuzzy sets and conditional probability

We refer to the state of information (at a given moment) of a real (or fictitious) person that will be denoted by "You". If X is a (not necessarily numerical) quantity with range C_X, let H_x be, for any $x \in C_X$, the event $\{$"$X = x''\}$. The family $\{H_x\}_{x \in C_X}$ is obviously a *partition* of the certain event $\Omega = C_X$. Now, let φ be any *property* related to the quantity X and E_φ="You claim that X has the property φ" be an event of interest. E_φ is also called a *a fuzzy event* [1].

Remark 2.2.3. For instance let E_φ="A doctor claims that a service is efficient" be an event related to the efficiency (i.e. φ) evaluation of a health care provider. Suppose that the proposition which defines the event E_φ is logically equal to the following: "A doctor claims that the duration time of service is efficient". In other words, in this example we are modeling the efficiency of service using the key performance indicator X:=duration time of service (minutes). The first proposition is more concise than the

second one whether You have many indicators which characterized the efficiency of a service. Continuing the example, suppose also that C_X is partitioned in two ranges: $X \leq 10$, $X > 10$ (minutes). With an abuse of notation, we can use the symbol X (or think to it) as an identifier of each range. So, in this respect, $X = 1$ corresponds to the range $X \leq 10$ and $X = 2$ corresponds to the range $X > 10$. The family $\{H_1, H_2\}$ is a partition of the certain event Ω, with H_1="The duration time X is less or equal to 10 minutes" and H_2="The duration time X is greater than 10 minutes" being the performance that can be observed during the service providing.

From a pragmatic point of view, it is natural to think that You have some information about possible values of X, which allows You to refer to a suitable *membership function* $\mu_\varphi(x)$ of the *fuzzy subset* E_φ of "elements of C_X with the property φ".

For example, if X is a numerical quantity and φ is the property "small", for You the membership function $\mu_\varphi(x)$ may be put equal to 1 for values x of X less than a given x_1, while it is put equal to 0 for values greater than x_2; then it is taken as decreasing from 1 to 0 in the interval from x_1 to x_2. This choice of the membership function implies the following two facts: 1) You are certain that elements of C_X less than x_1 have the property φ, those greater than x_2 do not; 2) You are uncertain that on having or not the property φ those elements of C_X between x_1 and x_2. Then the interest is in fact directed toward *conditional events* such as $\{E_\varphi|H_x\}_x$, where x ranges over the interval from x_1 to x_2. It follows that, while You may assign to each of these conditional events a degree of belief (subjective probability) $\{P(E_\varphi|H_x)\}_x$, You must not assign a degree of belief $1 - P(E_\varphi|H_x)$ to the event E_φ under the assumption H_x^c (the value of X is not x), since an additivity rule *with respect to the conditioning events* does not hold, i.e. $1 - P(E_\varphi|H_x) \neq P(E_\varphi|H_x^c)$. In other

words, it seems sensible to identify the values of the membership function $\mu_\varphi(x)$ with suitable conditional probabilities. In particular, putting H_F="X is greater than x_2" and H_T="X is less than x_1", one has that E_φ and H_F are incompatible and that H_T implies E_φ, so that, by the rules of a conditional probability, $P(E_\varphi|H_F) = 0$ and $P(E_\varphi|H_T) = 1$.

Remark 2.2.4. It is worth noting that this conditional probability $\left\{P(E_\varphi|H_x)\right\}_x$ is *directly* introduced as a function on the set of conditional events $\left\{E_\varphi|H_x\right\}_x$ (and without assuming any given algebraic structure). Is that possible? In the usual (Kolmorogovian) approach to conditional probability it is not possible since the introduction of $P(E_\varphi|H_x)$ would require the consideration (and the assessment) of $P(E_\varphi \wedge H_x)$ and $P(H_x)$ (assuming positivity of the latter), while it is possible in a coherent conditional probability setting [7][11].

2.2.3 Operations between fuzzy sets

Given two fuzzy subsets E_φ, E_ψ related to the same variable X, with the events E_φ, E_ψ logically independent with respect to X, it has been proved that the coherent values for $\mu_{\varphi\cap\psi}(x) = P(E_\varphi \wedge E_\psi|x)$ and $\mu_{\varphi\cup\psi}(x) = P(E_\varphi \vee E_\psi|x)$, for any given x in the range of X, can be obtained by Frank t-norms and their dual t-conorms [3][7][11]. The case of two fuzzy subsets E^*_φ, E^*_ψ, related to two distinct random quantities X_1 and X_2, respectively, has been studied [7] by assuming the following conditional independence condition: for every (x, x') belonging to the range $\mathcal{C}_{(X_1,X_2)}$ of the vector (X_1, X_2)

$$P(E_\varphi|x \wedge x') = P(E_\varphi|x) \quad , \quad P(E_\psi|x \wedge x') = P(E_\psi|x')\,, \tag{2.2.1}$$

being $x \wedge x'$ the event "$X = x$"$\wedge$"$X = x'$ ".

Now we note that the requirement of conditional independence (2.2.1) can be removed. By assuming that E_φ and E_ψ are logically independent we have

$$P(E_\varphi \vee E_\psi | x \wedge x') = P(E_\varphi | x \wedge x') + P(E_\psi | x \wedge x') - P(E_\varphi \wedge E_\psi | x \wedge x')\,, \quad (2.2.2)$$

and among the coherent values for $P(E_\varphi|x\wedge x')$ and $P(E_\psi|x\wedge x')$ there are those given by the assumption (2.2.1): conditional independence assumption in fact only restricts the set of coherent values [12]. Then, in this general case we can have coherent values for $E_\varphi \wedge E_\psi | x \wedge x'$ and $E_\varphi \vee E_\psi | x \wedge x'$ also outside the previous intervals, but all the values in those intervals are coherent.

Note that by using this interpretation of membership functions, all the results of the previous section can be used for making inference, when we take into account fuzzy and probabilistic information. This approach has been used for evaluating the quality of a health care provider [4] or to model some physique index to deal with an AVATAR application [3].

2.2.4 Similarity

We recall now the definition of *similarity* (also called "resemblance" [13]).

Given two fuzzy subsets $E_\varphi,\ E_\psi$ related to the same variable X, let $\mathcal{F}(C_X)$ be the family of fuzzy subsets of C_X. A *similarity* S is a mapping

$$S : \mathcal{F}(C_X) \times \mathcal{F}(C_X) \longrightarrow [0, 1]$$

such that

1. (Symmetry) $\quad S(E_\varphi, E_\psi) = S(E_\psi, E_\varphi)$;

2. (Reflexivity) $S(E_\varphi, E_\varphi) = 1$.

In the interpretation of fuzzy sets used here the concept of similarity can be reread as follow [9].

Proposition 2.2.6. *Let E_φ, E_ψ be fuzzy subsets of $\mathcal{C}_X$, with*

$$\mu_\varphi(\cdot) = P(E_\varphi|\cdot) \ , \ \mu_\psi(\cdot) = P(E_\psi|\cdot)\,,$$

and let $P(E_\varphi \wedge E_\psi|H_x)$ be a relevant coherent assessment. Then any coherent extension of $P(\cdot|\cdot)$ to the conditional event $E_\varphi \wedge E_\psi|E_\varphi \vee E_\psi$ is a similarity.

The existence of such a function is warranted by the fundamental extension Theorem recalled in section 2.2.1. The semantic behind this choice is the following: **the more** two fuzzy subsets are considered to be similar, **the more** if You claim *at least one* of the two corresponding properties You are willing to claim *both* properties.

How to compute $S(E_\varphi, E_\psi)$? Given $\mu_\varphi(\cdot) = P(E_\varphi|\cdot)$ and $\mu_\psi(\cdot) = P(E_\psi|\cdot)$, the membership functions $\mu_{\varphi\cup\psi}(\cdot)$ and $\mu_{\varphi\cap\psi}(\cdot)$ of the fuzzy sets $(E_\psi \cup E_\varphi)$ and $(E_\psi \cap E_\varphi)$ (corresponding to a T–conorm and a dual T–norm [11]) arise as coherent extensions of the assessment P given on $\{E_\psi|H_x\,,\ E_\varphi|H_x\ :\ H_x$="X=x", $x \in C_X\}$ with E_φ and E_ψ logically independents with respect to X.

Then, given a conditional probability $P(\cdot|\cdot)$ on $\mathcal{H}_X \times \mathcal{H}_X^o$ (which gives rise to a class $\{P_\alpha\}$ of coherent unconditional probabilities), we have (for simplicity we refer to a finite $\mathcal{C}_X$)

$$S(E_\varphi, E_\psi) = \frac{\sum_x \mu_{\varphi\cap\psi}(x)\lambda_\alpha(x)}{\sum_x \mu_{\varphi\cup\psi}(x)\lambda_\alpha(x)} \tag{1}$$

where $\lambda_\alpha(x) = P_\alpha(H_x)$, with α the so called zero-layer [7] of the event $E_\psi \vee E_\varphi$.

Notice that, contrary to what happens in the classic fuzzy framework, this approach to similarity is able to take into account possible different "weights" of the values x through the probability values $\lambda_\alpha(x)$.

Some classic similarity functions (the most used in applications and proposed in the relevant literature) are related to the above formula involving conditional probability [9].

2.2.5 Inference

Let E_φ ="You claim that the variable X has the property φ" be an event of interest. So, when $\mu_\varphi(\cdot)$ and $P(\cdot)$ on $\mathcal{C}_X$ are known or assessed, according to the rules of conditional probability (in particular, the theorem of total probability), we can easily compute the probability of E_φ as

$$P(E_\varphi) = \sum_x P(X = x)P(E_\varphi|x) = \sum_x P(X = x)\mu_\varphi(x) \ , \tag{2.2.3}$$

which coincides with Zadeh's definition of the probability of a "fuzzy event" [15].

Now our aim is to choose the most probable element of $\mathcal{C}_X$ by using both statistical and fuzzy information. More precisely, if we have a probability distribution on $\mathcal{C}_X$ and a fuzzy information expressed by a membership function $\mu_\varphi(\cdot) = P(E_\varphi|\cdot)$, we can choose the most probable element $x \in \mathcal{C}_X$ under the hypothesis E_φ.

By using Bayes theorem actually we can easily compute, for every $x \in \mathcal{C}_X$, the value $P(X = x|E_\varphi)$ as

$$P(X = x|E_\varphi) = \alpha P(X = x)\mu_\varphi(x) \ ,$$

where $\alpha = \left(\sum_x \mu_\varphi(x)P(X = x)\right)^{-1}$.

So, to reach our goal it is sufficient to find the events x with maximum posterior, i.e.

$$P(x^*|E_\varphi) = \alpha \max_x \big\{P(X = x)\mu_\varphi(x)\big\}. \tag{2.2.4}$$

Remark 2.2.5. It is important to notice that here the Bayesian procedure is applied in an unusual semantic way. In fact, the distribution, which plays the role of "prior" probability, is here usually obtained by statistical data, whereas the membership function, which plays the role of "likelihood" is a subjective evaluation.

2.3 Algorithms

2.3.1 Quality evaluation of a health care provider

Let **p**, φ, **s**, **k**, **x**, **t** be indexes used to univocally identify some relevant concepts, i.e., a provider is denoted by the index **p**, while φ, **s**, **k**, **x**, and **t** represent respectively a property (i.e. a quality aspect), a service, a stakeholder of a service, a performance condition and a period of observation.
Assuming one wishes evaluate the provider **p** over the period **t**, an algorithm to measure the quality is:

Step 1

- For each performance, define an event H_x="The service is delivered with performance **x**".
- For each stakeholder,
  - define an event M_k="Stakeholder **k** claims that service **s** has property φ",
  - assess probability $\mu_{M_k}(x) = P(M_k|H_x)$ with a survey, asking "With what probability do you claim that service **s** has property φ assuming that the service is delivered with performance **x** ?".
- Define an event E="An expert claims that service **s** has property φ".
- Compute probability of $E|H_x$, that is,

$$P(E|H_x) = \mu_E(x) = \sum_k \alpha_k \mu_{M_k}(x)\,, \tag{2.3.1}$$

where α_k is the (positive) relevance of each stakeholder **k** with respect to property φ such that $\sum_k \alpha_k = 1$

Step 2

- For each performance, compute probability $P(H_x)$ as the (relative) frequency of performance **x**, using the occurrence of **x** observed during the delivery of services **s**.

Step 3

- Compute the probability of E, that is, the service quality score

$$e_\varphi = P(E) = \sum_x P(E|H_x)P(H_x)\,. \tag{2.3.2}$$

- Compute the service quality for all services **s** and properties φ.

- The overall quality of service score can be evaluated as the expected value of the resulting quality of services scores e_φ, i.e.,

$$b_s = \sum_\varphi w_\varphi e_\varphi\,, \tag{2.3.3}$$

and the overall quality of provider score can be evaluated as the expected value of the overall quality of services scores b_s, i.e.

$$B = \sum_s g_s b_s\,, \tag{2.3.4}$$

where w_φ and g_s, respectively, are the (positive) relevance of each property φ and service **s**, such that $\sum_\varphi w_\varphi = 1$ and $\sum_s b_s = 1$.

2.4 Applications

2.4.1 Measuring and benchmarking the quality of vaccination services

Introduction

In past years many methods of analysis and improvement of quality have been used in health care, including increasingly models that give a value to the user's perspective [16] [17] [18] [19] [20] [21].

The Declaration of Alma-Ata (1978) establish the right and duty to participate individually and collectively in the planning and implementation of health care. Therefore the assessment of quality of care cannot leave the direct and active involvement of both those who participate in the delivery of health services and those who use it directly.

The lack of standardized indicators of quality and the desire to overcome the partiality of the doctor's "absolutist" point of view, or the patient's "individualistic" one, should stimulate the search for a new methodology. It seems necessary to look for an innovative approach aims to obtain a qualitative assessment of service quality by all those who are directly involved in the health care process and to deliver consistent and useful results for any comparison between different providers.

In this respect, and starting from the measurement needs of a hospital [4], the members of the "Measures of Quality in health care Services" (MQS)research project, which has involved since 2005 a multidisciplinary team of researchers such as doctors, engineers, mathematicians of the Fatebenefratelli Association for Biomedical Research (AFaR), University of Rome "Sapienza" and University of Rome "Tor Vergata", have

tried to realize a model based on the probability theory according to Bruno de Finetti [4][8].

The aim of this study is the illustration of an application of the MQS approach to the evaluation of vaccination services in the Lazio Region of Italy [22].

Matherials and methods

Preliminaries

The "Quality in Vaccination: Theory and Research" (QuaVaTaR) group, a multidisciplinary team made up of medical doctors, engineers, epidemiologists and mathematicians, i.e. an extension of the MQS research group, was designated to conduct an application of the MQS approach to vaccination services in the Lazio Region of Italy.

Three among the twelve Local Health Units (ASL) within Lazio Region were involved, on the basis of the willingness to participate: ASL1 located in the center of Rome, ASL2 located on the edge of the city, and ASL3 located outside. Thirteen vaccination centers were selected: two centers in two of the four Districts of the ASL1, two centers in one of the five Districts of the ASL2 and nine centers in the four Districts of the ASL3.

The services under evaluation in the ASL2 and ASL3 were the pediatric vaccinations for children under three years of age, delivered with an appointment (VACP) and without appointment (VACL), while the anti-Human Papillomavirus vaccination (HPVV) services were evaluated in all the ASLs.

