A case-based prognosis system for disabilities of neurological origin

TESI DOCTORAL

Programa de Doctorat en Informàtica

Departament de Ciències de la Computació

Universitat Autònoma de Barcelona

Bellaterra, 2015

Laia Subirats

Directors: Enric Plaza
Luigi Ceccaroni

Tutor: Jordi González

to my parents

Acknowledgements

I had been harbouring the idea of doing a joint-industrial PhD for some years, but it remained vague until I arrived to the eHealth group of Barcelona Digital Technology Centre, now Eurecat, and the group's head, Felip Miralles, proposed me to work with Institut Guttmann, a leading hospital in the medical treatment, surgery and full rehabilitation of neurological diseases. I had heard of Institut Guttmann when I was little, because the best friend of my grandfather lost his hands during the Spanish civil war. And I thought it would be nice to combine computer science and medicine.

I would like to thank especially my directors, Enric Plaza and Luigi Ceccaroni, both role-models for me, for all their time, support and invaluable comments on this Thesis. Enric for his patience and for being a source of wisdom. Luigi for being my "director at the industry" and helping me to use the work in the projects Circles of Health (TSI-040200-2009-73) and Rehabilita (CEN-2009-1043) for my PhD.

I would also like to thank all members of Eurecat and Institut Guttmann for their kindness and support. All of you have contributed to make these six years at Eurecat wonderful and unforgettable. I am also very grateful to "la Caixa" Foundation, who gave me financial support for my Master's (which was the former PhD courses); and to Generalitat de Catalunya for the Beatriu Pinós pre-doctoral scholarship.

I do not want to forget my family and closest friends. I am very grateful to my parents for their love and patience during my studies. Finally, special thanks to David for sharing our lives, both the bad moments and the happy ones, which have been many more.

Laia Subirats

Barcelona, September 2015

Abstract

In this thesis we focus on disabilities of neurological origin. According to the WHO, the diseases of neurological origin are one of the leading causes of burden of disease in the world. Two populations of patients are used in this thesis: people who suffer from spinal cord injury (SCI) and people who suffer from acquired brain injury (ABI). Both datasets have been kindly provided by Institut Guttmann. More specifically, we work in four prediction domains for patients with neurological disabilities: emotional and urination functions of people with SCI, and emotional and executive functions of people with ABI.

In the rehabilitation of people who suffer from neurological diseases, there are two basic stages: intra-hospital and extra-hospital. In the intra-hospital (acute) phase, patients who suffered from a traumatic or nontraumatic injury stay in hospital undergoing rehabilitation. After typically a few months in hospital, they return home and the extra-hospital phase starts. Thereafter they go once a year to the rehabilitation hospital for a periodic comprehensive evaluation (PCE), when they are administered several questionnaires (depending on the disease they are suffering from) with the aim of measuring functioning independence, psychological and social variables. This study is focused on the extra-hospital or chronic phase.

This thesis contributes to the improvement of the state of the art of healthcare systems by dealing with: lack of interoperability, limited capability for temporal or populational analysis, limited capability for prognosis, and limited capability for integrating time-dependent information. In particular, regarding interoperability, this thesis proposes a new automatic translation system to international standards promoted by the WHO and a new ontology.

Regarding the ability for temporal or population analysis, this thesis provides a monitoring system with novel visualization techniques that allows a better understanding of the evolution of clinical attributes in individuals and populations. More specifically, it provides a graphical representation of states and their evolution that permits easily comparing the characteristics of an individual with those of the whole population.

With respect to prognosis, in this thesis we have used case-based reasoning (CBR), an artificial intelligence technique based on solving new situations by learning from past similar situations already solved. Unlike our monitoring system, in CBR-based prognosis systems the target patient is related to a small set of similar patients (instead of the whole population of patients). This thesis designs and develops a CAse-based Prognosis using Temporal Abstractions (CAPTA), a prognosis system with a new case-based reasoning method with promising results.

Finally, regarding time, CAPTA integrates the temporal nature of clinical attributes through temporal abstractions. CAPTA is applied and experimentally evaluated in four prediction domains for patients with neurological disabilities.

Abstract (in Catalan)

En aquesta tesi ens centrem en les discapacitats d'origen neurològic. Segons l'OMS, les malalties d'origen neurològic són una de les principals causes de càrrega de malaltia al món. En aquesta tesi es fan servir dues poblacions de pacients: les persones que pateixen lesions medul·lars (LM) i les persones que pateixen dany cerebral adquirit (DCA). Ambdós conjunts de dades han estat proporcionats amablement per l'Institut Guttmann. Més específicament, treballem en quatre dominis de predicció per als pacients amb discapacitats neurològiques: funcions emocionals i urinàries de persones amb LM, i funcions emocionals i executives de persones amb DCA.

En la rehabilitació de les persones que pateixen de malalties neurològiques, hi ha dues etapes bàsiques: intrahospitalària i extrahospitalària. En la fase intrahospitalària (aguda), els pacients que han patit una lesió traumàtica o no traumàtica resten a l'hospital en procés de rehabilitació. Generalment, després de romandre uns pocs mesos a l'hospital, tornen a casa i comença la fase extrahospitalària. A partir d'aquest moment, els pacients van un cop l'any a l'hospital de rehabilitació per fer una avaluació integral periòdica, i és llavors quan han de complimentar diversos qüestionaris (depenent de la malaltia que pateixen) amb l'objectiu de mesurar la independència funcional i les variables psicològiques i socials. Aquest estudi es centra en la fase extrahospitalària o fase crònica.

Aquesta tesi contribueix a la millora de l'estat de l'art dels sistemes de salut en: la falta d'interoperabilitat, la capacitat limitada per a l'anàlisi temporal o poblacional, la capacitat limitada per al pronòstic, i la capacitat limitada per integrar la informació depenent del temps. En particular, en relació amb la interoperabilitat, aquesta tesi proposa un nou sistema de traducció automàtica a les normes internacionals promogudes per l'OMS i una nova ontologia.

En quant a la capacitat per a l'anàlisi temporal o de la població, aquesta tesi proporciona un sistema de seguiment amb tècniques de visualització noves, que permet una millor comprensió de l'evolució d'atributs clínics en individus i poblacions. Més específicament, ens dóna una representació gràfica dels estats i la seva evolució, que permet comparar fàcilment les característiques d'un individu amb les de tota la població.

Pel que fa al pronòstic, en aquesta tesi s'ha utilitzat el raonament basat en casos (RBC), una tècnica d'intel·ligència artificial que es fonamenta en la solució de situacions noves, aprenent de situacions similars anteriors ja resoltes. A diferència del nostre sistema de seguiment, en els sistemes de pronòstic basats en RBC el pacient objectiu està relacionat amb un petit grup de pacients similars (en lloc de tota la població dels pacients). Aquesta tesi dissenya i desenvolupa un pronòstic basat en casos utilitzant abstraccions temporals (CAPTA), un sistema de pronòstic amb un nou mètode de raonament basat en casos amb resultats prometedors.

Finalment, pel que fa al temps, CAPTA integra la naturalesa temporal d'atributs clínics a través d'abstraccions temporals. CAPTA s’avalua i s'aplica experimentalment en quatre dominis de predicció per als pacients amb discapacitats neurològiques.

Contents

Acknowledgements **5**

Abstract **7**

Abstract (in Catalan) **9**

List of Figures **15**

List of Tables **17**

List of Abbreviations **19**

List of Definitions **21**

Chapter 1 Introduction **23**

1.1 Research motivation 23

1.2 Thesis objectives 24

1.3 Thesis structure 25

Chapter 2 State of the art **27**

2.1 Medical background 27

2.1.1 Spinal cord injury 29

2.1.2 Acquired brain injury 30

2.2 Interoperability in the medical domain 31

2.3 Visualization in medicine 33

2.4 Representation of time in intelligent health systems 37

2.5 Summary 40

Chapter 3 Interoperability of medical health records in disabilities **41**

3.1 Automatic method for the generation of interoperable, multidimensional and holistic indicators of MHRs 41

3.2 Knowledge representation 46

3.2.1 Classes 47

3.2.2 Object properties 49

3.2.3 Difficulties of mapping clinical questionnaires into standard terminologies and ontologies 49

3.2.4 Development 52

3.3 Interoperability and information models 54

3.4 Summary 55

Chapter 4 Medical domain analysis 57

4.1 Monitoring system 57

4.1.1 Description of the monitoring system 57

4.1.2 Evaluation of the monitoring system 65

4.2 Prognosis analysis using available standard machine learning methods 67

4.2.1 Description of the prognosis analysis 67

4.2.2 Prognosis analysis 68

4.3 Summary 70

Chapter 5 Temporal case-based prognosis 71

5.1 Approach 72

5.1.1 Description of a case 72

5.1.2 Temporal representation 72

5.1.3 Case temporal structure comparison 74

5.2 Case-based reasoning 77

5.3 Evaluation 80

5.3.1 Emotional functions of people with SCI 81

5.3.2 Urination functions of people with SCI 83

5.3.3 Emotional functions of people with ABI 85

5.3.4 Executive functions of people with ABI 88

5.4 Visualization of the prognosis 89

5.5 Summary 94

Chapter 6 Conclusions and Future Work 97

6.1 Conclusions 97

6.2 Future work 98

6.3 Publications related to this thesis 99

6.4 Patents related to this thesis 100

Bibliography 101

Subject index **106**

Appendix I: Analysis of the socio-economic impact **107**

List of Figures

Figure 2:1 Global Burden Disease Cause Patterns, Institute for Health Metrics and Evaluation (IHME). Source: http://vizhub.healthdata.org/gbd-cause-patterns/ 28

Figure 2:2 Google Health interface. Source: Google's official blog [http://googleblog.blogspot.com.es/2010/09/google-health-update.html]. 33

Figure 2:3 Healthvault interface. Source: Microsoft Developer Network's blog [http://blogs.msdn.com/b/healthvault/archive/2012/07/19/healthvault-1206-release.aspx]. 34

Figure 2:4 PatientsLikeMe search tool to find similar patients. Source: PatientsLikeMe website [www.patientslikeme.com]. 35

Figure 2:5 Similar patient found in PatientsLikeMe. Source: www.patientslikeme.com. 36

Figure 2:6 Watson applied to CDSSs. Source: Ferrucci et al. (2013). 37

Figure 2:7 Thirteen base relations that model the possible relations between two intervals. Source: Allen (1983). 39

Figure 3:1 Architecture of the automatic generation of multidimensional indicators..... 41

Figure 3:2 Flow diagram of the automatic generation of multidimensional indicators. .. 43

Figure 3:3 Interaction of healthcare institutions in the automatic generation of multidimensional indicators. 45

Figure 3:4 Summary of ontology's classes and relationships. 47

Figure 3:5 Ontology's modifies relationship. 47

Figure 3:6 Ontology's is composed of relationship. 48

Figure 3:7 Ontological representation of people who has SCI. 53

Figure 3:8 Entity Relationship Modeling. 54

Figure 4:1 Identity management, creation of an account. 58

Figure 4:2 Identity management, profile. 59

Figure 4:3 Information exchange, management of questionnaires. 60

Figure 4:4 Graphical representation of the evolution of an individual across categories of the ICF. ICF values of difficulty, deficiency or barrier are represented with red/4 in complete levels, orange/3 in severe levels, yellow/2 in moderate levels, green/1 in mild levels and blue/0 in no difficulty, deficiency or environmental factors. 61

Figure 4:5 Graphical representation of the status of populations and individuals across categories of the ICF. 62

Figure 4:6 Graphical representation of the evolution of a population across categories of the ICF. 63

Figure 4:7 Graphical representation of the status of populations and individuals through the spinal cord injury measure (SCIM). 64

Figure 5:1 TA granularities. 73

Figure 5:2 Temporal distance for attribute a with respect to RLCS. 78

Figure 5:3 MAE and accuracy comparison in the prediction of emotional functions in SCI. 83

Figure 5:4 MAE and accuracy comparison in the prediction of urination functions in SCI. 85

Figure 5:5 MAE and accuracy comparison in the prediction of emotional functions in ABI. 87

Figure 5:6 MAE and accuracy comparison in the prediction of executive functions in ABI. 89

Figure 5:7 View of the prediction of a target case in the domain of “emotional functions of people with SCI” 92

Figure 5:8 View, when selecting the similar attribute *Immediate family,* comparing the target case and the most similar case.. 93

Figure 5:9 View, when selecting the dissimilar attribute *Remunerative employment,* comparing the target case and the most similar case. 94

Figure 1 Block diagram of the methodology used. 107

List of Tables

Table 2:1 Demographic and clinical data of the prognosis of SCI. 30

Table 2:2 Demographic and clinical data of the prognosis of ABI. 31

Table 2:3 Summary of temporal knowledge representations for A.I. systems. 38

Table 2:4 Summary of prognosis of neurological diseases. 39

Table 2:5 Benchmark of existing ML approaches applied to diseases of neurological origin. 40

Table 3:1 Object properties of the ontology. 49

Table 3:2 Encoding of indicator values into SNOMED CT and ICF. 50

Table 3:3 Encoding of indicator values into SNOMED CT and ICF. 50

Table 3:4 Relationships between ontology and vMR objects. 55

Table 4:1 Prediction of emotional functions of people with SCI without cost matrix. 68

Table 4:2 Prediction of emotional functions of people with SCI with cost matrix. 68

Table 4:3 Prediction of emotional functions of people with SCI. 69

Table 4:4 Prediction of urination functions in SCI. 69

Table 4:5 Prediction of emotional functions of people with ABI. 70

Table 4:6 Prediction of executive functions of people with ABI. 70

Table 5:1 Allen's temporal interval relations between intervals X=[x^-,x^+] and Y=[y^-,y^+]. .. 75

Table 5:2 Prediction of emotional functions of people with SCI without cost matrix minimizing MAE. 81

Table 5:3 Prediction of emotional functions of people with SCI without cost matrix maximizing accuracy. 81

Table 5:4 Prediction of emotional functions of people with SCI minimizing MAE. 82

Table 5:5 Prediction of emotional functions of people with SCI maximizing accuracy. ... 82

Table 5:6 Prediction of urination functions in SCI minimizing MAE. 84

Table 5:7 Prediction of urination functions in SCI maximizing accuracy. 84

Table 5:8 Prediction of emotional functions of people with ABI minimizing MAE. 85

Table 5:9 Prediction of emotional functions of people with ABI maximizing accuracy... 86

Table 5:10 Prediction (while minimizing MAE) of executive functions of people with ABI. 88

Table 5:11 Prediction of executive functions of people with ABI maximizing accuracy... 88

List of Abbreviations

ABI	Acquired brain injury
ADL	Activity of daily living
AI	Artificial Intelligence
CAPTA	CAse-based Prognosis using Temporal Abstractions
CBR	Case-based reasoning
CDSS	Clinical decision support system
DALY	Disability-adjusted life year
GBD	Global burden disease
ICD	International Classification of Diseases
ICF	International Classification of Functioning, Disability and Health
KNN	k-nearest neighbor
MHR	Medical Health Record
NB	Naïve Bayes
LCS	Longest common subsequence
PCE	Periodic comprehensive evaluation
QALY	Quality-adjusted life year
QoL	Quality of life
RBR	Rule-based reasoning
SCI	Spinal cord injury
SNOMED CT	Systematized nomenclature of medicine – clinical terms
SPARQL	Protocol and RDF Query Language (SPARQL)
SVM	Support vector machine
TA	Temporal abstraction
TBI	Traumatic brain injury
VMR	Virtual medical record
WHO	World Health Organization

List of Definitions

Core set	A selection of indicators of international standards such as the *International Classification of Functionality, Disability and Health* (ICF), relevant to patients affected by specific diseases or processes.
Index	A combination of indicators, questionnaires and possibly other indexes. The function representing this combination gives as summarizing result a score.
Indicator	A (subjective or objective) parameter or descriptor used to measure or compare activities and participation, body functions, body structures, environment factors, processes, and results.
Information	A representation of concepts and the relationships, constraints, rules, and operations to specify data model semantics for a chosen domain of discourse.
MHR	Set of both written and graphical documents, referring to information on health and illness of a patient, and on the healthcare activity related to this information, stored in electronic form.
Ontology	In the context of computer and information sciences, an ontology defines a set of representational primitives with which to model a domain of knowledge or discourse. The representational primitives are typically classes (or sets), attributes (or properties), and relationships (or relations among class members). The definitions of the representational primitives include information about their meaning and constraints on their logically consistent application. In the context of database systems, ontology can be viewed as a level of abstraction of data models, analogous to hierarchical and relational models, but intended for modeling knowledge about individuals, their attributes, and their relationships to other individuals. Ontologies are typically specified in languages that allow abstraction away from data structures and implementation strategies; in practice, the languages of ontologies are closer in expressive power to first-order logic than languages used to model databases. For this reason, ontologies are said to be at the "semantic" level, whereas database schema are models of data at the logical or physical level. Due to their independence from lower level data models, ontologies are used for integrating heterogeneous databases, enabling interoperability among disparate systems, and specifying interfaces to independent, knowledge-based services. In the technology stack of the Semantic Web standards, ontologies are called out as an explicit layer. There are now standard languages and commercial and open source tools for creating and working with ontologies.
Process	(in this thesis) A description and ordering of work activities across time and space that is designed to yield specific results or services while ensuring the rehabilitation's overall objectives.
Questionnaire	(or test) A set of questions answered using a scale.