The service quality aspects under evaluation were the communicational efficiency (CE), the organizational efficiency (OE) and the comfort (CO). For each of them

two quality indicators were selected. CE was represented as the way and time of communicating information related to the vaccine with oral or written support, before or during vaccination. OE was represented using the time a user spent in a waiting room and the time a user spent for vaccination. CO was represented by means of the opening times of the services, the presence of toys and nursing room for children, and the availability of magazines for adolescents and adults.

In the following sections the reference model used by the QuaVaTar group is defined and then used to make the evaluation of the above ASLs.

The reference model

If a health–care provider runs several services (e.g.: supply of clinical tests, delivering of medical therapies, pediatric vaccinations, etc.), a quality aspect or property φ of service, for example the efficiency, can be defined and characterized by a certain number n of key performance indicators (e.g.: waiting time necessary to get the required clinical test, cost of treatment per outpatient episode, etc.): denote by I_j the j-th indicator, which is then a component of the vector $I = (I_1, I_2, \ldots, I_n)$. We consider all the possible realizations of the vector I (a subset of the cartesian product of the ranges of $I_1, I_2, \ldots, I_n$) identified by the index X (and C_X the set of all possible indexes x).

Using the interpretation of membership of a fuzzy set as a suitable likelihood [4], we ask all the clinical staff (e.g. doctors or nurses) to claim their (subjective) judgment on a given service. Considering the event $D =$ "A doctor claims that the service is efficient" and x ranging over C_X, we assign the membership function $\mu_D(x)$ by taking it equal to the probability of the conditional events $\{D|\text{“}X = x\text{”}\}$. The events "$X = x$" can be also denoted by H_x. The probability of $D|H_x$ could be

assessed, for example, through the following procedure: the doctors of the medical staff are required to evaluate the degree of quality of the service (given H_x) by a number between 0 and 1, so we get a vector of opinions $(o_x^{(1)}, o_x^{(2)}, \ldots, o_x^{(N)})$ (where N is the number of doctors); then, putting

$$o_x = \frac{o_x^{(1)} + o_x^{(2)} + \cdots + o_x^{(N)}}{N} \ , \tag{2.4.1}$$

we could assess $P(D|H_x) = o_x$.

The same procedure could be applied to non staff people, i.e. clients, of the health–care provider (i.e. patients, parents, etc.) to find the analogous membership function $\mu_C(x)$ as a conditional probability, by considering, now, the event $C =$ "A client claims efficient the service".

We can obtain a fuzzy model relative to doctors and clients (that represent together a sort of expert) of the provider as

$$\mu_E(x) = \alpha_D \mu_D(x) + (1 - \alpha_D)\mu_C \, ,$$

that is a coherent extension of the above conditional probability $P(\cdot|\cdot)$ assessed on the set $\left\{D|H_x \, , C|H_x\right\}_x$ to a "new" set of conditional events $\left\{G|H_x\right\}_x$, with G such that $D \wedge C \subseteq G \subseteq D \vee C$, α_D and $1 - \alpha_D$ respectively the relevance (> 0) of a doctor and a client.

Now, denoting by f_x the distribution, for the clients attending the service, of the realization x, we can measure the quality φ of service s assessing

$$e_\varphi = P(E) = \sum_x f_x \mu_E(x) \, . \tag{2.4.2}$$

Finally the overall quality of service score can be evaluated as the expected value of the resulting quality of services scores e_φ, i.e.

$$b_s = \sum_\varphi w_\varphi e_\varphi \,, \tag{2.4.3}$$

and the overall quality of provider score can be evaluated as the expected value of the overall quality of services scores b_s, i.e.

$$B = \sum_s g_s b_s \,, \tag{2.4.4}$$

being w_φ and g_s respectively the (positive) relevance of each property φ and service **s** under evaluation (such that $\sum_\varphi w_\varphi = 1$ and $\sum_s b_s = 1$).

The application model: overview

The reference model described in the previous section is used to define the model that was realized for the evaluation of a network of vaccination centers (ASLs). Here we use the algorithm presented in section 2.3.1 at which we refer for more details.

Let $t = 1$ be the time period of the evaluation, i.e. April-June 2010.

Let $p = 1, 2, 3$ be the provider identifier for ASL1, ASL2 and ASL3.

Let $s = 1, 2, 3$ be the service identifier for HPVV, VACL, VACP.

Let $\varphi = 1, 2, 3$ be the quality identifier for CE, OE and CO (i.e. the quality aspects or properties under evaluation).

Assuming $t = 1$ (April-June 2010) and $p = 1$ (ASL1), for $s = 1$ (HPVV) and $\varphi = 1$ (CE), the resulting service quality score is given by the probability

$$e_\varphi = e_1 = P(E) = \sum_{x=1}^{8} P(E|H_x)P(H_x) \,,$$

being

E="An expert claims that service HPVV is communicatively efficient",

H_1="Information is provided through a brochure with the aid of a person in the waiting room during the HPVV providing",

H_2="Information is provided through a brochure with the aid of a person during the vaccination phase of HPVV providing",

H_3, H_4, H_5, H_6, H_7, H_8 as stated in Table 2.3,

$P(H_1) = 0$, $P(H_2) = 0$, $P(H_3) = 0$, $P(H_4) = 0$, $P(H_5) = 0$, $P(H_6) = 0$, $P(H_7) = 1$ and $P(H_8) = 0$ as stated in Table 2.5 (see column ASL1),

$P(E|H_x) = \mu_E(x) = \sum_{k=1}^{4} \alpha_k \mu_{M_k}(x)$, with $\alpha_1 = 0.25$ the relevance given to parent of vaccinated women (stakeholder k=1) in the evaluation of CE and $\alpha_2 = 0.3$, $\alpha_3 = 0.25$, $\alpha_4 = 0.2$, respectively the relevance of doctor (k=2), nurse (k=3) and vaccinated woman (k=4); see Table 2.4 to see the values obtained for $\mu_E(x)$.

Remark 2.4.1. The event E is such that $\bigwedge_k^4 M_k \subseteq E \subseteq \bigvee_k^4 M_k$, being M_1="A parent of a vaccinated women claims that service HPVV is communicatively efficient", M_2="A doctor claims that service HPVV is communicatively efficient", M_3="A nurse claims that service HPVV is communicatively efficient", M_4="A vaccinated women claims that service HPVV is communicatively efficient".

Remark 2.4.2. $\mu_{M_1}(1) = P(M_1|H_1)$ is the probability that a parent of a vaccinated women claims service HPVV communicatively efficient, assuming true the event H_1, i.e. supposing that information is provided through a brochure with the aid of a person in the waiting room during the HPVV providing. That probability was assessed asking the following question: "With what probability do you claim communicatively efficient service HPVV supposing that information is provided through a brochure with the aid of a staff member in the waiting room during the HPVV providing ?". Thus we assessed all the $\mu_{M_k}(x) = P(M_k|H_x)$ using formula (2.4.1), considering judgments resulting from a survey over a sample of stakeholders (see also Table 2.1).

We considered also properties OE and CO for the quality evaluation of service HPVV. The resulting scores e_2, e_3 were evaluated in the same way of e_1.

Then the HPVV overall quality of service score was evaluated as $b_s = \sum_{\varphi=1}^{3} w_\varphi e_\varphi$, being w_1, w_2, w_3 respectively the relevance of CE, OE and CO.

Finally the overall quality of provider score, obtained considering HPVV, VACP and VACL, was evaluated as $B = \sum_{s=1}^{3} g_s e_s$, being g_1 the relevance of HPVV, g_2 the relevance of VACL and g_3 the relevance of VACP evaluated as the corresponding frequency of occurrence during the period of observation.

Remark 2.4.3. Note that different key performance indicators were used to model the properties of services under evaluation (i.e. CE, OE and CO). In fact, for example, for the evaluation of OE we considered the waiting time and duration time related to service providing, while for CE we used two indicators, which represent the way and time of communicating information.

Remark 2.4.4. Stakeholders of VACP and VACL were also different from those of HPVV and exactly: parent of child (k=1), doctor (k=2) and nurse (k=3).

The above application model was applied to evaluate the quality of ASL1, ASL2 and ASL3, step by step, as suggested also by the algorithm shown in section 2.3.1.

Step 1 of the evaluation

In the period April-July 2010 a survey was performed in the ASLs in order to collect *subjective* data on the quality aspects of HPVV, VACP and VACL.

Remark 2.4.5. The adjective "subjective" has not to be interpreted as "without objectivity". It refers to the fact that a subject, i.e. a stakeholder as a doctor or a patient, can express its opinion using possibly all available data and experience using the maximum of objectiveness.

A questionnaire was administered to the services stakeholders: parents and escorts of immunized children (P/E) and adolescents and young women vaccinated against anti-Human Papillomavirus (V), medical doctors (D) and nurses (N) working in the ASL. There were three sections in a questionnaire: an anonymous demographic section (number 1) with questions about age, education, marital status and job, an informative section (number 2) about the functioning of a generic vaccination service

as stated by some institutional guide lines, and a judgments section (number 3, see for example Figure 2.1) containing 8 questions on CE, 16 questions on OE and 8 questions on CO.

Judgments were expressed by the stakeholders assessing probabilities (i.e. values ranging in the interval 0-100%) to some well–defined events related to CE, OE and CO (see Table 2.2).

For example a question was: "With what probability (in percentage) do you claim communicatively efficient the HPVV service assuming that a staff member gives information on the vaccination by means of a leaflet in the waiting room ?". And another question was: "With what probability (in percentage) do you claim organizationally efficient a service assuming that the waiting time is 15 minutes and the vaccination time is 10 minutes ?". A specific relevance was assigned to the judgments of each stakeholder. For VACL and VACP the weights (α_k) were 0.3, 0.4, 0.3 respectively for P/E, D, and N. For HPVV the weights were 0.25, 0.3, 0.25 and 0.2 respectively for P/E, D, N and V. Expert judgments (probabilities) were obtained summing the stakeholders judgments taking into account the mentioned weights. The resulting maximum and minimum probability values represent the best achievable service quality scores and the worst achievable service quality scores.

Step 2 of the evaluation

In the same period of the survey, April-July 2010, performance detection, regarding the same points investigated in section number 3 of the questionnaire, were observed in the ASLs during the service providing and recorded by some operators of the QuaVaTAR group. This were *objective* data (i.e. evidence) about the waiting times, the duration times of vaccination, the way of the communication and the comfort

features.

Step 3 of the evaluation

At the end of the period April-July 2010 all gathered data, objective and subjective, were entered in a spreadsheet which implements the mathematical algorithm illustrated in section 2.3.1. Here we recall the fundamental decision of assign a relevance to the quality aspects under investigation by means of some weights. The QuaVaTaR group evaluated the relevance of CE, OE and CO respectively equal to $w_1 = 0.5$, $w_2 = 0.3$ and $w_3 = 0.2$ that gave to CE more relevance mainly because this aspect was related to the necessity, by law constraints, of obtaining a patient's informed consent for the delivery of the vaccination service. So the values were arbitrary chosen such that $w_1 > w_2 > w_3$ and in respect to the mathematical constraints $w_\varphi > 0$ and $\sum_{\varphi=1}^{3} w_\varphi = 1$.

Results

From April to July 2010 the administered questionnaires were 678 and the performance detections were 304.

For each investigated service the main socio-demographic characteristics of the interviewees are summarized in Table 2.1: category, age, education level, marital and job status. For example for HPVV it is shown that 120 over 170 interviewed parents or escorts have a high education level with a mean age of 42.7 years (and a standard deviation of 5), while 23 over 63 interviewed vaccinated girls have a high education level with a mean age of 17.1 years (and a standard deviation of 5.7). The events related to quality of services were evaluated for CE, OE and CO for HPVV, VACP and VACL and are shown in Table 2.2 with identifiers from E1 to E9.

Table 2.3 shows the relevant conditional events related to the communication efficiency in HPVV providing assuming different performance conditions from H1 to H8; the resulting probabilities in Table 2.4 were investigated during the survey from different stakeholders' points of view; for example from an expert point of view (i.e. a combination of all stakeholders' points of view) the minimum value (7.16%) is assessed assuming true the performance H8 (information is not provided); on the contrary the maximum value (87.5%) is assessed assuming true the performance H1 (information is given through a brochure with the aid of a person in the waiting room).

Table 2.5 shows that ASL1 and ASL2 performed differently from ASL3 as the occurrence of the communicative performances from H1 to H8 were not the same.

Table 2.6 shows the probabilities of events from E1 to E8 , i.e. scores regarding the quality of services, their minimum and maximum values, for each service, aspect and provider under evaluation.

Finally Table 2.7 shows the quality scorecard for all the services under evaluation, that is the main output of the measurement process under study. For example the HPVV quality scores were 73.83%, 71.57%, 67.80% respectively for ASL1, ASL2 and ASL3 considering that the achievable values may range between a minimum of 20.82% and a maximum of 86.10%. In the table the results for VACP and VACL can be read in the same way.

2.5 Conclusions and future work

Evaluation of the quality of a health care provider can be conducted using a probabilistic approach. Such an approach provides an effective way to benchmark quality of providers when key performance indicators used to represent quality are different, are (possibly) described by means of "natural language" and it is necessary to take into consideration different stakeholders' points of view. Future work regards: 1) further development of such a general approach; 2) the evaluation of all vaccination centers in the Lazio region of Italy as an actual test. In fact, for example, instead of using the algorithm 2.3.1, the evaluation of services could be done considering in a separate way the stakeholders' points of view and then aggregating the information. For this aim, we need to refer to coherent t-norms. Then we can obtain both an evaluation by medical staff (doctors, nurses, etc.), and an evaluation by clients (patients, caregivers, etc.). We could compare such evaluations by means of similarity functions and try to manage the possible differences. Finally, we could aggregate those evaluations to form a final expert evaluation.

The Author intends also to investigate how to reuse and embed such mathematical tools into a decision support system for the clinical governance of a hospital, such as the Business Simulator for Health Care (BuS-4H) system, which the Author, who is the creator of the corresponding research project, wishes to develop in collaboration with an Italian company called "SiliconDev srl" and several clinical, mathematics and engineering departments of the University of Rome "Sapienza", the University of Rome "Tor Vergata", the "Fatebenefratelli" Association for Biomedical Research (AFaR) and the "San Giovanni Calibita Fatebenefratelli Isola Tiberina" Hospital in Rome.

Efficienza comunicativa IdQuest .....

Il servizio che si sta considerando è la **vaccinazione anti-HPV**

n°	Domanda	Giudizio tra 0 e 100
1	Tra 0 e 100 , quanto crede sia efficiente un servizio che si preoccupa di informarla attraverso un opuscolo con l'ausilio di una persona durante l'attesa in sala d'aspetto?	.....
2	Tra 0 e 100 , quanto crede sia efficiente un servizio che si preoccupa di informarla attraverso un opuscolo con l'ausilio di una persona durante la vaccinazione?	.....
3	Tra 0 e 100 , quanto crede sia efficiente un servizio che si preoccupa di informarla attraverso un opuscolo senza l'ausilio al momento della chiamata diretta?	.....
4	Tra 0 e 100 , quanto crede sia efficiente un servizio che si preoccupa di informarla attraverso un opuscolo senza l'ausilio di una persona durante l'attesa in sala d'aspetto?	.....
5	Tra 0 e 100 , quanto crede sia efficiente un servizio che si preoccupa di informarla attraverso un opuscolo senza l'ausilio di una persona durante la vaccinazione?	.....
6	Tra 0 e 100 , quanto crede sia efficiente un servizio che si preoccupa di informarla a voce con l'ausilio di una persona durante l'attesa?	.....
7	Tra 0 e 100 , quanto crede sia efficiente un servizio che si preoccupa di informarla a voce con l'ausilio di una persona durante la vaccinazione?	.....
8	Tra 0 e 100 , quanto crede sia efficiente un servizio che non le offre nessuna informazione?	.....