Scale	A mapping between some ordered (qualitative or quantitative) values (or grades) and their description. These values are used to answer questionnaires.
Terminology	A system of terms belonging or peculiar to a science, art, or specialized subject; nomenclature

Chapter 1 Introduction

1.1 Research motivation

Health interventions are slowly adopting and taking advantage of information technologies (IT), which will improve the ability to understand, model and redesign the processes addressing diseases and their evolution. Healthcare practitioners can greatly benefit from the use of IT, such as knowledge-representation techniques that are already available today, and from current capabilities of information extraction and analysis. Recent advances in artificial intelligence (AI) have raised great interest worldwide in their use in medicine, for example in computer-aided prognosis and in helping doctors personalize treatments. Nevertheless, this paradigm shift comes along with a host of new challenges that healthcare systems have to face, such as the following ones.

Healthcare systems often *lack interoperability*, the ability to interpret and share data. Terminologies, ontologies and information models are used to solve this problem. Most healthcare systems *lack graphical representations to show the temporal evolution of information or population analyses*. *Prognosis capabilities are currently limited*, but slowly adopting information technologies which improve the ability to understand the evolution of diseases. Often, *time is overlooked and not integrated into knowledge-based systems*. This simplification leads to poor reasoning. There is a *limited capability for recommending therapies* in many healthcare domains. Finally, when integrating healthcare systems, there are *difficulties in maintaining the knowledge bases*.

This thesis focuses on the aforementioned issues and improves the state of the art in several aspects.

We start with interoperability. Clinical centers use questionnaires depending on their cultural aspects, organizations and countries. Consequently, there are often different instruments for capturing the same type of information, with the consequent lack of interoperability. Therefore there is the need to standardize data using international standards. When speaking about interoperability, we also need to discuss ontologies, which define a set of representative classes, attributes and relationships. Due to their independence from lower level data models, ontologies are used for integrating heterogeneous databases, enabling interoperability among disparate systems, and specifying interfaces to independent, knowledge-based services.

Another challenge we would like to address is the limited ability for temporal or population analysis. It is a challenge to understand individual and population data, and to visualize their status and evolution. Most systems display aggregated information of patients, but they do not show similar patients in a way that allows easy extraction of temporal knowledge.

A third challenge concerns the design of prognosis models, which are used to predict the course and outcome of disabilities. Prognosis models include different techniques: hand-crafted, statistical or AI-based. There are already some initiatives that provide prognosis such as detecting risk patterns in diabetes (Armengol, 2001), temporal abstractions for hemodialysis (Montani, 2013) or sequence of patterns for stress (Funk, 2006). However, many difficulties remain such as the implementation of prognosis in a multi-dimensional domain such as the one of disabilities of neurological origin.

Finally, time plays an important role in medicine and sometimes is oversimplified in healthcare systems. As mentioned previously, this oversimplification can lead to loss of information and not interpreting the evolution of an event over time. How to integrate time with AI systems it is an open field of research.

1.2 Thesis objectives

In this thesis we focus in disabilities of neurological origin. According to the WHO, the diseases of neurological origin are one of the leading causes of burden of disease in the world. Two populations of patients are used in this thesis: people who suffer from spinal cord injury (SCI) and people who suffer from acquired brain injury (ABI). Both datasets have been kindly provided by Institut Guttmann. More specifically, we work in four prediction domains for patients with neurological disabilities: emotional and urination functions of people with SCI, and emotional and executive functions of people with ABI.

In the rehabilitation of people who suffer from neurological diseases, there are two basic stages: intra-hospital and extra-hospital. In the intra-hospital (acute) phase, patients who suffered from a traumatic or nontraumatic injury stay in hospital undergoing rehabilitation. After typically a few months in hospital, they return home and the extra-hospital phase starts. Thereafter they go once a year to the rehabilitation hospital for a periodic comprehensive evaluation (PCE), when they are administered several questionnaires (depending on the disease they are suffering from) with the aim of measuring functioning independence, psychological and social variables. This study is focused on the extra-hospital or chronic phase.

This thesis contributes to the improvement of the state of the art in these issues. In particular, regarding interoperability this thesis proposes a new automatic translation system to international standards promoted by the WHO and a new ontology.

Regarding the ability for temporal or population analysis, this thesis provides a monitoring system with novel visualization techniques that allows a better understanding of the evolution of clinical attributes in individuals and populations. More specifically, it provides a graphical representation of states and their evolution that permits easily comparing the characteristics of an individual with those of the whole population.

With respect to prognosis, in this thesis we have used case-based reasoning (CBR), an AI technique based on solving new situations by learning from past similar situations already solved (Armengol, 2001; Aamodt and Plaza, 1994). Unlike monitoring systems, in CBR-based prognosis systems the target patient is related to a small set of similar patients (instead of the whole population of patients). This thesis designs and develops a CAse-based Prognosis using Temporal Abstractions (CAPTA), a prognosis system with a new case-based reasoning method with promising results.

Finally, regarding time, CAPTA integrates the temporal nature of clinical attributes through temporal abstractions. CAPTA is applied and experimentally evaluated in four prediction domains for patients with neurological disabilities.

To sum up, and taking into account the limitations of current systems explained in Section 1.1, the aims of this thesis are:

- To provide a method to generate automated translation of Medical Health Records (MHRs) to international standards promoted by the WHO.
- To provide a standardized representation of all concepts involved using an ontology.
- To provide new visualization interfaces to allow clinicians to understand both the status and evolution of individuals and populations.
- To improve the prognosis of the status of people with disabilities of neurological origin through:
    - Developing a case-based reasoning system with a time-aware similarity measure; and
    - Providing a better understanding of the temporal context of attributes by displaying a target patient, the three most similar patients, and the main similar and dissimilar attributes of each retrieved case with respect to the target patient.

1.3 Thesis structure

The thesis is organized as follows:

- Chapter 2 provides some medical background and describes the state of the art of healthcare systems in terms of interoperability, usability of interfaces, CBR, and design of an appropriate representation of time.
- Chapter 3 addresses interoperability in the disability domain and proposes a new method and a new ontology to solve it. The results are published in Subirats et al. (2013a), Miralles et al. (2013), Ceccaroni and Subirats (2012), Subirats et al. (2012).
- Chapter 4 analyzes the medical domain by describing a new monitoring system and performs a prognosis analysis using standardized data. The prognosis analysis has been carried out using standard machine learning methods. The results are partially published in Subirats et al. (2015) and Subirats et al. (2013a).
- Chapter 5 presents our contribution to temporal case-based prognosis.
- Chapter 6 concludes the thesis outlining the main contributions and future work.
- Appendix I analyzes the potential impact of a prognosis system in the cardiac rehabilitation domain. The results of this experimentation are published in Calvo et al. (2013).

Chapter 2 State of the art

The first section of this chapter motivates the study of a prognosis system for people who have a disability of neurological origin. In addition, it illustrates the application scenarios and provides a detailed description of the four data sets which are used in this thesis.

Interoperability, usability of interfaces and the design of an appropriate representation of time are common requirements that current clinical decision support systems (CDSSs) have to meet. The Sections 2.2 to 2.4 of this chapter present the state of the art of these three problems in the medical domain.

2.1 Medical background

In 1990, the World Health Organization (WHO) introduced a new measure to quantify the burden of diseases, injuries and risk factors: the disability-adjusted life year (DALY). The DALY is based on years of life lost from premature death and years of life lived in less than full health. According to the World Health Organization (2004) (The global burden of disease: 2004 update), the diseases of neurological origin are among the 20 leading causes of burden of disease in the world, specifically:

- unipolar depressive disorders, 3rd leading cause, 66 million DALYs;
- cerebrovascular disease, 6th, 47 M DALYs;
- road traffic accidents, 9th, 41 M DALYs.

From 2004 to 2030, there will be substantial changes in the leading causes of DALYs globally. The projected DALYs in 2030 are:

- unipolar depressive disorders, 1st leading cause of burden of disease, 84 million DALYs;
- cerebrovascular disease, 4th, 58 M DALYs;
- road traffic accidents, 3rd, 67 M DALYs.

This data is especially relevant; if we consider that the global burden of disease is projected to decrease in 2030. Figure 2:1 shows a tool to visualize global burden disease in the different years.

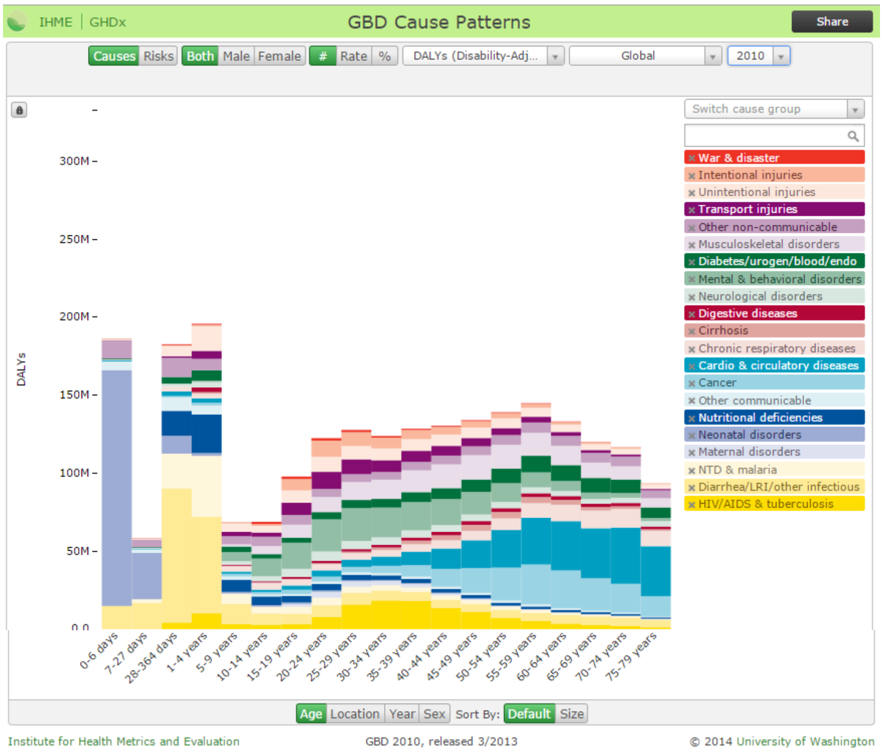


Figure 2:1 Global Burden Disease Cause Patterns, Institute for Health Metrics and Evaluation (IHME). Source: http://vizhub.healthdata.org/gbd-cause-patterns/

In the rehabilitation of people who suffer from neurological diseases, there are two basic stages: *intra-hospital* and *extra-hospital*. In the intra-hospital (acute) phase, patients who suffered from a traumatic or non-traumatic injury stay in hospital undergoing rehabilitation. After typically a few months in hospital, they return home and the extra-hospital phase starts. Thereafter they go once a year to the rehabilitation hospital for a *periodic comprehensive evaluation* (PCE), when they are administered several questionnaires (depending on the disease they are suffering from) with the aim of measuring functioning independence, psychological and social variables. In Section 2.1.1 and 2.1.2 we present examples of questionnaires administered in the PCE. This study is focused on the extra-hospital or chronic phase.

There are two main categories of CDSSs: those oriented to *assessment* and those oriented to *proposal*. The objectives of the assessment-oriented ones are: assessment of a patient's past, current and future status (this includes prognosis, i.e., the likely outcome of an illness); risk quantification; and classification of patients according to their functional diversity. The objectives of proposal-oriented ones are: risk prevention; and definition of therapeutic goals. We designed and developed a prognosis system oriented to *assessment* and, more specifically, to prognosis, in a clinical-system environment which requires reasoning under uncertainty, with standard indicators (e.g., *Ingestion functions* - b510) used in the representation of cases. The way of giving advice can be a *push* action, with warnings or alarms, or a *pull* action, giving support when explicitly requested. In the intra-hospital stage, support is pulled during all the rehabilitation process and pushed in the initial assessment and at the end of each rehabilitation activity. In the extra-hospital stage, advice is pulled in the prognosis support and pushed at the end of the PCE. In the style of communication and decision-making process applied to rehabilitation, both a consulting model and a critiquing model are used. In our case, the consulting model is used; the system provides an assessment of health future status and risks; and the final decision is taken by the health professional.

The proposed prognosis system uses two different datasets provided by Institut Guttmann: the first one is of people with spinal cord injury (SCI) and the second one of people with acquired brain injury (ABI) having as attributes demographic and clinical data. *Clinical data* varies over time and it is collected once a year.

2.1.1 Spinal cord injury

The following example corresponds to a real case scenario of a person with SCI.

Pol is the (anonymized) name of a man who suffered an accident 47 years ago while driving to Barcelona. Pol's accident resulted in paraplegia with a severe lifelong disability, which changed his life dramatically. After finishing his rehabilitation in Institut Guttmann, he returned to his hometown, but once a year he visits Institut Guttmann to perform his PCE. There health professionals, more specifically a psychologist and a doctor, use a monitoring system to assess and predict his status. The prediction allows professionals foresee complications and to act accordingly. For example, in one periodic evaluation, his psychologist suspects he will have emotional problems (depression and/or anxiety). She then consults the monitoring tool that predicts that he will have a higher deficiency in emotional functions next year. So she derives him to an external consultation to avoid problems. Then the doctor checks his state. The doctor recommends that he uses a collector instead, but Pol argues that he prefers to perform intermittent self-catheterization. The doctor then uses the prognosis tool to predict Pol's urinary function problems based on previous cases. After seeing the prediction of his evolution, the patient follows the doctor recommendation of using a collector.

The questionnaires used to gather *clinical data* in the PCE for people who suffer SCI are: Institut Guttmann social scale (ESIG), Personal well being index (PWI), Community integration questionnaire (CIQ), Craig Hospital inventory of environmental factors (CHIEF), Hospital Anxiety and depression (HAD), Patient health questionnaire (PHQ9), WHO quality of life questionnaire (WHOQOL), Functional independence measure (FIM), Spinal cord injury measure (SCIM) and ASIA scale.

Regarding *demographic data*, the data used in the two predictions are described in Table 2:1.

Table 2:1 Demographic and clinical data of the prognosis of SCI.

Attribute\Prognosis	**Urination functions** (469 people)	**Emotional functions** (493 people)
Age	[18,86], mean = 51.9, stdDev = 15.4	[18,86], mean = 51.6, stdDev =15.0
Gender	female (151), male (318)	female (152), male (341)
Years from diagnosis	[4,52], mean = 17.4, stdDev = 10.0	[4,52], mean = 17.4, stdDev = 9.9
Disease	Complete paraplegia (166), incomplete paraplegia (156), complete tetraplegia (44), incomplete tetraplegia (103)	Complete paraplegia (178), incomplete paraplegia (166), complete tetraplegia (45), incomplete tetraplegia (104)
Origin	Traumatic (290), medical (179)	Traumatic (308), medical (185)
ASIA classification (see Marino et al. (2003))	A(210), B(57), C(56), D(120), E(4), ?(22)	A(224), B(57), C(60), D(126), E(4), ?(22)
Neurological level (see Marino et al. (2003))	C1 (6), C2 (7), C3 (6), C4 (36), C5 (37), C6 (23), C7 (21), C8 (8), T1 (2), T2 (5), T3 (22), T4 (30), T5 (24), T6 (24), T7 (13), T8 (16), T9 (11), T10 (25), T11 (27), T12 (26), L1 (19), L2 (14), L3 (22), L4 (6), L5 (5), S1 (2), S2 (3), NO (7), ?(22)	C1 (6), C2 (7), C3 (6), C4 (35), C5 (40), C6 (22), C7 (21), C8 (8), T1 (3), T2 (6), T3 (23), T4 (35), T5 (23), T6 (24), T7 (14), T8 (20), T9 (11), T10 (27), T11 (32), T12 (27), L1 (19), L2 (15), L3 (22), L4 (6), L5 (5), S1 (3), S2 (3), NO (8), ?(22)
Length of the time series	[3, 7], mean = 4.3, stdDev = 1.0	[3, 7], mean = 4.3, stdDev = 1.0
Missing values of the time series in the predicted attribute	2006 (52%), 2007 (49%), 2008 (49%), 2009 (39%), 2010 (41%), 2011 (55%), 2012 (0%)	2006 (58%), 2007 (50%), 2008 (48%), 2009 (39%), 2010 (39%), 2011 (55%), 2012 (0%)
Predicted attribute (urination \ emotional functions)	No deficiency (61), mild deficiency (268), moderate deficiency (18), severe deficiency (11), complete deficiency (111)	No deficiency (35), mild deficiency (325), moderate deficiency (81), severe deficiency (52), complete deficiency (0)

2.1.2 Acquired brain injury

The following example corresponds to a real case scenario of a person with ABI.

> Anna is the (anonymized) name of a 43-year woman from Barcelona who suffered a head injury due to a car accident 26 years ago. After completing her regular assessment at the Institut Guttmann, a psychologist visits her and displays a graphical monitoring system that represents the state of the assisted person individually and compared with people with the same disease. The prognosis system helps her psychologist to predict problems in emotional functions and executive functions.

The questionnaires used to gather *clinical data* in the PCE for people who suffer ABI are: Institut Guttmann social scale (ESIG), Community integration questionnaire (CIQ), Patient competency rating scale (PCRS), Behavioral scale, PCRSi (informer), Rancho scale levels of cognitive functioning, Barthel index, Disability rating scale (DRS) and Extended Glasgow outcome scale (GOSE).