Figure 2.1: Part of the original questionnaire (in Italian language) regarding the communication efficiency (CE).

Survey Results (demographic characteristics)															
	HPVV					VACL					VACP				
Stakeholders	n°	Age (SD)	High School	Married	Job	n°	Age (SD)	High School	Married	Job	n°	Age (SD)	High School	Married	Job
P/E	170	42.7 (5.0)	120	159	98	162	34.9 (5.7)	125	147	98	226	34 (6.7)	200	213	173
N	11	45.3 (9.0)	11	8	10	8	39.7 (13.2)	8	5	6	8	39.2 (8.4)	8	5	8
D	18	49.1 (9.4)	18	16	17	3	39 (9.8)	3	2	3	9	46.1 (9.6)	9	7	9
V	63	17.1 (5.7)	23	3	3	-	-	-	-	-	-	-	-	-	-
Total	262	36.0 (12.9)	172	186	128	173	35.2 (6.2)	136	154	107	243	35.6 (7.2)	217	225	190

Table 2.1: Interviewees' demographic characteristics (parents/escorts (P/E), nurses (N), medical doctors (M) and vaccinated women (V)).

Quality of Services	
Event	**Event proposition**
E1	An expert claims communicatively efficient the HPVV
E2	An expert claims organizationnaly efficient the HPVV
E3	An expert claims confortable the HPVV
E4	An expert claims communicatively efficient the VACP
E5	An expert claims organizationnaly efficient the VACP
E6	An expert claims confortable the VACP
E7	An expert claims communicatively efficient the VACL
E8	An expert claims organizationnaly efficient the VACL
E9	An expert claims confortable the VACL

Table 2.2: Main events (i.e. quality of services HPVV, VACP and VACL) under evaluation.

Survey Questions		Service Performances	
Event	**Event proposition**	**Event**	**Event proposition**
E1 | H1	You* claim communicatively efficient the HPVV assuming H1 true	H1	information is provided through a brochure with the aid of a person in the waiting room during the HPVV providing
E1| H2	You* claim communicatively efficient the HPVV assuming H2 true	H2	information is provided through a brochure with the aid of a person during the vaccination phase of the HPVV providing
E1 | H3	You* claim communicatively efficient the HPVV assuming H3 true	H3	information is provided through a brochure call without the aid of a person at the time of the HPVV service direct
E1 | H4	You* claim communicatively efficient the HPVV assuming H4 true	H4	information is provided through a brochure without the aid of a person in the waiting room during the HPVV providing
E1 | H5	You* claim communicatively efficient the HPVV assuming H5 true	H5	information is provided through a brochure without the aid of a person during the vaccination phase of the HPVV providing
E1 | H6	You* claim communicatively efficient the HPVV assuming H6 true	H6	information is provided orally by a person in the waiting room during the HPVV providing
E1 | H7	You* claim communicatively efficient the HPVV assuming H7 true	H7	information is provided orally by a person during the vaccination phase of the HPVV providing
E1 | H8	You* claim communicatively efficient the HPVV assuming H8 true	H8	information is not provided during the HPVV providing

Table 2.3: Conditional events (i.e. quality of services) and their conditioning events (i.e. service performances) for the HPVV quality evaluation. Probabilities of the resulting conditional events $\{E|H_x\}_x$ were investigated through a questionnaire that was administered to a sample of stakeholders. For example, about $E|H_1$, the question was: "With what probability (in percentage) do you claim communicatively efficient the HPVV service assuming that a person (i.e. an operator of the provider's staff) gives information on the vaccination by means of a brochure in the waiting room ?".

.

HPVV Survey Results					
Event ID	**PROBABILITIES**				
	Patient/ Escort	**Nurse**	**Doctor**	**Woman**	**Expert**
E1 | H1	88.43%	88.18%	91.19%	77.7%	**87.05%**
E1| H2	80.3%	78.18%	84.38%	74.97%	79.93%
E1 | H3	59.79%	66.36%	53.75%	57.46%	59.16%
E1 | H4	54.13%	65.45%	49.38%	50.16%	54.74%
E1 | H5	46.13%	48.18%	34.06%	43.13%	42.42%
E1 | H6	75.8%	76.36%	77.5%	68.85%	75.06%
E1 | H7	71.49%	70,00%	76.25%	66.64%	71.57%
E1 | H8	7.44%	5.45%	3.75%	14.05%	**7.16%**

Table 2.4: Results of the survey on the communicative efficiency of HPVV service. The values of last column were obtained through a formula (2.3.1) and it represents the expert evaluation obtained as aggregation of stakeholders' evaluations (columns 2, 3, 4 and 5).

HPVV Performances			
Event ID	**Probabilities**		
	ASL 1	**ASL 2**	**ASL 3**
H1	0%	0%	0%
H2	0%	0%	0%
H3	0%	0%	**38.5%**
H4	0%	0%	0%
H5	0%	0%	0%
H6	0%	0%	**23%**
H7	**100%**	**100%**	**38.5%**
H8	0%	0%	0%

Table 2.5: Probabilities of conditioning events (i.e. service performances) under evaluation for the HPVV service providing by three providers (ASL1, ASL2, ASL3).

Quality of Services						
Main Events		**Probabilities**				
Event ID	**Event proposition**	**ASL1**	**ASL2**	**ASL3**	***Min***	***Max***
E1	An expert claims communicatively efficient the HPVV	71.57%	71.57%	67.6%	*7.16%*	*87.05%*
E2	An expert claims organizationnaly efficient the HPVV	67.79%	69.12%	64.66%	*22.15%*	*82.91%*
E3	An expert claims confortable the HPVV	88.52%	75.23%	73,00%	*52.99%*	*88.52%*
E4	An expert claims communicatively efficient the VACP	-	80.88%	70.53%	*1.84%*	*90.95%*
E5	An expert claims organizationnaly efficient the VACP	-	78.94%	80.33%	*17.8%*	*94.91%*
E6	An expert claims confortable the VACP	-	68.04%	65.92%	*29.14%*	*95.61%*
E7	An expert claims communicatively efficient the VACL	-	67.7%	68.61%	*6.3%*	*90.22%*
E8	An expert claims organizationnaly efficient the VACL	-	78.88%	72.8%	*23.31%*	*81.22%*
E9	An expert claims confortable the VACL	-	70.84%	70.46%	*26.46%*	*90.17%*

Table 2.6: Main events (i.e. quality of services) under evaluation for three providers (ASL1, ASL2, ASL3) and their probabilities of being true from an expert's point of view. The values in this table were obtained using values from tables 2.5 and 2.4 and formula (2.3.2). The minimum and maximum probabilities were obtained supposing that services were 100% delivered in the worst or best performance condition.

Quality Scorecard					
Services	**ASL1**	**ASL2**	**ASL3**	***Min***	***Max***
HPVV	73.83%	71.57%	67.80%	*20.82%*	*86.10%*
VACP	-	77.73%	72.55%	*12.09%*	*93.07%*
VACL	-	71.68%	70.23%	*17.44%*	*87.51%*

Table 2.7: Service overall quality of a network of vaccination centers (ASL1, ASL2, ASL3) in the Lazio Region of Italy. The values in this table were obtained using values from table 2.6 and formula (2.4.3). Note that ASL1 did not provide VACP and VACL services.

Bibliography

[1] Zadeh LA, *Fuzzy sets*, Information and Control 8(1965).

[2] Coletti G, Scozzafava R, *Conditioning in a coherent setting: theory and applications*, Fuzzy Sets and Systems 155, (2005).

[3] Coletti G, Gervasi O, Tasso S, Vantaggi B, *Generalized Bayesian inference in a fuzzy context: From theory to a virtual reality application*, Computational Statistics & Data Analysis *56(4)* (2012).

[4] Coletti G, Scozzafava R, Paulon L, Vantaggi B, *Measuring the Quality of Health-Care Services: A Likehood-Based Fuzzy Modeling Approach*, Symbolic and Quantitative Approaches to Reasoning with Uncertainly Lecture Notes in Computer Science (2007).

[5] Coletti G, Scozzafava R, *Characterization of coherent conditional probabilities as a tool for their assessment and extension*, International Journal of Uncertainty Fuzziness And Knowledge-Based Systems *4(2)* (1996).

[6] Coletti G, Scozzafava R, *Conditioning and Inference in Intelligent Systems*, Soft Computing *3* (1999).

[7] Coletti G, Scozzafava R, *Probabilistic logic in a coherent setting*, Trends in Logic *15* (2002).

[8] de Finetti B, *Theory of Probability*, Wiley series in probability and mathematical statistics (1993).

[9] Scozzafava R, Vantaggi B, *Fuzzy Inclusion and Similarity through Coherent Conditional Probability*, Fuzzy Sets and Systems 160 (2009).

[10] Baioletti M, Capotorti A, Tulipani S, Vantaggi B, *Simplification rules for the coherent probability assessment problem*, Annals of Mathematics and Artificial Intelligence *35(1-4)* (2002).

[11] Coletti G, Scozzafava R, *Conditional probability, fuzzy sets, and possibility: a unifying view*, Fuzzy Sets and Systems *144(1)* (2004).

[12] Vantaggi B, *The role of coherence for handling probabilistic evaluations and independence*, Soft Computing, *9(8)* (2005).

[13] Rifqi M, Bothorel S, Bouchon-Meunier B, *Toward general measures of comparison of objects*, Fuzzy Sets and Systems *84* (1996).

[14] De Baets B, De Meyer H, *Transitivity-preserving fuzzification schemes for cardinality-based similarity measures*, European Journal of Operational Research *160(3)* (2005).

[15] Zadeh LA, *Probability measures of fuzzy events*, Journal Of Mathematical Analysis and Applications *23* (1968).

[16] Parasuraman A, Zeithaml VA, Berry LL, *Servqual - A Multiple-Item Scale For Measuring Consumer Perceptions Of Service Quality*, Journal of Retailing *64(1)* (1988).

[17] Tamil Mani B, Arun Kuma S, Prabhakara S, Ravichandran K, *Influence of Service Quality on Customer Satisfaction: Application of Servqual Model*, International Journal of Business and Management *4* (2010).

[18] Lin DJ, Sheu IC, Pai JY, Bair A, Hung CY, Yeh YH, Chou MJ, *Measuring patient's expectation and the perception of quality in LASIK services*, Health And Quality Of Life Outcomes *7* (2009).

[19] Ryan M, *Discrete choice experiments in health care - NICE should consider using them for patient centred evaluations of technologies*, British Medical Journal *328(7436)* (2004).

[20] Ryan M, Skatun D, *Modelling non-demanders in choice experiments*, Health Economics *13(4)* (2004).

[21] Donebedian A, *Evaluating the Quality of Medical Care*, Milbank Memorial Fund Quarterly *44(1)* (1966).

[22] Meleo C, Campolongo A, D'Anna C, Mangia ML, Sgricia S et al., Maurici M, *Application of the QuavaTAR model to vaccination services in Latium, Italy*, Ig. Sanit Pubbl. *66* (2010).

Chapter 3

Case mix analysis using multiobjective optimization

3.1 Introduction

There is general agreement that hospital case mix complexity, i.e. the distribution of patients (i.e., cases) treated in hospital over a period of interest (for instance one year), is a key factor in determining hospital costs. Furthermore, case mix is intimately involved in the classification of hospitals, in rate-setting, in reimbursement mechanisms and in the national debate over the efficiency and effectiveness of health care.

Hospital care activity in Italy actually consumes approximately the 45% of the country's overall economic health resources. Thus, clinical and economical multi–criteria case mix analysis is a very important task for managers of hospitals.

Thus, in this chapter, we face the problem of performing a case mix analysis using a multiobjective optimization approach. In this particular context, in fact, optimization of clinical and business administration functions often conflict. We will evaluate in detail case mix different from to the actual one, using Pareto optimal

solutions of a well–defined multiobjective optimization problem that we derived from the competing points of view of the Chief Executive Officer and the Chief Hospital Officer.

The target application at the end of this chapter is a first step toward an effective innovation goal, which is the building of a decision support system, called the Business Simulator for Health Care (BuS-4H) that is able to perform forecasting through the using of data–oriented functionalities, among which are storage and reporting activities (i.e., hospital dashboards), such as case mix analysis, discrete event simulations and optimization.

To realize this work, the Author worked at the University of Rome "Sapienza", University of Rome "Tor Vergata" and "Fatebenefratelli" Association for Biomedical Research, in collaboration with, as listed in a publication under preparation, Alessandra Campolongo, Carlo Maria Cellucci, Mauro Cenerelli, Maurizio Ferrante, Alessandro Giordani, Francesco Lagana, Lorena Lagana, Massimo Maurici, Daniele Milazzo and Francesco Rinaldi.

3.2 Background

3.2.1 A multiobjective optimization problem

This and the following three sections contain basic facts based on books [1][2][3].

Definition 3.2.1. We study a *multiobjective optimization problem* of the form

$$minimize\{f_1(x), f_2(x), ..., f_k(x)\},$$

$$subject\ to\ x \in S,$$

in which we have k (≥ 2) *objective functions* $f_i : \mathbb{R}^n \to \mathbb{R}$.

We denote the vector of objective functions by $f(x) = (f_1(x), f_2(x), ..., f_k(x))^*$. The *decision vector* $x = (x_1, x_2, ..., x_n)^*$ belong to the (nonempty) *feasible region* (set) S, which is a subset of the *decision variable space* $\mathbb{R}^n$. We do not yet fix the form of the *constraints functions* forming S, but refer to S in general.

From now on, we denote the image of the feasible region by Z and call it a *feasible objective region*. It is a subset of the *objective space* $\mathbb{R}^k$. The elements of Z are called *objective (function) vectors* or *criterion vectors* and are denoted by $f(x)$ or $z = (z_1, z_2, ..., z_k)^*$, where $z_i = f_i(x)$ for all $i = 1, ..., k$ are *objective (function) values* or *criterion values*. The words in the parentheses above are usually omitted for short.

Definition 3.2.2. When all the objective functions and the constraint functions forming the feasible region are linear, the multiobjective optimization problem is called *linear*. In brief, it is a multiobjective linear programming (MOLP) problem. If at least one of the objective or constraints functions is nonlinear, the problem is called a nonlinear multiobjective optimization problem.

3.2.2 Aspiration levels and utility functions

Definition 3.2.3. Objective function values that are satisfactory or desirable to the decision maker are called *aspiration levels* and are denoted by $\overline{z}_i, i = 1, ..., k$. The vector $\overline{z} \in \mathbb{R}^k$, which consist of aspiration levels, is called a *reference point*.

Definition 3.2.4. A function $U : \mathbb{R}^k \to \mathbb{R}$ representing the preferences of the decision maker among the objective vectors is called a *value function*. So, let z_1 and $z_2 \in Z$ be two different objective vectors. If $U(z_1) > U(z_2)$, then the decision maker prefers z_1 to z_2. If $U(z_1) = U(z_2)$, then the decision maker finds the objective vectors equally desirable, that is, they are indifferent.

Remark 3.2.1. Sometimes the term *utility function* is used instead of the value function.