Regarding *demographic data*, the used data in the two predictions are described in Table 2:2:

Table 2:2 Demographic and clinical data of the prognosis of ABI.

Attribute\Prognosis	**Emotional functions** (419 people)	**Executive functions** (477 people)
Age	[17, 90], mean = 46.7 stdDev = 15.5	[17, 90], mean = 47.2 stdDev = 15.6
Gender	female (145), male (274)	female (164), male (313)
Years from diagnosis	[4,67], mean = 17.9 stdDev = 15.7	[2, 72], mean = 19.0 stdDev = 17.0
Disease	Not assigned (2), Guillain-Barre (18), polio (14), plexus (5), mielomeningocele (20), traumatic brain injury (213), multiple sclerosis (43), other progressive diseases (22), children cerebral palsy (103), hemorrhagic stroke (122), thrombotic stroke (26), embolic stroke (12), undetermined ischemic brain stroke (24), other ischemic brain stroke (9), other degenerative diseases not traumatic (84), muscular dystrophy (1), poliradiculoneuritis (7), other (4)	Not assigned (2), Guillain-Barre (14), polio (7), plexus (3), mielomeningocele (14), traumatic brain injury (133), multiple sclerosis (23), other progressive diseases (13), children cerebral palsy (73), hemorrhagic stroke (87), thrombotic stroke (22), embolic stroke (10), undetermined brain stroke (14), other ischemic brain stroke (5), other degenerative diseases not traumatic (51), muscular dystrophy (1), poliradiculoneuritis (3), other (2)
Origin	Traumatic (131), medical (208), undefined (80)	Traumatic (134), medical (231), undefined (112)
Length of the time series	[3, 7], mean = 4.2, stdDev =1.0	[3, 7], mean = 4.2, stdDev = 1.0
Missing values of the time series in the predicted attribute	2007 (99%), 2008 (80%), 2009 (54%), 2010 (42%), 2011 (39%), 2012 (52%), 2013 (0%)	2007 (77%), 2008 (69%), 2009 (69%), 2010 (45%), 2011 (37%), 2012 (47%), 2013 (0%)
Predicted attribute (emotional \ executive functions)	No deficiency (120), mild deficiency (130), moderate deficiency (112), severe deficiency (39), complete deficiency (18)	No deficiency (100), mild deficiency (69), moderate deficiency (103), severe deficiency (91), complete deficiency (114)

2.2 Interoperability in the medical domain

In physical medicine and rehabilitation there is a lack of interoperability among representations of the knowledge about patients' status as explained in the introduction chapter. Interoperable representations are valuable for certain artificial intelligence systems whose reasoning is based on experience or analogy, because their reasoning result improves if it is based on knowledge accumulated through a wide and diverse range of cases spanning as many healthcare organizations as possible. International standards and standard proposals for nomenclatures, ontologies and information models exist in this domain, but the scope of most of them is too general to cope with the specificities that characterize rehabilitation.

In the domain of medicine in general and physical medicine and rehabilitation in particular, several standard terminologies and classifications exist, which can be used for knowledge representation and integration. Some examples are: the *systematized nomenclature of medicine – clinical terms* (SNOMED CT) (College of American Pathologists and National Health Service (2001); the *unified medical language system* (UMLS); GALEN; the *international classification of diseases* version 10 (ICD-10); and the *international classification of functioning, disability and health* (ICF) defined by the *World Health Organization* (WHO). In particular, the use of the ICF for measuring the functioning status and diversity with *multidimensional* indicators, at both individual and population levels, can contribute to solve interoperability problems among health institutions that employ different measuring questionnaires. The ICF framework classifies concepts of functioning, disability and health and specifies their range of values. In rehabilitation, which is a multidimensional process, the ICF is useful to achieve comparable and interoperable data collections and is suitable for knowledge representation in *clinical decision support systems* (CDSSs), together with information models such as the *virtual medical record* (vMR), to contribute to solve interoperability problems in the electronic exchange of clinical information.

Several classifications and terminologies are used in CDSSs. In rehabilitation, ICF is used for encoding patients' health status and professionals' *recommendations*. Recommendations are activities and changes in environmental factors suggested by professionals to improve the quality of life of the patient. ICD-10 is used for representing diseases and SNOMED CT for other attributes; in particular, SNOMED CT has the potential of getting ICF-related terms mainstreamed in clinical systems, providing more contexts when necessary and including semantic relationships.

The ICF belongs to a family of international classifications developed by the WHO. Its aim is to provide a unified and standard language and framework for the description of health and health-related statuses; it defines components of health and some health-related components of well-being. According to the WHO, disability derives from the interaction between functional limitations and an unaccommodating environment. People are not described as having a disability based upon a medical condition, but rather are described according to a detailed representation of their functioning, which uses these main classes: body functions, body structures, activities, participation and environmental factors.

- *Body Functions* is the domain most closely related to a medical model as it refers to the physiological and psychological functions of body systems.
- *Body Structures* are defined by the ICF as anatomic parts of the body such as organs, limbs and their components. *Activities* refer to a wide range of deliberate actions performed by an individual. They are actions undertaken in order to accomplish a task, such as walking or climbing stairs.
- *Participation* refers to activities that are integral to economic and social life, such as attending school or holding a job.
- *Environmental factors* make up the physical, social and attitudinal environment in which people live, including not only the physical environment, but the cultural and political environment as well.

A given level of impairment in body function will not necessarily translate into activity or participation limitations if the environment accommodates a person's functional condition. Disability comes from participation restrictions that result from the interaction of many factors. Each main domain includes more detailed subclasses up to 5 levels, so a person's health condition can be defined with different precision by an array of components of the ICF rated using the same 0–4 scale, representing the level of impairment, limitation, restriction or barrier encountered, expressed as: (4) complete; (3) severe; (2) moderate; (1) mild and (0) no problem.

Regarding the use of information models in rehabilitation *CDSSs*, vMR is still in process of improvement and evolution by VMR Project Team, and there are several CDSSs that use it. Moreover, vMR is designed to reduce development costs and time responses in CDSSs. As a consequence, although it is not widely implemented in hospitals and there are not any tools available to facilitate its implementation (unlike the information model EN/ISO 13606), it is probably the most appropriate information model for CDSSs today.

2.3 Visualization in medicine

Interfaces play an important role in medicine because they can help health professionals to interact friendly with them and this can lead to a better understanding of the information they provide. Existing interfaces in medicine represent time in several different ways:

- Some display a chart summarizing the evolution of variables over time and detail of their last values. An example is the permanently discontinued product Google Health (see Figure 2:2).
- Others depict a full temporal chart. An example is HealthVault (see Figure 2:3 and Figure 2:4).
- Others use colours in a temporal chart. An example is PatientsLikeMe (Frost, 2008) (see Figure 2:5).
- Some do not show temporal information at all. An example is Watson (see Figure 2:6).

Figure 2:2 Google Health interface. Source: Google's official blog [http://googleblog.blogspot.com.es/2010/09/google-health-update.html].

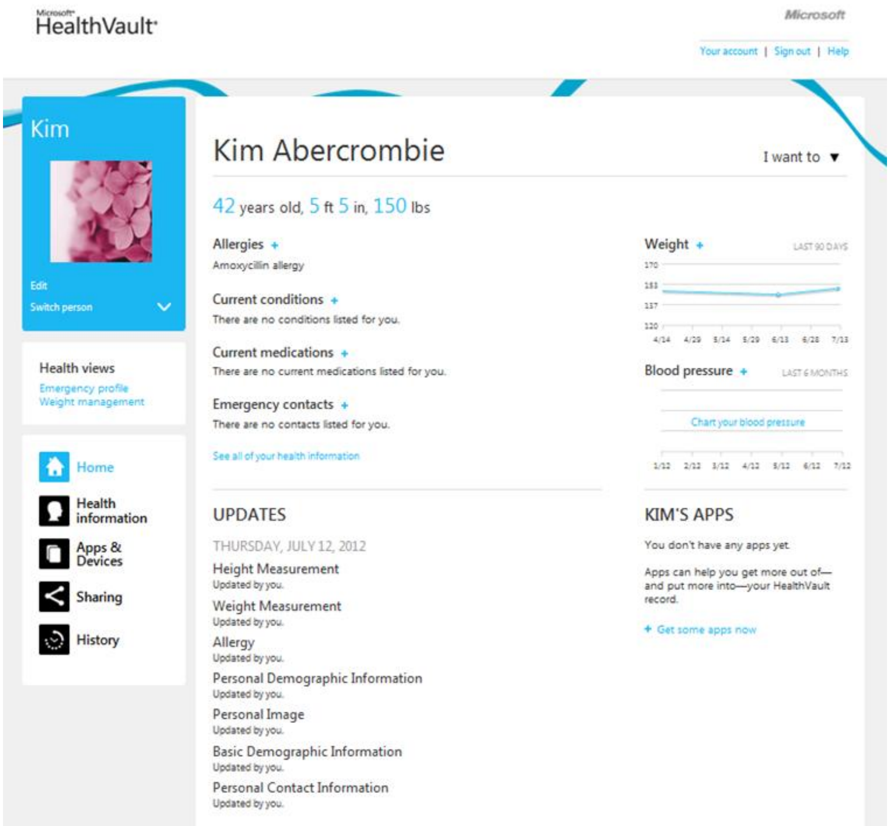


Figure 2:3 Healthvault interface. Source: Microsoft Developer Network's blog [http://blogs.msdn.com/b/healthvault/archive/2012/07/19/healthvault-1206-release.aspx].

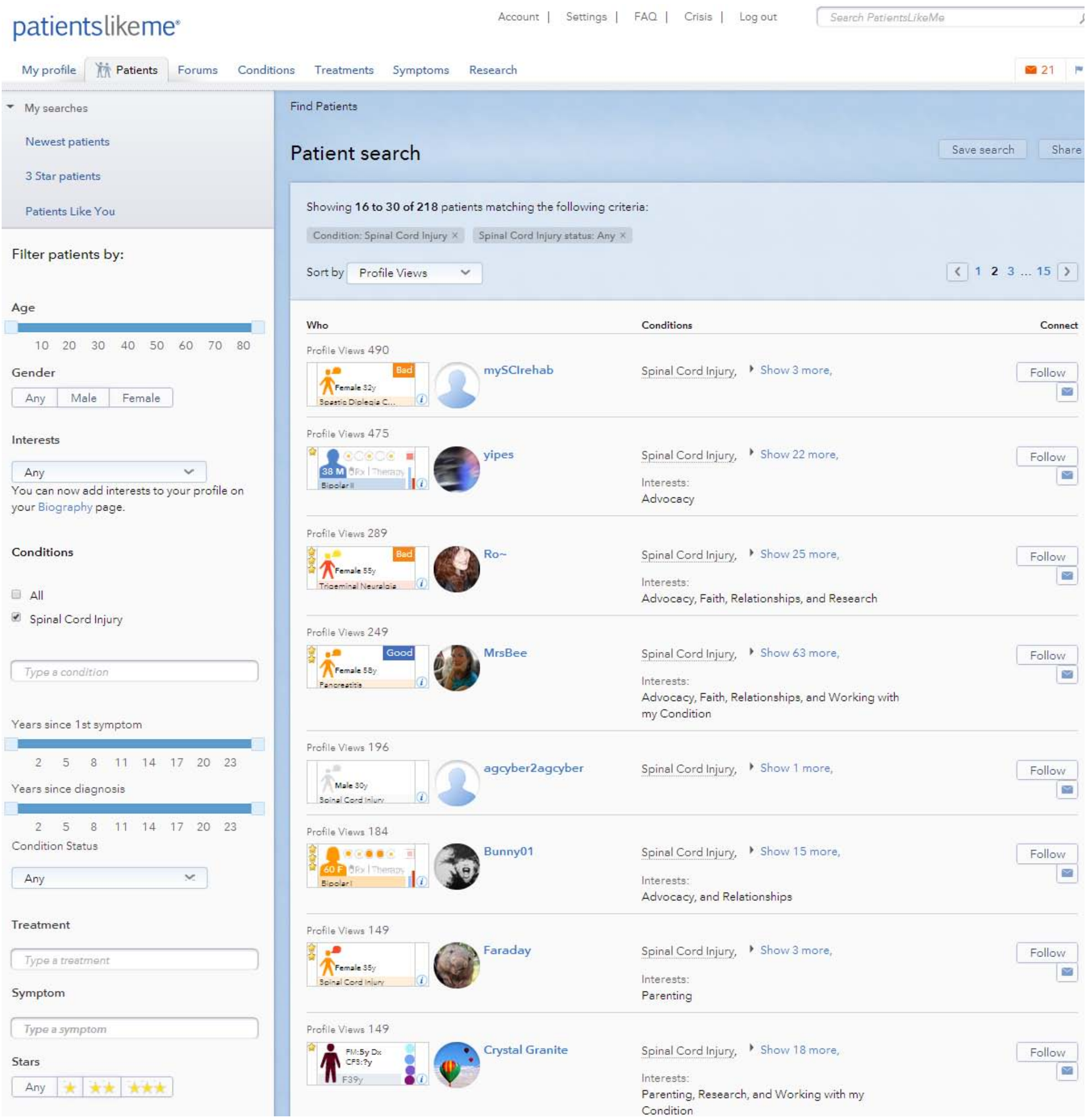


Figure 2:4 PatientsLikeMe search tool to find similar patients. Source: PatientsLikeMe website [www.patientslikeme.com].

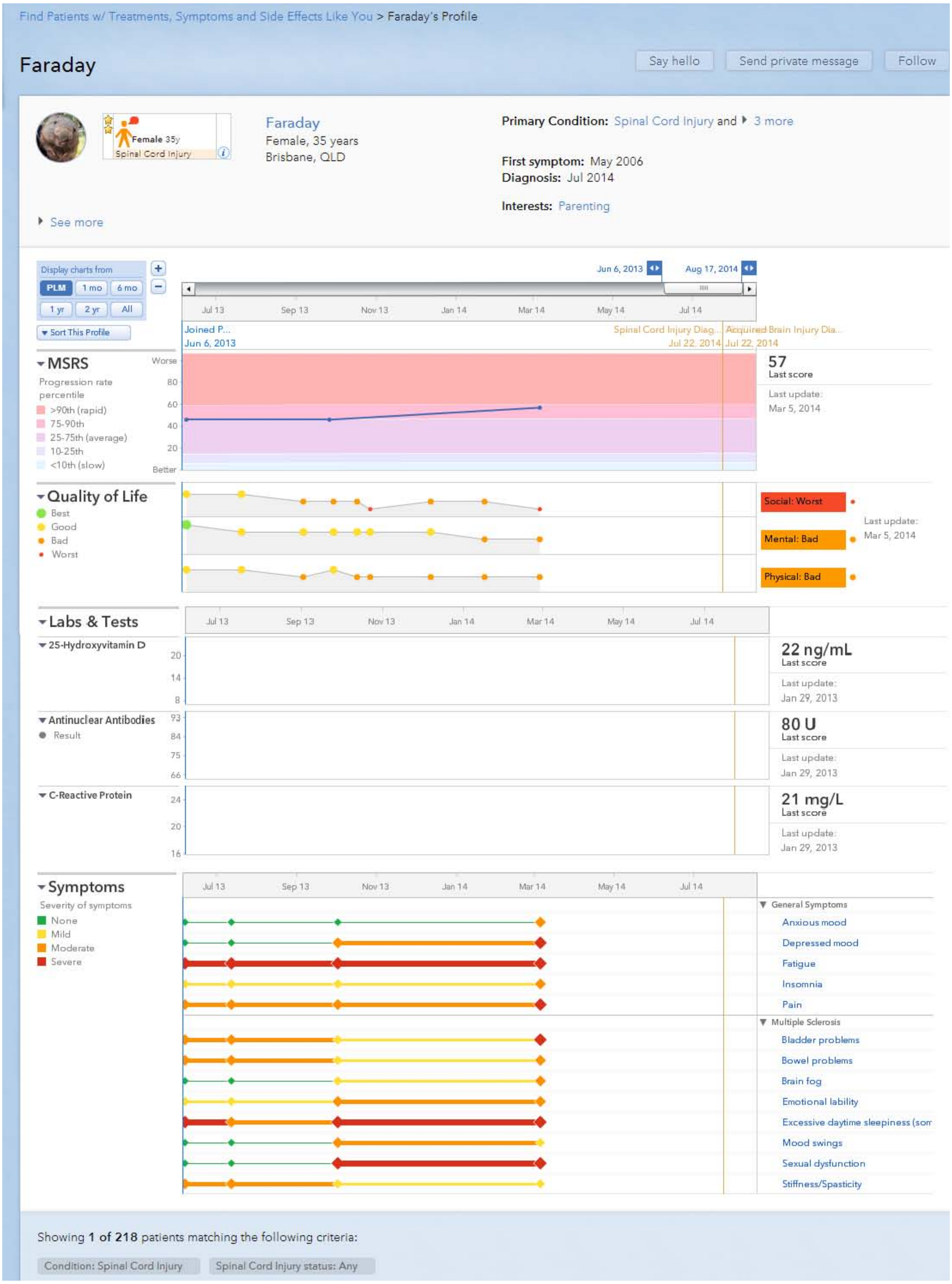


Figure 2:5 Similar patient found in PatientsLikeMe. Source: www.patientslikeme.com.

Figure 2:6 Watson applied to CDSSs. Source: Ferrucci et al. (2013).