3.2.3 Pareto optimality

Definition 3.2.5. A decision vector $x' \in S$ is Pareto optimal if there does not exist another decision vector $x \in S$ such that $f_i(x) \leq f_i(x')$ for all $i = 1, ..., k$ and $f_j(x) < f_j(x')$ for at least one index j .

Definition 3.2.6. An objective vector $z' \in Z$ is Pareto optimal if there does exist another objective vector $z \in Z$ such that $z_i \leq z'_i$ for all $i = 1, ..., k$ and $z_j < z'_j$ for at least one index j; or equivalently, z' is Pareto optimal if the decision vector corresponding to it is Pareto optimal.

Definition 3.2.7. A decision vector $x' \in S$ is weakly Pareto optimal if there does not exist another decision vector $x \in S$ such that $f_i(x) < f_i(x')$ for all $i = 1, ..., k$.

Definition 3.2.8. An objective vector $z' \in Z$ is weakly Pareto optimal if there does exist another objective vector $z \in Z$ such that $z_i < z'_i$ for all $i = 1, ..., k$; or equivalently, if the decision vector corresponding to it is weakly Pareto optimal.

Definition 3.2.9. A decision vector $x' \in S$ is locally Pareto optimal if there exists $\delta > 0$ such that x' is Pareto optimal in $S \cap B(x', \delta)$, B being an n-dimensional interval (ball) around x'.

Definition 3.2.10. An objective vector $z' \in Z$ is locally Pareto optimal if the decision vector corresponding to it is locally Pareto optimal.

3.2.4 The dominance relation

Definition 3.2.11. A solution x_1 is said to dominate an other solution x_2, if both conditions 1 and 2 are true:

1. The solution x_1 is no worse than x_2 in all objectives, or $f_j(x_1) \not\triangleright f_j(x_2)$ for all $j = 1, 2, ..., k$.

2. The solution x_1 is strictly better than x_2 in at least one objective, or $f_j(x_1) \lhd f_j(x_2)$ for at least one $j \in 1, 2, ..., k$.

If either of the above conditions is violated, the solution x_1 does not dominate the solution x_2. If x_1 dominates the solution x_2 (or mathematically $x_1 \preceq x_2$), it is also customary to write any of the following:

- x_2 is dominated by x_1;
- x_1 is non-dominated by x_2;

- x_1 is non-inferior to x_2 .

Remark 3.2.2. The dominance relation is *transitive*. This is because if $x_1 \preceq x_2$ and $x_2 \preceq x_3$, then $x_1 \preceq x_3$.

Remark 3.2.3. The dominance relation has another interesting property: if solution x_1 does not dominate solution x_2, this does not imply that x_2 dominates x_1. In this case we say that such solutions are incomparable.

Definition 3.2.12. A solution x_1 strongly dominates a solution x_2 (or $x_1 \prec x_2$), if solution x_1 is strictly better than solution x_2 in all k objectives.

3.2.5 Pareto front

Definition 3.2.13. The Pareto optimal set, denoted P_x, is a set of solutions $x = (x_1, ..., x_n)$ that are not dominated by other solutions.

Definition 3.2.14. The Pareto optimal front, denoted P_f, is the image of P_x using the objective functions $f(x) = (f_1(x), ..., f_k(x))$.

Remark 3.2.4. It must still be kept in mind that in many cases only a finite set of locally weakly Pareto optimal solutions may be computationally available (see also Figure 3.1).

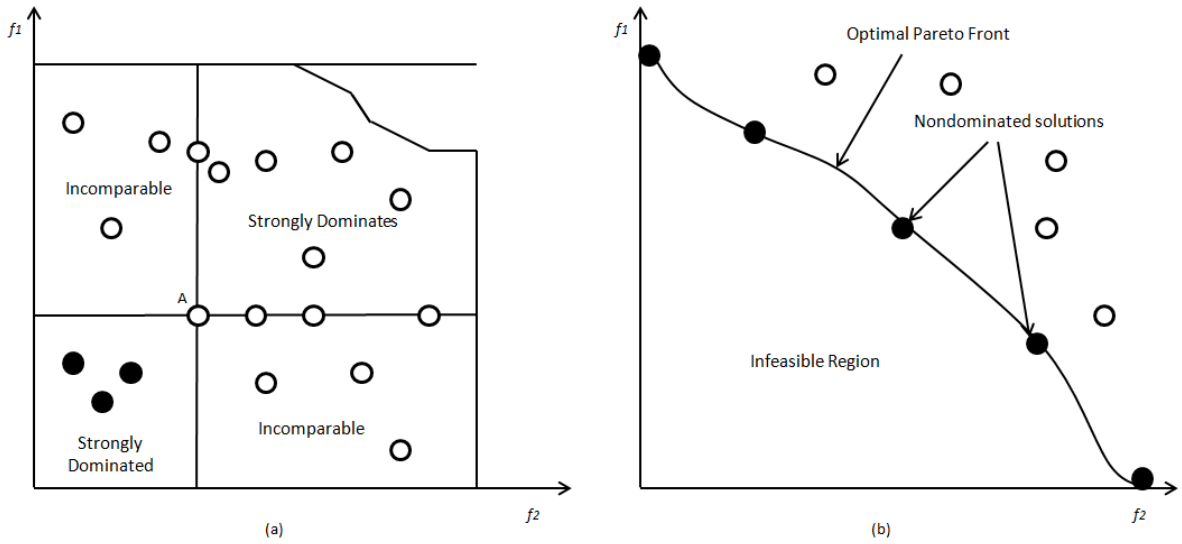


Figure 3.1: Illustration of (a) the Pareto Dominance relationship between candidate solutions relative to solution A and (b) the relationship between the approximated Pareto front and the true Pareto front.

3.3 Algorithms

Multiobjective evolutionary algorithms (MOEA) represent one of the more popular stochastic search methodologies for solving an MOOP. In particular, we are interested in finding a good approximation of the Pareto front P_f, and this approximate (finite) set, $\hat{P}_f$, should satisfy the following optimization goals:

- Minimize the distance between $\hat{P}_f$ and P_f.
- Obtain a good distribution of generated solutions along the $\hat{P}_f$.
- Maximize the spread of the discovered solutions.

The general MOEA framework can be represented in the pseudocode shown in 3.2 and it can be shown that most MOEAs fit into this framework.

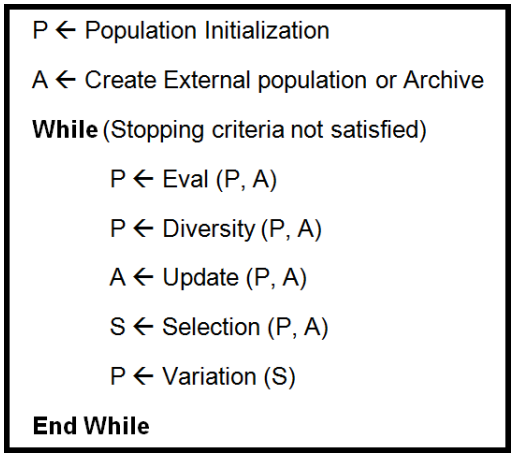


Figure 3.2: Structure of a generic MOEA.

For instance, the non-dominated sorting genetic algorithm II (NSGA-II) [3][4] is a MOEA available in the Optimization Toolbox of Matlab.

3.4 Applications

3.4.1 Case mix analysis of a gynecological ward

The purpose of this study is to show precisely how to perform, with the support of some mathematical tools [6] [7] [8] [9], a case mix adjustment for an obstetrics and gynecology ward, i.e., an evaluation of the number of patients that would be sustainable to treat per year, both from an economic and a clinical point of view, using also the Diagnosis Related Groups (DRG) standard for the classification of health care activities.

This evaluation takes into account the most important supplied services of the ward under study: Caesarean section without complications and comorbidities (DRG 371), and vaginal childbirth without complications (DRG 373).

Taking into account different contrasting goals, it is shown how to identify a better case mix in respect to the actual; the optimized case mix would both reduce the Caesarean section rate (a clinical goal) and increase profits (an economic goal).

Materials and methods

Preliminaries

This study was carried out in the obstetrics and gynecological ward of the Fatebenefratelli San Giovanni Calibita (FBF-SGC) Hospital in Rome, one of the most important Italian hospitals in terms of number of childbirth cases both at the regional and the national level. The results of the study were critical for the hospital, both in qualitative and quantitative terms. A research group, designated the Business Administration Simulator (BAS) team and composed of doctors, engineers, statisticians

and other experts in health care was formed. The services under evaluation were classified as DRG 371 and DRG 373 using version 24 of the DRG classification system [10]; they were, respectively, Cesarean section without complications or comorbidities and vaginal childbirth without complications. The top managers of the FBF-SGC Hospital, respectively, the Chief Executive Officer (CEO) and the Chief Hospital Officer (CHO), were involved in the evaluation. The year 2010 was considered the datum point of the actual case mix for which improvement was requested.

Step 1 - Data Definition, Gathering and Analysis

First of all it was necessary to support the top managers in defining and evaluating the most important key performance indicators (KPIs) and goals related to the services under evaluation, taking into account both national and regional government recommendations regarding the most relevant indicators [5][11].

The main FBF-SGC Hospital data flow related to hospitalizations, such as SDO (resulting from hospital discharge forms) and CEDAP (resulting from hospital childbirth), as well as some other services–related data (such as costs and incomes) were imported and integrated into a single database made expressly for this study. To build a report, some useful queries were then defined and uses to create a sort of dashboard for the ward under evaluation (see Table 3.2). This report contained, among other items, the following KPIs: Caesarean section rate, profit (difference between income and costs), rate of low–length–of–stay (LOS) admissions (considering the number of hospitalizations with length of stay between 0 and 2 days), bed occupation rate, overall number of supplied birth, minimum number of supplied births, and minimum number of Caesarean sections.

Step 2 - Multiobjective Optimization

Starting from the KPIs and goals chosen by the top managers of the ward under evaluation (see Table 3.3), the problem of finding a better case mix with respect to the actual case mix can be mathematically formulated as a multiobjective optimization problem (MOOP).

Using matrix notation, the problem is

$$\begin{aligned} &minimize\big\{f_1(x), f_2(x)\big\}\,, \\ &subject\ to\ Ax \leq b\,. \end{aligned} \tag{$P1$}$$

The vector of unknowns x_1, x_2, x_3 represents the case-mix: x_1 is the number of vaginal childbirths per year with low LOS; x_2 is the number of Caesarean sections per year; and x_3 is the number of vaginal childbirth per year without low LOS. The linear system $Ax \leq b$ represents the constraints of the problem (in the sequel, it is explained why it is linear), while $f_1(x)$ and $f_2(x)$ are the contrasting objective functions expressed as a function of the vector of unknowns x.

In more detail, $f_1(x)$ is the Caesarean section rate, i.e,

$$f_1(x) = \frac{x_2}{x_1 + x_2 + x_3}.$$

$f_2(x)$ is the overall profit of the ward under evaluation, i.e,

$$f_2(x_1, x_2, x_3) = -\big(R(x_1, x_2, x_3) - C(x_1, x_2, x_3)\big)\,,$$

where R is the overall childbirth income per year (in euros), i.e,

$$R(x_1, x_2, x_3) = 1434x_1 + 2382x_2 + 1482x_3\,,$$

and C is the overall childbirth cost per year (in euros), i.e,

$$C(x_1, x_2, x_3) = 1000x_1 + 1700x_2 + 1200x_3\,.$$

Remark 3.4.1. Note that the cost of a single childbirth was estimated at about 1100 euros for a vaginal childbirth and 1700 euros for a Caesarean section. The income obtained for a single childbirth was estimated at about 1450 euros for a vaginal childbirth and 2382 euros for a Caesarean.

Focusing on the linearity of the constraints, we emphasize that this property is a consequence of the KPI formulas (see Table 3.3). For example, the bed occupation rate, i.e., the ratio (%) between the overall LOS for childbirth hospitalizations and the overall maximum LOS for childbirth hospitalization, is expressed by the following formula

$$I_2 = I_2(x_1, x_2, x_3) = \sum_{k=1}^{3} \frac{x_k\, t_k}{N_b\, 365},$$

where t_k is an estimate of the LOS for childbirth hospitalization, N_{bed} the number of available beds per day in the ward, and 365 is the number of days the hospital is open per year.

To set a goal for I_2 means to fix a tolerability range $[I_{2min}, I_{2max}]$ for the bed occupation rate and, in mathematical terms, to write the following inequality:

$$I_{2min} \leq I_2 = \frac{\sum_k x_k\, t_k}{N_b\, 365} \leq I_{2max}\,.$$

After some simple algebra, this expression becomes

$$I_{2min}\; N_b\; 365 \leq \sum_k x_k\; t_k \leq I_{2max}\; N_b\; 365\;.$$

Then, coefficients t_k are exactly in a row of the matrix A, while $I_{2min}\; N_b\; 365$ and $I_{2max}\; N_b\; 365$ are in b.

Proceeding in the same way for I_2 and taking into consideration all the KPIs and goals chosen by the top managers, we can write

$$\begin{cases} 55\% \leq I_1(x_1, x_2, x_3) = \dfrac{x_1}{x_1 + x_3} \leq 80\% \\ \\ 75\% \leq I_2(x_1, x_2, x_3) = \dfrac{t_1 x_1 + t_2 x_2 + t_3 x_3}{N_b \; 365} \leq 100\% \\ \\ 3800 \leq I_3(x_1, x_2, x_3) = x_1 + x_2 + x_3 \leq 4200 \\ \\ 20\% \leq I_4(x_1, x_2, x_3) = \dfrac{x_2}{x_1 + x_2 + x_3} \leq 100\% \\ \\ x_k > 0 \; , \; k = 1, 2, 3 \end{cases}$$

which can be split as follows

$$
\begin{cases}
\dfrac{x_1}{x_1 + x_3} \geq 0.80 \\[2ex]
\dfrac{x_1}{x_1 + x_3} \leq 0.55 \\[2ex]
\dfrac{t_1 x_1 + t_2 x_2 + t_3 x_3}{N_b \; 365} \leq 1.00 \\[2ex]
\dfrac{t_1 x_1 + t_2 x_2 + t_3 x_3}{N_b \; 365} \geq 0.75 \\[2ex]
x_1 + x_2 + x_3 \leq 4200 \\[2ex]
x_1 + x_2 + x_3 \geq 3800 \\[2ex]
\dfrac{x_2}{x_1 + x_2 + x_3} \geq 0.20 \\[2ex]
\dfrac{x_2}{x_1 + x_2 + x_3} \leq 1.00 \\[2ex]
x_k > 0 \; , \; k = 1, 2, 3
\end{cases}
$$

The above expressions can be rewritten, after some simple algebra (denominators positive), as

$$\begin{cases} x_1 \leq 0.55\ (x_1 + x_3) \\ x_1 \geq 0.80\ (x_1 + x_3) \\ t_1 x_1 + t_2 x_2 + t_3 x_3 \leq 1.00\ N_b\ 365 \\ t_1 x_1 + t_2 x_2 + t_3 x_3 \geq 0.75\ N_b\ 365 \\ x_1 + x_2 + x_3 \leq 4200 \\ x_1 + x_2 + x_3 \geq 3800 \\ x_2 \geq 0.20\ (x_1 + x_2 + x_3) \\ x_2 \leq 1.00\ (x_1 + x_2 + x_3) \\ x_k > 0\ ,\ k = 1, 2, 3 \end{cases}$$

And, finally, written simply in matrix notation as

$$Ax \leq b.$$

A variety of algorithms can be used to solve problem (P1). In the multiobjective optimization literature the solutions techniques are typically divided into two groups:

- solution techniques with a prior articulation of preferences: objective functions are combined into a single function. In this case the typical output is a single non–dominated solution;

- posteriori articulation techniques: objective functions are analyzed separately. In this case, the methods attempt to capture the whole Pareto front, and a list of non–dominated solutions is given as output.