2.4 Representation of time in intelligent health systems

Time is crucial in medical domains, and several approaches to include time in intelligent healt systems will be reviewed in this section. Specifically, we will focus on CBR and ML approaches that either deal specifically with temporal data or are applied to diseases of neurological origin. CBR (Aamodt and Plaza, 1994) has many advantages as it reduces the knowledge acquisition effort, requires less maintenance effort, improves problem solving performance through reuse, makes use of existing data (e.g. in databases), improves over time, adapts to changes in the environment and has high user acceptance.

Table 2:3 summarizes existing temporal knowledge representations from several domains, the representation of temporal knowledge in CBR being an open field of research today (Sànchez-Marrè, 2005). Work conducted by Jaere (2002), shows how *Allen's temporal intervals* (see Figure 2:7) can be incorporated into the semantic network representation of the *Creek system*. Allen's temporal intervals are also used by Ibrahim et al. (2011) to compute similarity between audiovisual documents. Other authors use *temporal ab-*

straction (TA) (Shahar (1997) and Bottighi et al (2010)) that are *interpretations* of temporal sequence of past and present values, actions or interactions in a more abstract form that are relevant for the given set of goals. For example, TA can provide a tendency about the status of a patient (see Schmidt (2003), Schmidt (2001), Montani (2013) and Chausa (2009)). In Armengol (2001), the risk pattern of each diabetic patient is obtained using a CBR method called *lazy induction of descriptions* (LID). In some studies, such as Funk (2006), cases containing series are classified into a number of categories and the knowledge discovery approach is able to identify *key sequences*; while in others, hierarchical representations are used to find temporal patterns (such as Raj (2009) and Montani (2013)). Finally, for the prediction of functioning of people with disability of neurological origin, existing ML classifiers are used in Serrà (2013) (although without time-series data) and *clustering based on rules* is used in Gibert (2009).

Table 2:3 Summary of temporal knowledge representations for A.I. systems.

Domain	Temporal knowledge representation	Reasoning method	References
Oil-Well Drilling	Allen intervals and Creek temporal representation	CBR	Jaere (2002)
Kidney function	Temporal abstraction (TA)	CBR	Schmidt (2003)
Spread of Diseases	Temporal abstraction (TA)	CBR	Schmidt (2001)
Diabetes	Risk pattern	Lazy induction of descriptions, a CBR method	Armengol (2001)
Hemodialysis	Temporal abstraction (TA)	CBR	Montani (2013)
Audiovisual documents	TA and Allen intervals	Similarity-based	Ibrahim (2011)
Stress	Sequence of patterns	CBR	Funk (2006)
Treatment outcomes of *human immunodeficiency virus* (HIV)	Temporal abstraction (TA)	Rule-based reasoning (RBR)	Chausa (2009)
Drug resistance of HIV	Hierarchical	RBR	Raj (2009)
Insulin-dependent diabetes	Temporal abstraction (TA)	-	Shahar (1997)

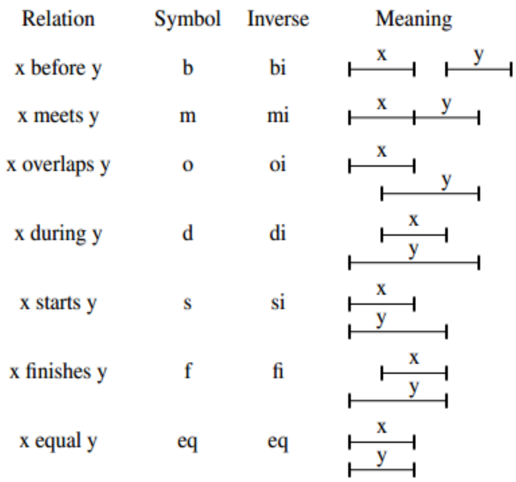


Figure 2:7 Thirteen base relations that model the possible relations between two intervals. Source: Allen (1983).

Table 2:4 summarizes existing ML and data-based approaches to predict functioning of people with neurological diseases. In Serrà (2013), accuracies achieved are practically always above 60% for all cognitive functions. The highest ones correspond to 61% for attention, 67% for memory, 61% for executive functions, and 64% for any using different types of classifiers. In Rovlias and Kotsou (2004), CART decision trees are used and the overall cross-validated predictive accuracy is 86.84%. In Pang (2007), the overall prediction accuracy ranges from 49.79% to 81.49%. On the other hand, in Brown (2005) the percentage of subjects misclassified is about 40%. In Gibert (2009), clustering based on rules by states is used to extract knowledge patterns about the evolution over time of the quality of life of people with spinal cord injury. More than half of the sample follows one of the four most frequent patterns of quality of life over time recommended by the system, while the remaining 47.7% are scattered among the other 21 observed patterns. In Whiteneck et al. (2004), regression is used to determine the impact of environmental barriers in functioning 4% or less in participation; and 10% in life satisfaction. Finally, in Westra et al (2011), a decision tree classifies 75.4% of patients' improvement in bowel incontinence.

Table 2:4 Summary of prognosis of neurological diseases.

Domain	Reasoning method	References
Acquired Brain Injury	Classifiers (Tree, k-Nearest Neighbour, Naïve Bayes, Support Vector Machine)	Serrà (2013)
Traumatic Brain Injury	Classifiers (CART decision tree)	Rovlias and Kotsou (2004)
	Classifiers (Tree)	Pang (2007)
	Classifiers (Tree)	Brown (2005)
Spinal Cord Injury	Clustering based on rules by states	Gibert (2009)
	Regression analysis	Whiteneck et al. (2004)
	Classifiers (Tree)	Westra et al (2011)

Table 2:5, shows the accuracy values of other ML approaches in the domain of diseases of neurological origin. After reading chapters 4 and 5, it can be seen that accuracies obtained in this study are lower than

the ones of other existing CDSS in the domain of diseases of neurological origin due to the nature of the data.

Table 2:5 Benchmark of existing ML approaches applied to diseases of neurological origin.

References	Evaluation
Rovlias and Kotsou (2004)	Accuracy of 86.84%, with a cross-validated relative error of 0.308.
Brown (2005)	Accuracy about 60%.
Pang (2007)	Accuracy achieved ranged from 49.79% to 81.49%.
Gibert (2009)	Half of the sample follows one of the four most frequent patterns of quality of life (QoL) over time recommended by the system, while the remaining 47.7% are scattered among the other 21 observed patterns.
Westra et al (2011)	Accuracy 75.4%
Serrà (2013)	Accuracies achieved are practically always above 60% in all cognitive functions. The highest ones correspond to 61% for attention, 67% for memory, 61% for executive functions, and 64% for any using different types of classifiers.

2.5 Summary

This chapter has illustrated the clinical motivation and problem and it has summarized the SoA of interoperability, existing interfaces and temporal representation in intelligent health systems. The solution, that it is a prognosis system, is described through the use of scenarios and used data sets are described in tables.

Regarding *interoperability*, the ICF is the most appropriate terminology to provide a unified and standard language and framework for the description of health and health-related status; because it defines components of health and some health-related components of well-being. On the other hand, the ICD is the most appropriate terminology to define diseases. Regarding information models, vMR is designed to reduce development costs and time responses in CDSSs. As a consequence, although it is not widely implemented in hospitals and there are not any tools available to facilitate its implementation (unlike the information model EN/ISO 13606), it is probably the most appropriate information model for CDSSs today.

Regarding existing *interfaces*, we can see that there are not interfaces that compare evolution of two cases in time, in order to help health professionals to understand the similarities and dissimilarities of patients' evolution.

Finally, regarding *temporal representation*, after reviewing the state of the art we can see that there are several limitations about how to represent time in intelligent systems. Temporal abstractions seem a promising field to represent time in the medical domain as they are *interpretations* of temporal sequences of past and present values, actions or interactions in a more abstract form that are relevant for the given set of goals. However, there are not studies about how temporal abstractions and Allen intervals are applied in a domain composed by quality values.

The next chapter will explain how the interoperability problem has been solved in the prognosis system.

Chapter 3 Interoperability of medical health records in disabilities

Terminologies, *ontologies* and *information models* are used in the health care information exchange to solve the interoperability problem. This chapter describes how our approach proposes new ways to address these three issues. The first section deals with the terminology and it explains a methodology to automatically collect, transform, share and graphically represent standardized and multidimensional indicators. The second section deals with the proposed ontology, how it is built, what problems have been faced when building it and how it has been implemented. Finally, the last section explains how information models such as vMR can be related to the proposed ontology.

3.1 Automatic method for the generation of interoperable, multidimensional and holistic indicators of MHRs

Figure 3:1 depicts the architecture of an automatic method for the generation of interoperable, multidimensional and holistic indicators from MHRs, which is comprised of an extractor, a transformer, an inference engine, a selector, a filter and a presentation system (Subirats et al. (2013a) and Miralles et al. (2011)).