The goal we have is that of finding one or more case mixes (the so–called Pareto optimal solutions). Because this is a difficult task in practice, we try to find a new feasible solution that improves all objective functions with respect to a given starting solution (the actual case mix).

A multiobjective evolutionary algorithm called NSGA-II [4] was chosen. This algorithm is available in Matlab as part of the Optimization Toolbox; in a single run it may produce many solutions that approximate the Pareto optimal solutions. The consequence is that the solution contains more than one case mix; the top managers can identify the case-mix that best suits their organizational needs.

Results

Results of Step 1

Table 3.1 shows the number of hospitalizations provided by the FBF-SGC Hospital and by the gynecology and obstetrics ward during one year (2010), the overall income, the number of Caesarean sections without complications (DRG 371) and the number of vaginal childbirths without complications (DRG 373). Table 3.2 shows the main KPIs that describe the performance of the gynecology and obstetrics ward over the same period. Finally, focusing on DRG 371 and 373, Table 3.3 shows the KPIs and goals chosen by the top managers.

	Incomes (Year 2010)	**Hospitalizations (Year 2010)**
FBF-SGC Hospital	100%	20378 patients
Obstetric and gynaecology ward	22% of overall hospital	5850 patients
DRG 371 + DRG 373	14% of overall hospital incomes, 65% of overall ward incomes	3946 childbirths

Table 3.1: Relevance of the services under study (DRG 371 and 373) delivered by the FBF-SGC Hospital during 2010.

Results of Step 2

The mathematical formulation of the problem, as stated by the top managers using Table 3.3, is the following:

$$minimize$$

$$f_1(x_1, x_2, x_3) = \frac{x_2}{x_1 + x_2 + x_3},$$

$$f_2(x_1, x_2, x_3) = -[(1434x_1 + 2382x_2 + 1482x_3) - (1000x_1 + 1700x_2 + 1200x_3)]$$

$$subject\ to$$

$$Ax = \begin{pmatrix} -0.45 & 0.00 & 0.55 \\ -0.20 & 0.00 & 0.80 \\ 1.00 & 1.00 & 1,00 \\ -1.00 & -1.00 & -1.00 \\ 2,00 & 5.50 & 3.70 \\ -2.00 & -5.50 & -3.70 \\ 0.20 & -0.80 & 0.20 \\ -1.00 & 0.00 & -1.00 \\ -1.00 & 0.00 & 0.00 \\ 0.00 & -1.00 & 0.00 \\ 0.00 & 0.00 & -1.00 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \leq b = \begin{pmatrix} 0 \\ 0 \\ 4200 \\ -3800 \\ 16790 \\ -12593 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Figure 3.3 shows the approximated Pareto–front: the x axis is the Caesarean birth rate f_1, and the y axis is the profit f_2 resulting from the overall childbirths.

Table 3.4 shows the actual case mix (see row 1) and, among others, the Pareto optimal case mix obtained using the NSGA-II. Case mix number 5 appears very interesting for improving the performance of the ward under evaluation, as does case mix 13, which may be more feasible to realize compared to the actual case mix (case mix 1).

3.5 Conclusions and future work

A case mix analysis, i.e. an analysis of the distribution of patients (i.e., cases) treated in a hospital over a period of interest, can be effectively performed using a multiobjective optimization approach. The results obtained from a preliminary case study applied to the case mix adjustment of an obstetrics and gynecology ward show how it is possible to take into consideration clinical and economic interests that are often in contrast so as to achieve a global improvement. Such an analysis may be useful in and of itself or as a preliminary step before conducting a further discrete-event simulation (DES) and optimization (Opt) [12][13][14]. In other words, with a case-mix analysis (at least) one case mix target is determined, while by means of DES and Opt, the means of reaching a defined target in terms of processes and resources is investigated.

In future work, the Author intends to investigate the use of DES and Opt in combination with a case mix analysis.

The final goal is the embedding of mathematical tools of this type into a decision support system for the clinical governance of a hospital, such as the Business Simulator for Health Care (BuS-4H) system, which the Author, who is the creator of the corresponding research project, wishes to develop in collaboration with an Italian company called "SiliconDev srl" and several clinical, mathematics and engineering departments of the University of Rome "Sapienza", the University of Rome "Tor Vergata", the "Fatebenefratelli" Association for Biomedical Research (AFaR) and the "San Giovanni Calibita Fatebenefratelli Isola Tiberina" Hospital in Rome.

KPIs of Obstetrics and Gynecology ward	
Definition	**Actual value**
Amount of time patients remain in the ward	20238 (days)
Ratio (%) between the overall LOS and the overall maximum LOS	120.54%
Ratio between number of ward discharged patients and ward number of beds	8.76
Overall LOS divided by the overall number of ward hospitalizations	4.18 (days)
Total services (in terms of DRGs weights) delivered by the ward divided by the overall number of ward hospitalizations	0.59
Percentage of non surgical DRGs delivered by the ward with respect to the overall number of ward hospitalizations	59.92%
Percentage of surgical DRGs delivered by the ward with respect to the overall number of ward hospitalizations	47.08%
Percentage of short not surgical DRGs (with length of stay less or equal to 2) delivered by the ward with respect to the overall number of ward hospitalizations	14.41%
Percentage of the overall patients re-hospitalized by the ward in respect to the overall number of ward hospitalizations	0.44%
Ratio between overall emergencies managed by the ward and the overall number of ward hospitalizations	0.15%
Percentage of the Caesarean childbirths witg respect to the overall childbirths	45%
Ward mortality, i.e. percentage of overall deaths occurred in the ward with respect to the overall number of ward hospitalizations	0%
Overall childbirths	4045
Overall ward cost	10.63% of the overall hospital cost
Overall ward income	22.21% of the overall hospital income

Table 3.2: Report of the actual (year 2010) performance of the obstetrics and gynecology ward of the FBF-SGC Hospital.

KPIs of services under evaluation		
Definition	**Symbol**	**Goal**
Percentage of vaginal childbirth delivered by the ward with respect to the overall vaginal childbirths	$I_1(x_1,x_2,x_3)$	Between 55% and 80%
Ratio (%) between the overall LOS for childbirth hospitalizations and the overall maximum LOS for childbirth hospitalization	$I_2(x_1,x_2,x_3)$	Between 75% and 100%
Overall childbirths	$I_3(x_1,x_2,x_3)$	Between 3.800 and 4.200
Percentage of Caesarean childbirths with respect to overall childbirths	$I_4(x_1,x_2,x_3)$	Greater than 20%
Percentage of Caesarean childbirths with respect to the overall childbirths	$f_1(x_1,x_2,x_3)$	Minimize
Profit due to childbirths	$f_2(x_1,x_2,x_3)$	Maximize

Table 3.3: KPIs and goals chosen by the top managers for improving the obstetrics and gynecology ward with respect to provision of the services under evaluation (DRG 371 and 373). KPIs and goals can be stated in the function that describes the case mix and contains the variables x_1, x_2, x_3 :
- x_1 is the number of vaginal childbirths with low LOS (between 0 and 2 days);
- x_2 is the number of Caesarean sections;
- x_3 is the number of vaginal childbirths without low LOS (greater than 2 days).

Case-mix ID	x1	x2	x3	x1+x2+x3	f1 (x1, x2, x3)	f2 (x1, x2, x3)	Dominant (YES/NO)
(Actual) 1	**697**	**1760**	**1491**	**3948**	**45%**	**1.923**	**-**
2	2165	850	974	3990	21%	**1.794**	NO
3	2165	905	974	4043	22%	**1.831**	NO
4	2163	1009	973	4145	24%	**1.901**	NO
5	2148	1095	929	4172	26%	**1.941**	YES
6	2110	1174	906	4190	28%	**1.972**	YES
7	2048	1254	856	4158	30%	**1.986**	YES
8	2054	1304	840	4198	31%	**2.018**	YES
9	2047	1361	771	4180	33%	**2.034**	YES
10	2008	1395	761	4164	34%	**2.038**	YES
11	2042	1480	667	4188	35%	**2.083**	YES
12	1951	1550	660	4161	37%	**2.090**	YES
13	1879	1712	579	4170	41%	**2.146**	YES
14	1809	1769	590	4168	42%	**2.158**	YES
15	1751	1887	529	4166	45%	**2.196**	NO
16	1686	1962	528	4177	47%	**2.219**	NO
17	1632	2086	483	4200	50%	**2.267**	NO

Table 3.4: Actual case mix (year 2010) and new case mixes of the obstetrics and gynecology ward of the FBF-SGC Hospital

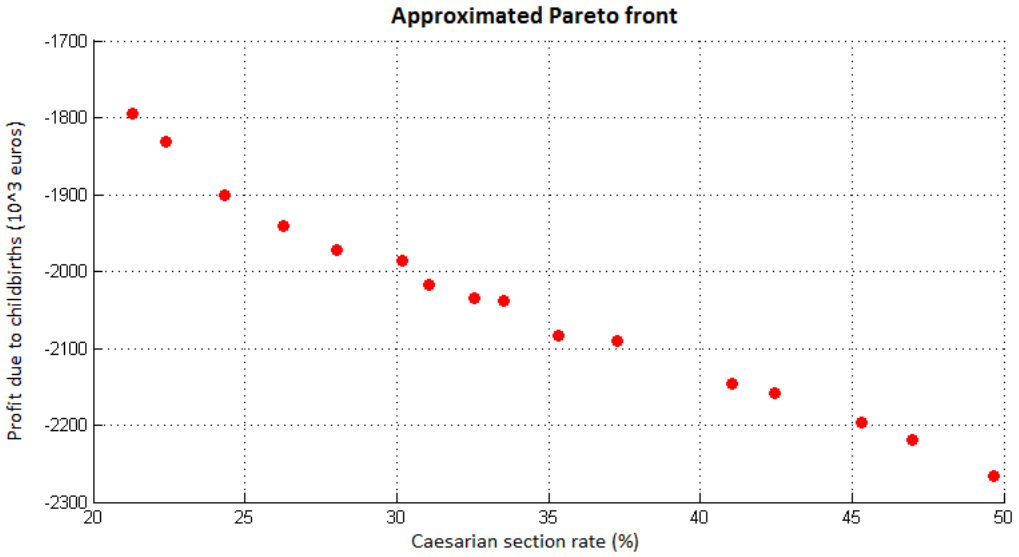


Figure 3.3: Approximated Pareto front $\hat{P}_f$. The graph represents the economic and clinical sustainability of the case mix related to the services under evaluation (DRG 371 and 373) that are provided by the obstetrics and gynecology ward of the FBF-SGC Hospital. The profit (y axis) is expressed in thousands of euros; the minus sign is due to the formulation of the resulting multiobjective optimization problem as a minimization problem.

Bibliography

[1] Miettinen K, *Nonlinear multiobjective optimization*, International Series in Operations Research & Management Science, Springer-Verlag Gmbh,(1998).

[2] Chi-Keong Goh, Kay Chen Tan, *Evolutionary Multi-objective Optimization in Uncertain Environments*, Studies in Computational Intelligence, Springer-Verlag Gmbh (2009).

[3] Deb K, *Multi-objective optimization using evolutionary algorithms*, Wiley (2001).

[4] Deb K, Pratap A, Agarwal S, Meyarivan T, *A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II*, IEEE Transactions on Evolutionary Computation *6(2)* (2002).

[5] France G, Taroni F, Donatini A, *The Italian health-care system*, Health Econ *14(suppl. 1)* (2005).

[6] Pareto V, *Manuel d'conomie politique*, Paris: Giard et Brire (1909).

[7] Osborne MJ, Rubenstein A, *A Course in Game Theory*, MIT Press (1994).

[8] Kaplan RM, Babad YM, *Balancing influence between actors in healthcare decision making*, BMC Health Serv Res 19 (2011).

[9] Baesler FF, Sepulveda JA, *Multi-objective simulation optimization: a case study in healthcare management*, International journal of industrial engineering-theory applications and practice *13(2)* (2006).

[10] Ministero della Salute, *DRG Medicare versione 24.0*, [http://www.salute.gov.it].

[11] Ministero della Salute, *Direzione generale e della programmazione sanitaria, dei livelli essenziali di assistenza e dei principi etici di sistema - Il dipartimento*, [http://www.salute.gov.it].

[12] Jacobson SH, Hall SN, Swisher JR, *Discrete-event simulation of health care systems*, In Patient Flow: Reducing Delay in Healthcare Delivery (International Series in Operations Research and Management Science). 1st edition (2006).

[13] Gunal M, Pidd M, *Discrete event simulation for performance modelling in healthcare: A review of the literature*, Working paper. Department of Management Science, Lancaster University (2010).

[14] Fu MC, *Optimization for Simulation: theory vs practice*, Informs Journal on Computing *14(3)* (2002).

[*This page was intentionally left blank*].

Chapter 4

Making prognoses using competing risk analysis

4.1 Introduction

Medical studies have demonstrated that many doctors are overly optimistic when giving prognostic information; in fact, most tend to overestimate how long a patient might live.

An estimator that is commonly used to describe prognoses is the survival time, i.e., the remaining duration of time before an event of interest (not necessarily death). If not otherwise specified, the survival time generally starts from the time of diagnosis.

In the presence of competing events, with respect to an event of interest, the usual survival analysis, for instance, the type of analysis performed using the KaplanMeier method [1], should be applied with caution, and one should to be aware of the consequences of its use.

In this chapter, the aim is to introduce the basic concepts of competing risks analysis, which is useful for the survival evaluation of patients affected by a relevant pathology of interest in the presence of other competing events.

A large number of data sets, gathered, for instance, from hospital information systems or clinical experimental studies, can be analyzed in this manner. Each record represents a subject exposed to a "treatment" (hospitalization or therapy) that is conjectured to be somehow related to an event of interest referred to as "failure" or "relapse" and also to be related to some other competing events. The observation can terminate in one of two ways: either failure is observed, or the individual is lost from observation (i.e., censored) before failure occurs. The crucial difficulty in using this kind of data is that the time to the failure's occurrence is not known for those subjects whose observation is terminated by a competing risk or by censoring.

The target application at the end of this chapter is a competing risk analysis for the identification of increased risk of stroke in asymptomatic subjects with severe internal carotid artery stenosis.

To realize this work the Author worked at the Fatebenefratelli Association for Biomedical Research (AFaR), in collaboration with, as listed in a related publication recently accepted on the "Journal of Cerebral Blood Flow & Metabolism", Mauro Silvestrini, Claudia Altamura, Raffaella Cerqua, Patrizio Pasqualetti, Giovanna Viticchi, Leandro Provinciali and Fabrizio Vernieri.

4.2 Background

4.2.1 Preliminaries

In the competing risk situation, either an event (the event of interest or some competing event) is observed during the follow-up of a subject, or, in the case of censoring (i.e., when the subject is lost from observation), nothing is observed.

The set (or family) $\mathcal{E} = \{E_0, E_1, ...E_p\}$ contains the main possible events under evaluation, which are $p >= 2$ and E_0 =“No event (failure or relapse) occurs during the follow-up of the subject”, E_1 =“The event of interest occurs during the follow-up of the subject”, or E_j =“The competing event of type j occurs during the follow-up of the subject”, with $1 < j \leq p$.

In the following we suppose that $\mathcal{E}$ forms a partition of the certain event Ω, i.e. the events of $\mathcal{E}$ are incompatible and $\bigvee_{i=0}^{p} E_i = \Omega$.

Otherwise, it is possible to redefine $\mathcal{E}$ as the family of ($s \leq 2^{(p+1)}$) atoms A_k (with $k = 1, ..., s$) of the events $E_0, E_1, ...E_p$, that is a partition. For instance, supposing $p = 2$ and E_1, E_2 logically independent, the atoms are $A_1 = E_1 \bigwedge E_2$, $A_2 = E_1^c \bigwedge E_2$, $A_3 = E_1 \bigwedge E_2^c$ and $A_4 = E_1^c \bigwedge E_2^c$, being $\bigwedge$ the logical product of events and E_1^c,E_2^c the contrary events of E_1, E_2.

Remark 4.2.1. One event of the partition $\mathcal{E}$ is true, but it is unknown which one, and the remaining events are false.

There are two approaches to competing risks analysis: the bivariate random variable approach and the latent variables approach.

It is possible to use the analysis package *crr*, which is available for the open source software R, to perform a competing risk analysis based on the bivariate random variable approach. Therefore, this approach is investigated in the following sections, which are mainly based on book [1], in which the latent variables approach is also shown.

4.2.2 The bivariate random variable approach

In the bivariate random variable approach, survival data can be represented as a bivariate random variable or pair (T', C).

Definition 4.2.1. T' is the time at which the event of type i occurred, with $i = 0, ..., p$.

Definition 4.2.2. The censoring variable C is defined as $C = i = 0$ if the observation is censored (i.e., if E_0 is true), as $C = i = 1$ if the observation is the occurrence of the event of interest (i.e., if E_1 is true), and $C = i$ (if E_i is true), for $1 < i \leq p$, otherwise.

Remark 4.2.2. Using the partition $\mathcal{E}$ introduced in the previous section, T' and C can also be stated as

$$T' = t_0|E_0| + t_1|E_1| + ... + t_p|E_p|, \text{ with } t_i > 0 \text{ and } i = 0, 1, ...p,$$

$$C = 0|E_0| + 1|E_1| + ... + p|E_p|,$$

where $|E_i|$ is the indicator of the event E_i (i.e. 1 if the event is true, 0 otherwise). From this formula it is clear that T' and C are stochastic variables that can assume, respectively, a real positive value and an integer value $0, 1$ or p. Furthermore T' can also be defined as the minimum of the times $T_0, ..., T_p$, T_i being the time at which the event E_i occurs. Note that T_0 is the time of the censoring event, while the other times refer to failure or relapse events.

In a survival analysis, the time under evaluation is $T = \min(T_1, ..., T_p)$, which represents the survival time of a subject until a failure of any type occurs.

It follows that, while T' is the time to any event of failure or censoring, T is exactly the time to any event of failure (not considering the time T_0 of censoring).

To describe the uncertainty on T, we use the following functions.

Definition 4.2.3. The survivor function,

$$S(t) = P(T > t),$$

also called the survival or reliability function, is the probability that an event of failure occurs by time t (i.e., the subject will survive beyond a specified time), such that $\lim_{t \to +\infty} S(t) = 0$, and $\lim_{t \to -\infty} S(t) = 1$.

Definition 4.2.4. The cumulative distribution function (cdf),

$$F(t) = P(T \leq t),$$

is the probability that an event of failure occurs before or at time t (i.e., the subject will not survive beyond a specified time), such that $\lim_{t \to +\infty} F(t) = 1$, and $\lim_{t \to -\infty} F(t) = 0$.

Definition 4.2.5. The probability density function (pdf),

$$f(t) = \frac{dF(t)}{dt} = -\frac{dS(t)}{dt},$$

is such that $f(t) > 0$, $\forall\, t > 0$, and $f(t) = 0$, $\forall\, t \leq 0$.

Remark 4.2.3. Note that

$$P(\Omega) = P(T \leq t \bigvee T > t) = P(T \leq t) + P(T > t) = 1\ ,$$

being $\bigvee$ the logical sum of events and Ω the certain event. Thus $S(t) = 1 - F(t)$.

Remark 4.2.4. $F(t) = \int_0^t f(\tau) d\tau$ and $S(t) = \int_t^\infty f(\tau) d\tau$.

The following function, which is called the hazard function, is useful to characterize the time periods in which a subject (1) improves, (2) gets worse or (3) does neither as a function of time. It is interpreted as the instantaneous failure rate for an individual who has survived to time t.

Definition 4.2.6. The hazard function is the function

$$h(t) = \frac{f(t)}{S(t)},$$

such that $h(t) > 0$, $\forall\, t > 0$, $h(t) = 0$, $\forall\, t \leq 0$ and

$$\int_0^{+\infty} h(\tau)d\tau = +\infty.$$

Definition 4.2.7. The cumulative hazard function of T is the function

$$H(t) = -\int_0^t h(\tau)d\tau.$$

Remark 4.2.5. The survivor function can be expressed in terms of the (cumulative) hazard function

$$S(t) = \exp^{-H(t)}.$$

Focusing on the representation of the stochastic vector $\Gamma = (T_1, ..., T_p)$, being $T_1, ..., T_p$ the so–called latent times (of failure or relapse), it is necessary to introduce the following functions.

Definition 4.2.8. The joint survivor function is the function

$$S(t_1, ..., t_p) = P(T_1 > t_1 \ , \ ... \ , \ T_p > t_p),$$

in which "$T_1 > t_1$, ... , $T_p > t_p$" denotes the logical product event "$T_1 > t_1$" $\bigwedge$ ... $\bigwedge$ "$T_p > t_p$".

Remark 4.2.6. Note that $S(t_1, ..., t_p) = S(t) = 1 - F(t)$ by definition of $T = \min(T_1, ..., T_p)$. Furthermore the uncertainty on the single component T_i of Γ can be represented in the same way as for T with a minor change in the terms used to designate the resulting functions, i.e., by using the prefix "marginal". For instance $S_{T_1}(T_1 > t)$ is the marginal survivor function related to the event of interest E_1.

To decompose the survivor function $S(t)$ as a sum of cause-specific survivor functions $S_i(t)$, it is necessary to state the following definitions.

Definition 4.2.9. The cumulative incidence function (CIF) of cause i

$$F_i(t) = P(T \leq t, C = i),$$

also called the subdistribution function, is the probability that an event of type i occurs at or before time t.

Definition 4.2.10. The cause-specific survivor function

$$S_i(t) = P(T > t, C = i),$$

also called the subsurvivor function, is the probability that an event of type i does not occur by time t.

Remark 4.2.7. The CIF can take values only up to $P(C = i)$ because

$$\lim_{t\to\infty} F_i(t) = P(C = i).$$

Therefore, $F_i(t)$ is not a proper distribution, hence the term 'subdistribution'. Furthermore note that

$$F_i(t) + S_i(t) = P(C = i).$$

Definition 4.2.11. The subhazard

$$\tilde{h}_i(t) = \lim_{\delta t\to 0} \left\{ \frac{P(t < T \leq t + \delta t, C = i|T > t)}{\delta t} \right\} = \frac{f_i(t)}{S(t)}$$

is the instantaneous failure rate in the competing risks situation.

Remark 4.2.8. The hazard function can be found by summing over all subhazards, as shown:

$$h(t) = \sum_{i=1}^{p} \tilde{h}_i(t).$$

Definition 4.2.12. The hazard of the CIF, also called the subdistribution hazard for cause i, is the function

$$\gamma_i(t) = \lim_{\delta t\to 0} \left\{ \frac{P\big(t < T \leq t + \delta t, C = i|T > t \text{ or } (T \leq t \text{ and } C \neq i)\big)}{\delta t} \right\}.$$

Definition 4.2.13. The cause-specific hazard is

$$h_i(t) = -\frac{\partial log(S_i(t))}{\partial t} = \frac{f_i(t)}{S_i(t)}.$$

Remark 4.2.9. Note the difference between $h_i(t)$ and $\gamma_i(t)$. In fact it holds that

$$1 - F_i(t) = S_i(t) + P(C \neq i),$$

being $F_i(t) + S_i(t) + P(C \neq i) = P(\Omega) = 1$.

Finally, the survival function $S(t)$ can be computed as the sum of the cause-specific survivor functions, that is

$$S(t) = P(T > t) = \sum_{i=1}^{p} P(T > t, C = i) = \sum_{i=1}^{p} S_i(t),$$

or, alternatively, as

$$S(t) = 1 - F(t) = 1 - \sum_{i=1}^{p} P(T \leq t, C = i) = 1 - \sum_{i=1}^{p} F_i(t).$$

Remark 4.2.10. Note that $F_i(t)$ can be evaluated from $\gamma_i(t)$, being

$$F_i(t) = 1 - \exp\left(\int_0^t \gamma_i(\tau)d\tau\right)$$

and $\int_0^t \gamma_i(\tau)d\tau = H_i(t)$ the cumulative hazard of the CIF, also called cumulative subdistribution hazard.

4.2.3 Modeling the hazard of the CIF

In evaluating $F_i(t)$, it is possible to adopt a semiparametric proportional model for the corresponding subdistribution hazard [1], that is

$$\gamma_i(t) = h_{0i}(t)e^{\beta_i x},$$

where $h_{i0}(t)$ is the so–called baseline hazard, x is the vector of the variables (also called covariates) under evaluation (i.e., predictive or risk factors that influence the survival of a subject) and β_i is the vector of the coefficients that represent the weights of the above factors.

This model is semi-parametric because, while the baseline hazard can take any form, the covariates enter the model linearly. It is said to be proportional with respect

to the assumption that, given two observations that differ in their x-values, say x' and x'' (corresponding to two subjects under evaluation), the resulting hazard ratio for these two observations is independent of time t, i.e., it is equal to $\exp(\beta_i x')/\exp(\beta_i x'')$.

Estimation for this model follows the partial likelihood approach used in a standard Cox model [2]. Thus, each component of β_i, denoted by β_{ik}, with $k = 1, \ldots, dim(x)$, is obtained solving the maximization of a score statistic that is the log of the partial likelihood written for each component x_k of x. See [1] for more details.

Finally, given a certain value of the covariates, we can evaluate

$$F_i(t) \simeq \hat{F}_i(t) = 1 - \exp(-\hat{H}_i(t)),$$

where $\hat{H}(t)$ is an estimate of the cumulative hazard of the CIF obtained, for instance, using a Breslow-type estimator. See [1] for more details.

4.3 Algorithms

Using the open source statistical software R available at www.R-project.org and distributed under the GNU (www.gnu.org) General Public License, it is possible to conduct a competing risks analysis [1] based on the bivariate random approach and a semiparametric proportional modeling of the subdistribution hazards, i.e. $\gamma_i(t) = h_{0i}(t)e^{\beta_i x}$. We refer mainly to [3].

The main steps in this analysis are:

- Download and install the package *crr*.
- Use the library *aod* and the source code *crr-addson*[3].
- Load the (cleaned) input data set containing only the subjects included in the analysis.
- Define and set all the variables, including the censoring variable C and the time T' to any event (failure or censoring).
- Build the so–called design matrix including all the covariates representing the factors under evaluation; this matrix is an extended version of the input data set useful for the elaboration.
- Build the first candidate regression model for $\gamma_i(t)$ using the design matrix.
- Build the second candidate model using a submatrix of the design matrix, including only the significative covariates resulting from the first model.
- Build other candidate models as refinements of the second model, adding just one less significative covariate.

- Select a candidate model using a ranking score criterion (for instance the Bayesian information criterion).
- Refine the candidate models, adding covariates relevant to the particularly study.
- Select the best model.
- Test the best model.
- Compute the cumulative incidence functions $F_i(t)$ of interest.

In more detail, the function *crr* is used to estimate $\hat{\beta}_i$, the parameters of $\gamma_i(t)$, with a confidence interval $\hat{\beta}_{ik} \pm \hat{\sigma}_{\beta_{ik}} z_{(1-\alpha/2)}$, where z_α is the α quantile of the standard normal distribution.

As output of the *crr* script, we obtain also:

- the p-value for each test of significance, which compares the hypothesis $H_0 =$"$\beta_{ik} = 0$" versus the alternative $H_a =$"$\beta_{ik} \neq 0$", using as a test statistics the ratio $\hat{\beta}_{ik}/\hat{\sigma}_{ik}$ which follows the standard normal distribution [1][3];
- a matrix of the so–called Schoenfeld residuals, that can be drawn to visually test the proportionality assumption of the $\gamma_i(t)$ [1][3].

4.4 Applications

4.4.1 Ultrasonographic markers of vascular risk in patients with asymptomatic carotid stenosis

In the present study we aimed at investigating reliable markers of increased risk of stroke ipsilateral to severe internal carotid artery (ICA) stenosis, taking into account also contralateral stroke, transient ischemic attack (TIA), myocardial infarction (MI) and vascular death, by integrating ultrasonographic data concerning the carotid wall and plaque characteristics. For this purpose, we conducted a prospective investigation of asymptomatic subjects who were being treated with the best available medical therapy for each vascular risk condition, paying particular attention to adherence to treatment [4] [5].

Matherials and methods

During a four-year period (January 2005 to December 2008), we diagnosed ICA stenosis of 60% or more in 819 asymptomatic subjects among patients referred by their primary care physicians or by other specialists to receive an ultrasound screening for carotid atherosclerotic disease according to international guidelines [6].

158 patients who considered carotid endarterectomy (CEA) were referred to our Vascular Surgery Department and were excluded from the analyses. The remaining 661 subjects underwent neurological and cardiological examination.

Blood analysis and clinical history, with particular attention to the major vascular risk factors (hypertension, diabetes, smoking and hyperlipidemia) were also obtained.

The diagnostic criteria and the pharmacological treatment of vascular risk factors, in addition to behavioral recommendations (smoking cessation, regular physical exercise, weight control) were planned in accordance with international guidelines [4] [5] [7].

Patients were followed-up by telephone interviews every 3 months and re-evaluated clinically every 6 months.

End points were stroke ipsilateral and contralateral to ICA stenosis, TIA, MI and vascular death. Stroke or TIA diagnoses had to be confirmed by brain computed tomography or magnetic resonance imaging.

The neurologists who defined primary events on the basis of clinical charts were blind to the neurosonological findings.

The study was approved by the local ethics committees. Each subject gave informed consent to participation in the study.

Ultrasonographic examination .

Carotid arteries were assessed and defined by continuous wave Doppler and Color flowB-mode Doppler ultrasound (Philips iU22, Bothell, USA). The degree of carotid stenosis was established according to combined criteria considering blood flow velocities and morphological characteristics.

Steno-occlusive lesions in the common carotid arteries (CCAs), carotid bulb and ICA were assessed and defined in three categories: $60-70\%$, $71-90\%$ and $91-99\%$ according to validated criteria [8] [9].

To measure intima-media thickness (IMT), semiautomatic software was used to improve measurement reliability and reproducibility [10]. Because of the ICA stenosis, we considered CCA IMT [5].

Stenotic lesions were characterized on the basis of their echogenicity and surface. The assessment of plaque echolucency was based on the modified version of the Gray-Weale classification [11]. Plaque surface was defined as (1) smooth and regular; (2) mildly irregular; or (3) ulcerated [12] [13]. Progression of plaque was defined as any change to a higher category of ICA stenosis from baseline at the 6 or 12-month follow-up evaluations.

All ultrasonographic assessments were performed by sonographers blinded to clinical information.

Competing risk analysis: overview .

The statistical analysis aimed to identify which factors could predict ipsilateral stroke events in patients with carotid stenosis. Since no vascular deaths were observed during follow-up, only ipsilateral stroke, contralateral stroke, TIA and MI defined the typical scenario for competing risks. In this study, we considered ipsilateral stroke the event of interest. Contralateral stroke, TIA and MI were defined as competing events because their occurrence brings relevant clinical changes (i.e. diagnostic and therapeutic changes) that can modify the natural history of ICA stenosis. In particular, patients who had TIA or stroke ipsilateral to carotid stenosis were referred to our Vascular Surgery Department.

Competing risk analysis, based on the bivariate random variable approach and on modeling of the hazards of subdistribution using a semiparametric proportional model, was applied using the open source statistical software R (32bit) version 2.13.1. [1][3].

Competing risk analysis: mathematical details .

From a probabilistic point of view, the logical events under evaluation were the following:

E_1 =“Ipsilateral stroke occurs during the follow-up of the subject”;

E_2 =“Contralateral stroke occurs during the follow-up of the subject”;

E_3 =“TIA occurs during the follow-up of the subject”;

E_4 =“MIA occurs during the follow-up of the subject”;

E_0 =“No event occurs during the follow-up of the subject”.

The family $\mathcal{E} = \{E_0, E_1, E_2, E_3, E_4\}$ is supposed to be a partition of the certain event, i.e. just one event is true, but it is unknown which, and the others are false.

Therefore, the competing risk situation was represented as the stochastic pair (T', C), with:

$$T' = t_0|E_0| + t_1|E_1| + t_2|E_2| + t_3|E_3| + t_4|E_4|, \text{ with } t_i > 0 \text{ and } i = 0, 1, ...4,$$

$$C = 0|E_0| + 1|E_1| + 2|E_2| + 3|E_3| + 4|E_4|.$$

The survival time of a subject was defined as the stochastic variable

$$T = \min(T_1, T_2, T_3, T_4),$$

being:

T_1 =time to ipsilateral stroke (months);

T_2 =time to contralateral stroke (months);

T_3 =time to TIA (months);

T_4 =time to MIA (months).

In particular, we were interested in the estimation of the cumulative incidence function of ispilateral stroke, that is

$$F_1(t) = P(T \leq t, C = 1),$$

which can be evaluated from the corresponding subdistribution hazard $\gamma_1(t)$ by means of the relation

$$F_1(t) = 1 - \exp\left(\int_0^t \gamma_1(\tau)d\tau\right).$$

Based on the foregoing, the analysis of the available data set aimed to build a semiparametric proportional model for $\gamma_1(t)$, that is

$$\gamma_1(t) = h_{01}(t)e^{\beta_{i1}x_1+\beta_{i2}x_2+\ldots+\beta_{im}x_m},$$

by looking for a subset of covariates that globally characterize the subject and their corresponding weights, the coefficients β_{ik}, that were most significative for the subdistribution hazard modeling.

A baseline coding was performed for those factors, such as the gender of the subject, that assumed a finite number J of values. If a factor was not binary, i.e., $J > 2$, it was split into $J - 1$ binary indicators.

For this reason the covariates under evaluation were 23 in number; they included

x_1 = "Gender of the subject (M or F)";

x_2 = "Age of the subject (years)";

x_3 = "Indicator of smoking habit (1 if the subject is a smoker, 0 otherwise)";

x_4 = "Indicator of diabetes (1 or 0)";

x_5 = "Indicator of dyslipidemia (1 or 0)";

x_6 = "Indicator of hypertension (1 or 0)";

x_7 = "Indicator of cardiopathies (1 or 0)";

x_8 = "Indicator of concurrent antihypertensive therapy(1 or 0)";

x_9 = "Indicator of concurrent antidiabetic therapy (1 or 0)";

x_{10} = "Indicator of concurrent statin therapy (1 or 0)";

x_{11} = "Indicator of concurrent antiplatelet therapy (1 or 0)";

x_{12} = "Indicator of ispilateral IMT (1 or 0)";

x_{13} = "Indicator of contralateral IMT (1 or 0)";

x_{14} = "Indicator of carotid stenosis at level $[71 - 90\%)$ (1 or 0)";

x_{15} = "Indicator of carotid stenosis at level $[90 - 99\%)$ (1 or 0)";

x_{16} = "Indicator of hyper-echogenicity of the lesions (1 or 0)";

x_{17} = "Indicator of mixed-echogenicity of the lesions (1 or 0)";

x_{18} = "Indicator of hypo-echogenicity of the lesions (1 or 0)";

x_{19} = "Indicator of ulcerated plaque surface (1 or 0)";

x_{20} = "Indicator of irregular plaque surface (1 or 0)";

x_{21} = "Indicator of plaque progression (1 or 0)";

x_{22} = "Mean CCA IMT (millimeters)".

Table 4.1 shows all of these factors and some useful descriptive statistics of the data under evaluation. Note that, using the baseline coding, some relevant findings are implicitly represented. For instance the observation of a regular plaque surface in a subject was represented by giving zero values to both x_{19} and x_{20}.

Patients (n=621)						
Demographic characteristics			*Ultrasound findings*			
Follow-up, months (median, min-max)	27	6-68		ipsilateral	1.04	0.25
Age, years (mean, SD)	72.2	8.9	IMT (mean, SD)	contralateral	1.01	0.24
Sex (n,% of males)	348	56%		mean	1.03	0.24
Risk factors and therapy (n,%)				60-70%	414	67%
Smoking	161	26%	Carotid Stenosis (n,%)	71-90%	173	28%
Diabetes	155	25%		91-99%	34	5%
Dislypidemia	348	56%		hyper	382	62%
Hypertension	416	67%	Echogenicity (n,%)	mixed	178	29%
Cardiopathies	124	20%		hypo	61	10%
Antihypertensives	391	63%		regular	463	75%
Antidiabetics	155	25%	Plaque Surface (n,%)	irregular	130	21%
Statins	279	45%		ulcerated	28	5%
Antiplatelets	329	53%	Plaque progression (n,%)		137	22%

Hyper = Hyperechogenic plaque ; Hypo = Hypoechogenic plaque.

Table 4.1. Baseline characteristics and frequency of occurrence of vascular risk factors in the cohort study.

Competing risk analysis: implementation details .

The open source statistical software R, version 2.13.1 (32bit), was used to perform the following steps that follow the algorithm in section 4.3 and the related references.

Assuming that the data file "stenosi.csv" is located on the current working directory, it can be read as follows:

> *data=read.csv("stenosi.csv")*

with > being the prompt of the R console and *data* the new matrix variable containing the entire input data set.

If the data file is correctly read, the first records can be visualized:

> *head(data)*

The mathematical variable T, which represents the time to ipsilateral stroke in this particular application, was modeled in R using the variable *time* and the mathematical censoring variable C using the variable *status*.

A variable for each component of the vector of predictive covariates x, was also defined. That is, demographic characteristics (age and sex), risk factors (smoking, diabetes, etc.), concomitant therapies (antidiabetic, antihypertensives, etc.) and specific ultrasound findings (carotid stenosis, type of plaque surface, etc.) were each represented by at least one ad–hoc variable.

The baseline codification was applied where necessary. For instance, the factor (carotid) Stenosis was coded by running the following commands:

> *Stenosis=data$Stenosis*

> *Stenosis=factor(Stenosis)*

> *Stenosis=factor2ind(Stenosis,"1")*

given that the matrix *data* contained a column called Stenosis and values 1, 2 or 3 to indicate, respectively, carotid stenosis at the $60-70\%$ level, the $70-90\%$ level or the $90-99\%$ level. To test the effect of the baseline coding onto the factor Stenosis, we ran the command

> *head(Stenosis)*

and checked the first records of the matrix Stenosis.

The so called design matrix was obtained:

> *designMatrix=cbind(Age,Sex,...,PlaqueProgression)*

The parameters passed in the *cbind* function call, that is

Age, Sex,...,PlaqueProgression,

were exactly the implementation, in R, of the covariates x_1,x_2,...,x_{22} under evaluation for the modeling of the hazard of the subdistribution $\gamma_1(t)$.

The first regression model was obtained considering all the factors, i.e. using the design matrix:

> *model1=crr(time, status, designMatrix)*

The covariates of this model were classified as significative, quasi-significative or not significative according to the p-values of the resulting estimated weights. A Wald test was used to obtain the overall p-values for factors, such as Stenosis, that were baseline coded.

Thus, a second candidate model was built considering only those covariates that were significative (p-value$<$ 0.02) using the *crr* command and a submatrix of the design matrix.

Other candidate models were obtained by adding to the second model just one quasi–significative–covariate.

The evaluation of the candidate models was performed on the basis of a well–defined ranking score called BIC (meaning Bayesian information criteria):

$>$ *modsel.crr(mod1, mod2,...)*

Because the magnitude of any information criterion is not relevant, differences with respect to the smallest value are usually computed. Interpretation is then based on the so called *general rules of thumb*: values of $\Delta BIC = BIC_i - \min(BIC) > 10$ provide very strong evidence against the i-th model, but values of $0 < \Delta BIC_i < 2$ suggest that the i-th model has substantial support and should receive consideration in making inferences.

Once a model was selected, we built other candidate models as refinements of the selected model, taking into consideration also the meaning of the covariates under evaluation and the aims of the study.

The best model was selected, still using the BIC criteria.

The output of the *crr* function also provides a matrix of Schoenfeld residuals. If the proportional hazard subdistribution assumption holds the residuals should have locally a constant mean across time, that can be checked visually using ad–hoc scatterplots.

Finally, the predicted cumulative incidence function for ipsilateral stroke was computed:

> *CIF1=predict(bestModel, newSubject)*

with *newSubject* a vector of values that characterized a hypothetical subject according to the covariates of the best model.

Results

A preliminary analysis was performed to compare 661 subjects with the 158 subjects that were excluded from the study because of their decision to be referred to the Vascular Surgery Department. No differences in demographic characteristics, vascular risk profile or ultrasound findings were found. At baseline, subjects with modifiable vascular risk factors were prescribed the corresponding treatment, although a large variability in compliance was observed.

During follow-up, regular therapy assumption was observed in 88% of hypertensive patients, 99% of diabetic patients and 68% of patients with dyslipidemia. Subjects with scarce or lack of compliance were considered as non–users. The median follow-up period was 27 months (min=6, max=68). Table 1 shows the study participants' baseline characteristics.

During follow-up, 99 subjects (15.9%) experienced a vascular event: 72 (72.7%) cerebrovascular events (39 strokes, 27 TIA ipsilateral to the most stenotic carotid artery and 6 contralateral strokes) and 27 (27.3%) MI.

Figures 4.1 and 4.2 summarize the findings of the Competing Risk Regression analysis, a simple model that uses only carotid stenosis indicators, and the finding of the best model according to the available data; the latter uses the most significative ultrasound findings, that is, the plaque surface indicators, the plaque progression indicator, the carotid stenosis indicators and the mean CCA IMT. Both models were successfully tested using the diagnostic test of Schoenfield residuals (see Figures 4.3 and 4.4). To sum up, given a subject with a median IMT of 1 millimeter, evidence of plaque progression and plaque irregular surface, the best model well shows the relevance of a carotid stenosis (compare Figure 4.2 with Figure 4.1).

4.5 Conclusions and future work

Competing risk analysis can support physicians in giving prognostic information to patients. The probability of surviving after a given time, i.e., the survival probability function, can be decomposed as a sum of cause-specific survivor functions. In this way, it is possible to understand the role of each competing risk in the possible failure. The use of tools such as those included in the open source statistical software R may seem sufficient, but this is so only in appearance. In fact, in conducting a risk analysis of this type, it is not only necessary to have programming skills, but more importantly, it is necessary to possess a solid mathematical background so as to be able to consider all the relevant functions and assumptions made at the theoretical level. In this case, knowledge of the latent variable approach and of semiparametric proportional modeling of cause-specific hazard functions is needed.

The future aim related to this work is the application of competing risk analysis to hospitalization data sets resulting from the standard information system of a hospital, for instance, to investigate the time to rehospitalization.

The final goal of the work described in this section is the embedding of this type of survival analysis into a decision support system for the clinical governance of a hospital, such as the Business Simulator for Health Care (BuS-4H) system, which the Author, who is the creator of the corresponding research project, intends to develop in collaboration with an Italian company called “SiliconDev srl” and several clinical, mathematics and engineering departments of the University of Rome “Sapienza”, the University of Rome “Tor Vergata”, the “Fatebenefratelli” Association for Biomedical Research (AFaR) and the “San Giovanni Calibita Fatebenefratelli Isola Tiberina” Hospital in Rome.

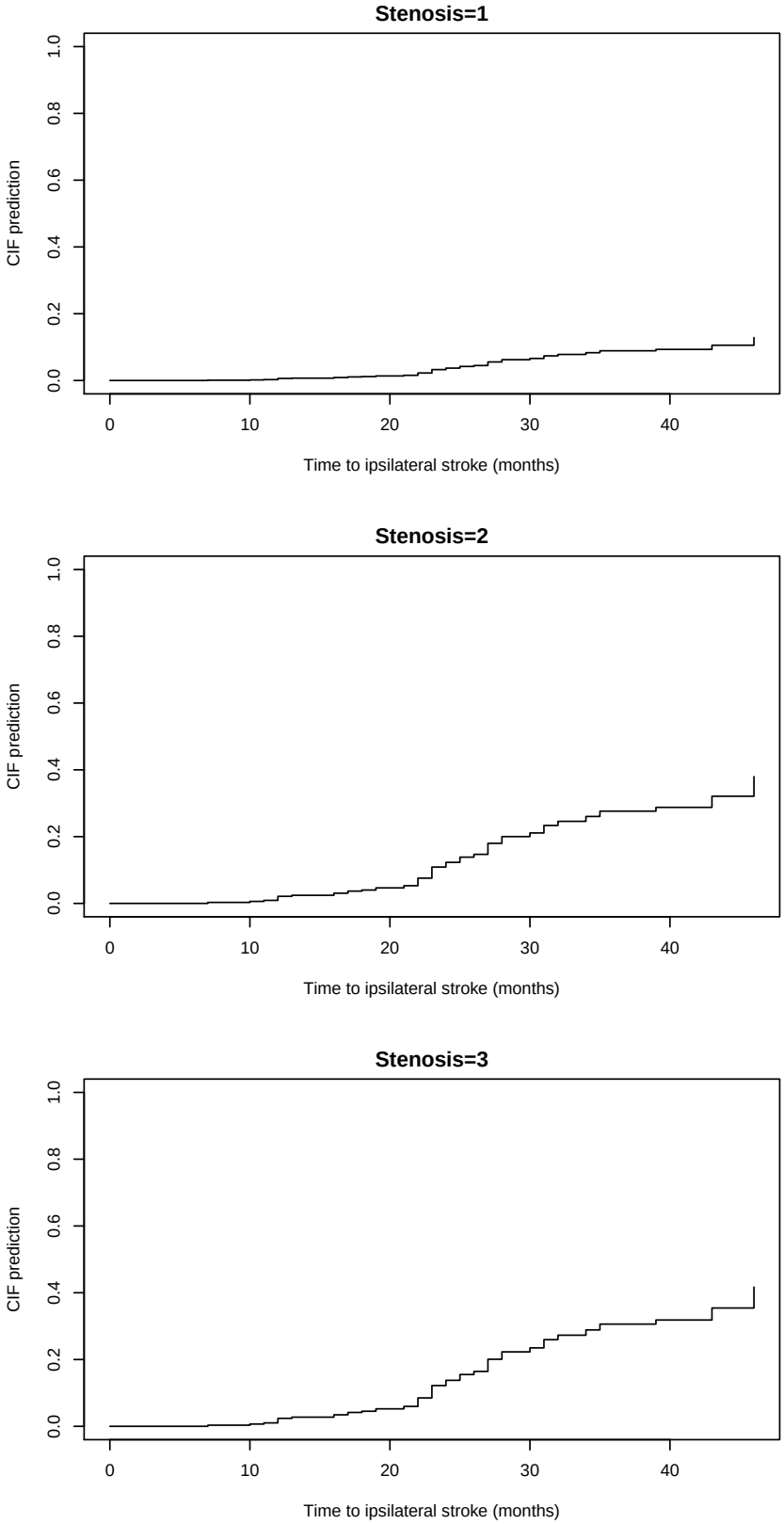


Figure 4.1. $F_1(t)$ obtained obtained using a model with one categorial covariate representing the level of (carotid) stenosis. The increase in the CIF, which varies with the level of stenosis, corresponds to an increase in the hazard of ispilateral stroke.

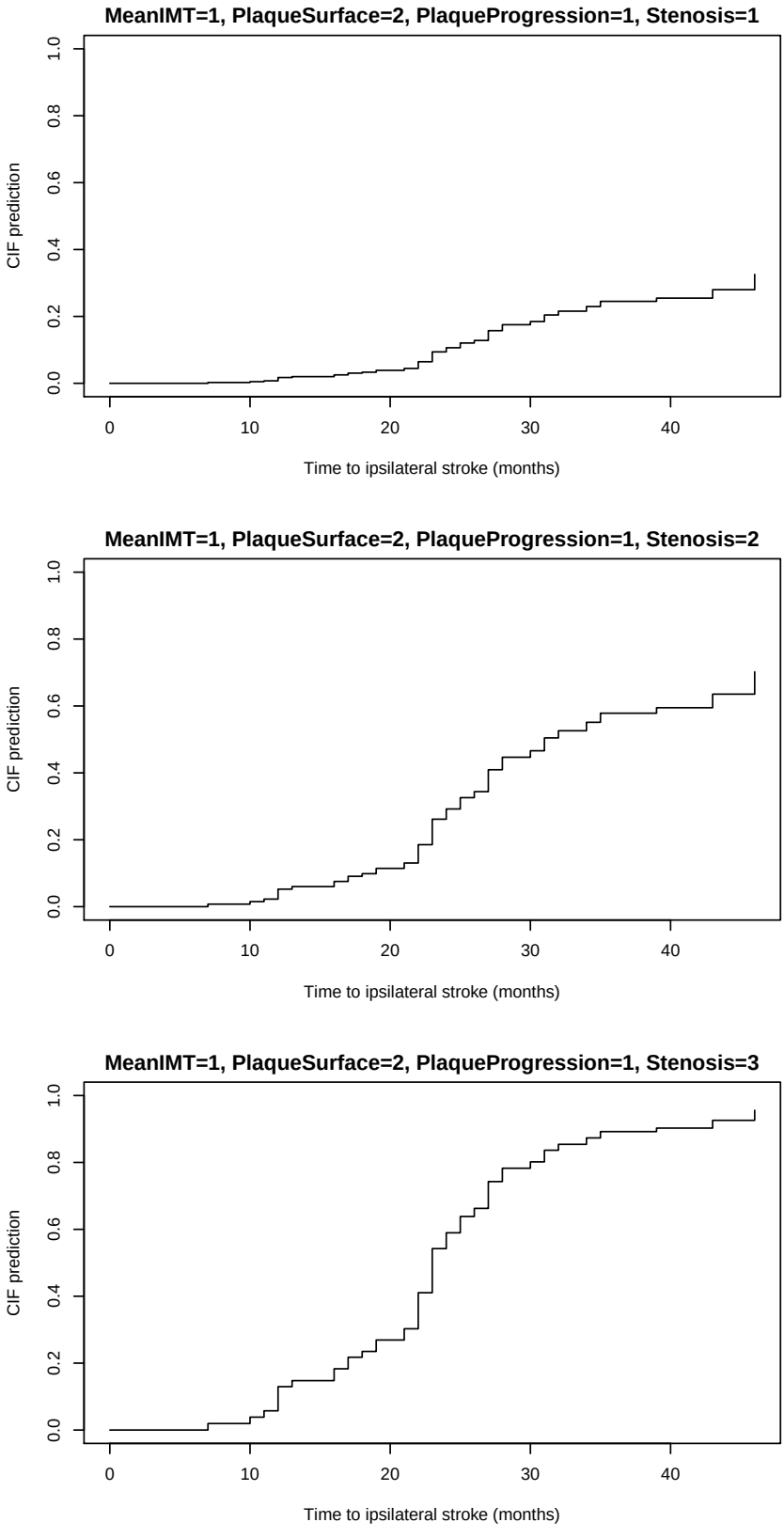


Figure 4.2. $F_1(t)$ obtained using the best model, which uses as covariates the mean of IMT (millimeters), the plaque surface, the plaque progression and the level of carotid stenosis. The role of the level of carotid stenosis is now more evident.

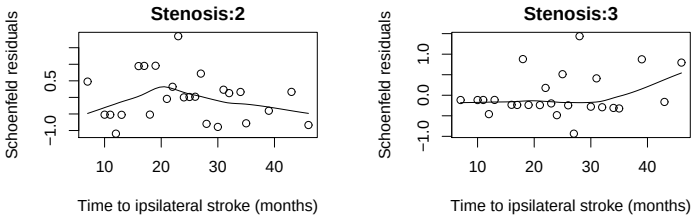


Figure 4.3. Diagnostic of a simple regression model, which uses only the indicators of carotid stenosis.

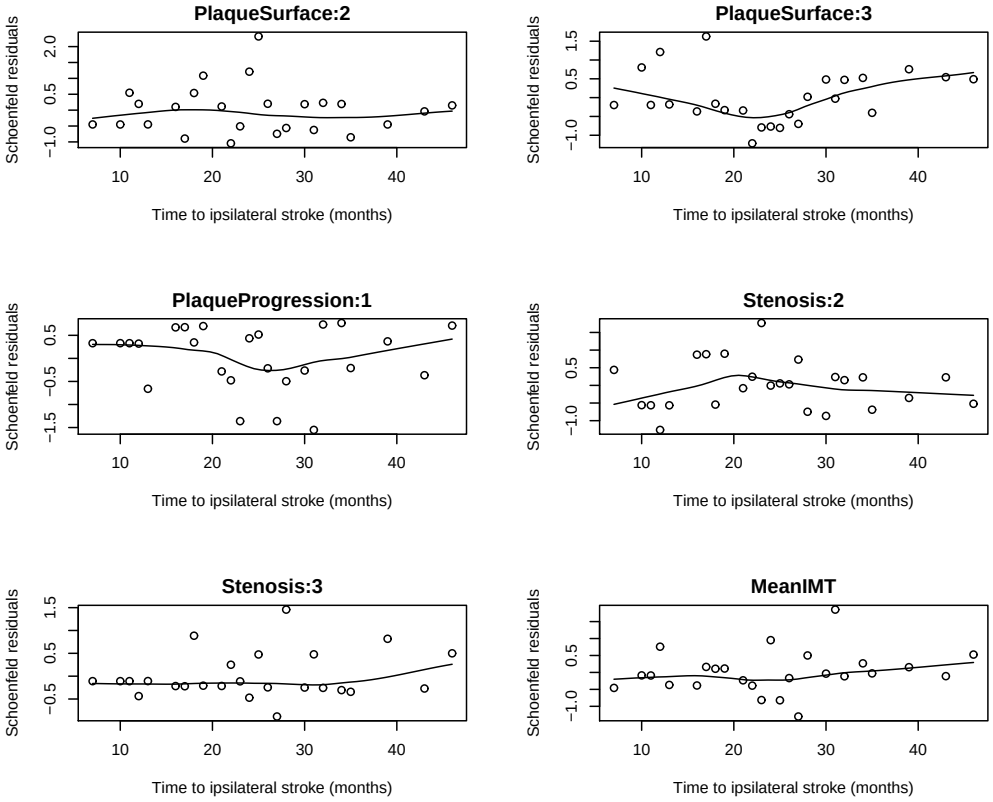


Figure 4.4. Diagnostic graphs of the best model, which uses as covariates the plaque surface indicators, the plaque progression indicators, the carotid stenosis indicators and the mean of IMT (millimeters).

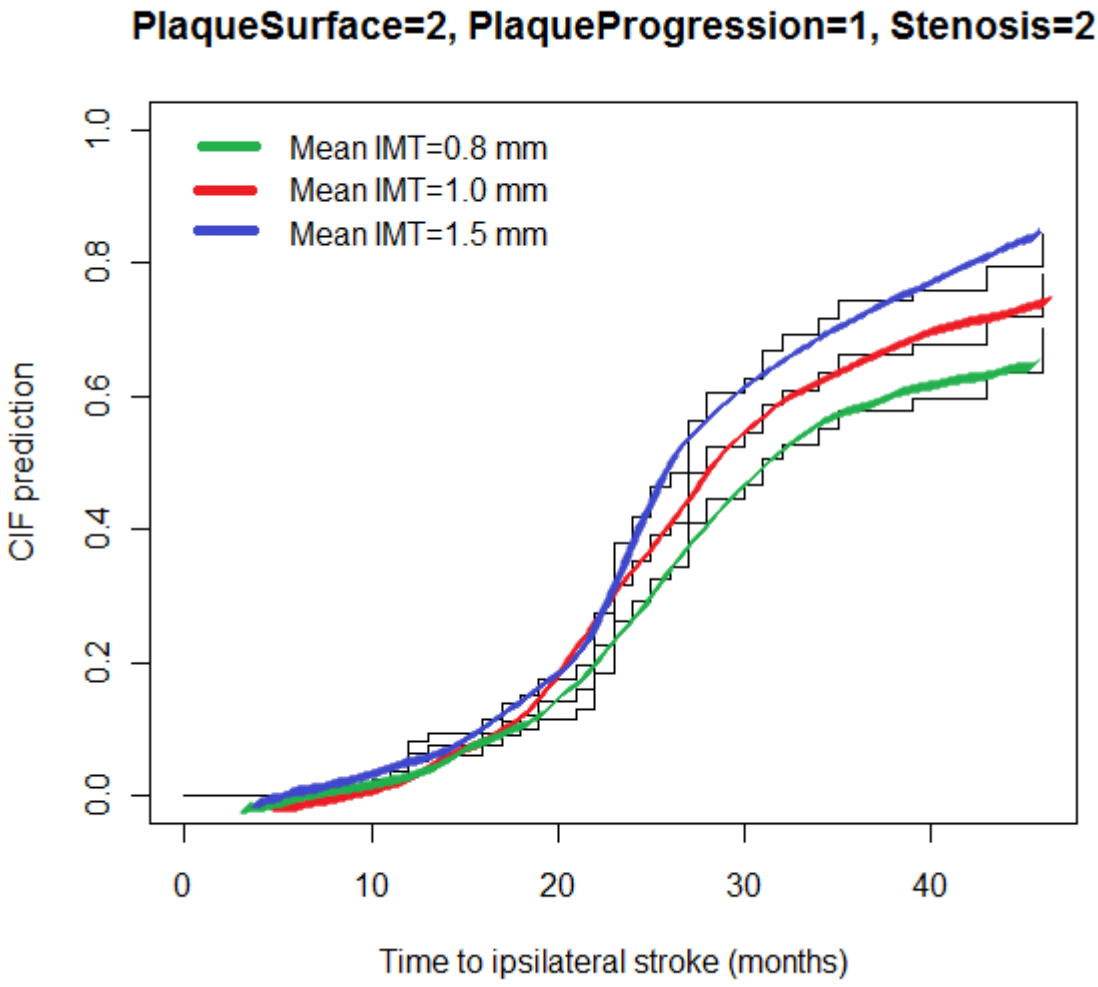


Figure 4.5. $F_1(t)$ obtained using the best model, which uses as covariates the mean of IMT (millimeters), the plaque surface, the plaque progression and the level of carotid stenosis. The role of the mean of IMT is qualitatively shown (below and over the median value, i.e., 1 mm).

Bibliography

[1] Pintilie M, *Competing Risks: A practical Perspective*, Wiley, (2006).

[2] Cox D. R., *Regression models and life-tables*, Journal of the Royal Statistical Society, Series B, 34, *187-220*, (1972).

[3] Scrucca L, Santucci A, Aversa F, *Regression modeling of competing risk using R: an in depth guide for clinicians*, Bone Marrow Transplantation, *45*, (2010), 1388-1395.

[4] Goldstein LB, Adams R, Becker K et al, *Primary prevention of ischemic stroke: A statement for healthcare professionals from the Stroke Council of the American Heart Association*, Stroke *32* (2001), 280-299.

[5] Goldstein LB, Bushnell CD, Adams RJ et al, *Guidelines for the primary prevention of stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association*, Stroke *42* (2011), 517-584.

[6] Qureshi AI, Alexandrov AV, Tegeler CH, et al, *Guidelines for screening of extracranial carotid artery disease: a statement for healthcare professionals from the multidisciplinary practice guidelines committee of the American Society of Neuroimaging; cosponsored by the Society of Vascular and Interventional Neurology*, J Neuroimaging *17* (2007), 19-47.

[7] Leys D, Kwiecinski H, Bogousslavsky J et al, *EUSI Executive Committee; EUSI Writing Committee. Prevention. European Stroke Initiative*, Cerebrovasc Dis *17(Suppl. 2)* (2004), 15-29.

[8] Gortler M, Niethammer R, Widder B, *Differentiating subtotal carotid artery stenoses from occlusion by colour-coded duplex sonography*, J Neurol *241* (1994), 301-305.

[9] de Bray JM, Baud JM, Delanoy P et al, *Reproducibility in ultrasonic characterization of carotid plaques*, Cerebrovasc Dis *8* (1998), 273277.

[10] Touboul PJ, Hennerici MG, Meairs S et al, *Mannheim carotid intimamedia thickness consensus (20042006). An update on behalf of the Advisory Board of the 3rd and 4th Watching the Risk Symposium, 13th and 15th European Stroke Conferences, Mannheim, Germany, 2004, and Brussels, Belgium, 2006*, Cerebrovasc Dis *23* (2007), 7580.

[11] Gray-Weale AC, Graham JC, Burnett JR, et al, *Carotid artery atheroma: comparison of preoperative B-mode ultrasound appearance with carotid endarterectomy specimen pathology*, J Cardiovasc Surg *29* (1988), 676-681.

[12] Mathiesen EB, Bnaa KH, Joakimsen O, *Echolucent plaques are associated with high risk of ischemic cerebrovascular events in carotid stenosis: the Troms study*, Circulation *103* (2001), 2171-2175.

[13] Kern R, Szabo K, Hennerici M, Meairs S, *Characterization of carotid artery plaques using real-time compound B-mode ultrasound*, Circulation *35* (2004), 870-875.