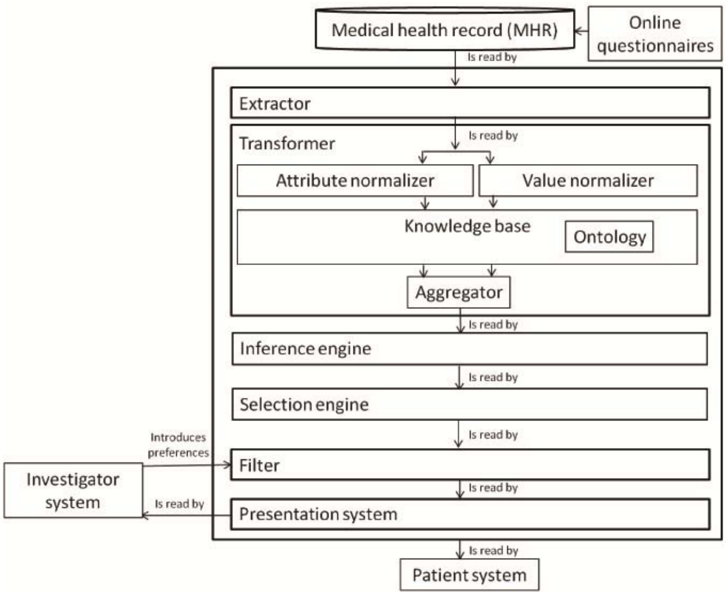


Figure 3:1 Architecture of the automatic generation of multidimensional indicators.

The *extractor* is configured to access databases and to selectively fetch information. Each time a value or piece of information is modified in a database, the extractor automatically updates the transformer, inference engine and selector. In the extraction of the information, patient-specific profiles are created. Patient's profiles can be based, for example, on diseases. The extractor then provides data to the transformer.

The *transformer* comprises an attribute normalizer, a value normalizer, and an aggregator. The normalizer however is not applied when standards such as SNOMED CT, which does not specify the range of values of the attributes, are used. In the same way, the aggregator is not used if there are not overlaps in collected data from questionnaires.

The *attribute normalizer* maps each extracted attribute to one or more normalized attributes. For this mapping one or a combination of international standards are selected (such as ICF or SNOMED CT). The mapping of extracted attributes to such systematic medical terminologies enables the recording of clinical data effectively and in a harmonized manner recognizable worldwide.

The *value normalizer* also receives the attributes from the extractor and proceeds to normalize their values to a standard scale. Values of attributes are normalized to values between 0 and 4, representing degrees of deficiency, from no deficiency to complete deficiency, respectively. This value normalization enables all patient-data to be evaluated on the same scale. Standards such as SNOMED CT do not specify a range of values in their attributes. In this case, it is not necessary and the original value type of the attribute is maintained. Not normalizing values has the advantage of not introducing potential artificial distortions in the original measurement, especially when they are represented as Booleans or strings. After normalization, an attribute can contain zero or more values, is classified into three classes, and is processed differently depending on its class.

- A normalized attribute of type-1 comprises no values.
- A normalized attribute of type-2 comprises only one value.
- A normalized attribute of type-3 comprises more than one value. In the case of the type-3, the pluralities of values are combined by aggregator automatically. This is done because although questionnaires are optimized for the problem they evaluate, there are overlaps in the normalized data collected from several questionnaires.

Questionnaires attributes can be normalized to one or more attributes of the selected nomenclature or nomenclatures. Therefore, using the appropriate aggregation function for a normalized attribute, the values of questionnaires attributes are combined into a single value providing the optimized representation of the information contained originally within the plurality of attributes. In the state representation of a person, each standardized attribute can have only one value.

Therefore, the *aggregator* manages overlaps among items which have been measured through different questionnaire questions. The aggregating function can follow different approaches. In a statistic approach, the aggregating function would be the average, the median or the mode value. In another, more optimistic approach, the best value; while, in a more pessimistic approach, the worst value. Furthermore, not all attributes need to have the same aggregating function. Depending on the nature of the attribute, one aggregation function can be more appropriate. The *aggregator* could be located in an external component, for instance, in an external presentation system. This enables centralized aggregation of data located in different health institutions. On the other hand, not all questionnaires are administered in the same time, therefore, an interval of study of the status of the patient is chosen. In this field, the interval chosen was a year, as questionnaires are administered in the PCE once a year. In the case of type-2, a single value complies with the minimum information necessary for evaluation. In the case of type-1, the problem exists that an empty entry can cause that the final evaluation of the patient's health is incomplete and even erroneous. In order to maximize the probability of correct health evaluation, inference engine is used to determine a value from related attributes, which are structured in a hierarchy.

The aim of the inference engine is to perform, when necessary and possible, the inference of the value of type-1 attributes from other attributes. After attribute normalization, the value of more general attributes can be inferred from more specific attributes.

After completing the processing of the normalized attributes according to their type, *selector* selects the relevant information depending on the particular profile being analyzed. Examples of relevant attributes that depend on each patient's values are non-empty values, extreme values, values which change over time, and values which are relevant for classifying patients in dysfunctional profiles. In addition, users are able to personalize the attributes they want to monitor depending on their personal experience or preferences.

The *filter* enables filtering population's status and the representation of each patient's evolution by a plurality of parameters: disease, cause of disease, lesion level and gender. When patients read health data from the representation system, the filter is not visible for them and population data are automatically filtered with population of their profile. Without the filter, data visualization is not personalized.

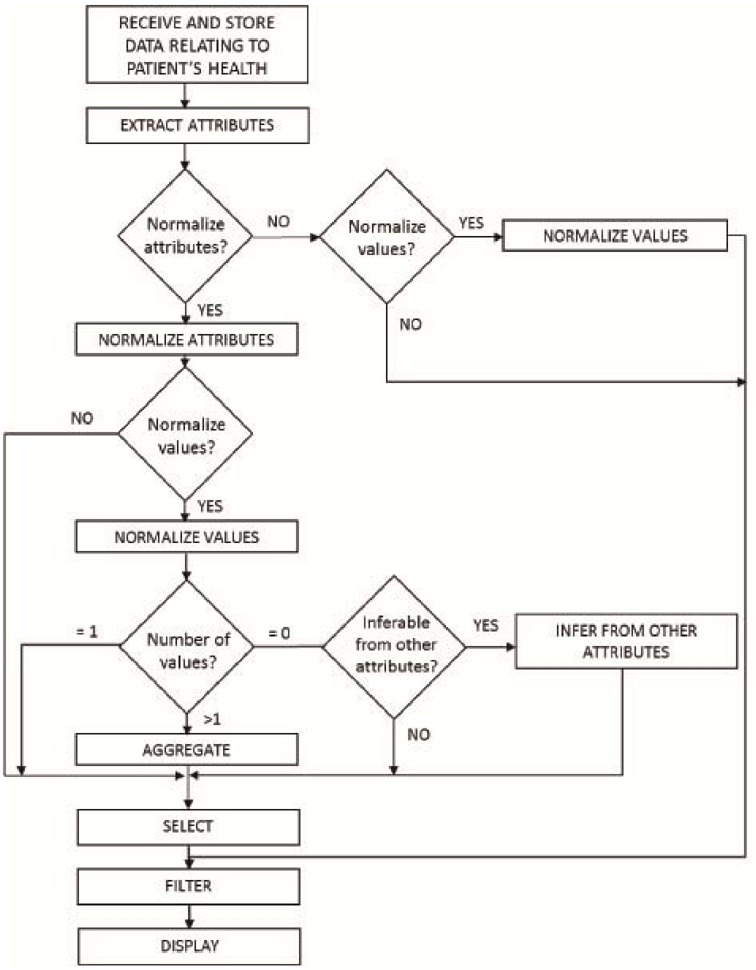


Figure 3:2 Flow diagram of the automatic generation of multidimensional indicators.

Figure 3:2 summarizes the described flow diagram, which is illustrated by the following real example:

- *Extract attributes*. One of the questionnaires people who suffer SCI are assessed is the FIM.
- *Normalize attributes*? Yes. The item *dressing upper body* is normalized to *dressing* (d540) and *putting on clothes* (d5400).
- *Normalize values*? Yes. Attributes of this scale have 7 levels for each item: complete independence, modified independence, supervision, minimal assistance, moderate assistance, maximal assistance and total assistance. As a consequence, values are normalized to the 0-4 scale of ICF. The normalization of values has been the following: 7 corresponds to 0; 5 and 6 correspond to 1; 3 and 4 correspond to 2; 2 correspond to 3; and 1 correspond to 4. The real example's value corresponds to 2, so the value of the dressing is 3.
- *Number of values > 1*? Yes. Attribute d540 is extracted also from SCIM scale. The value from SCIM scale is 2. Therefore, the value of this attribute is the average of values from both scales (2.5), which is rounded to 3.
- *Select*. Attribute d540 is a non-empty value, therefore it is displayed.
- *Filter*. In the representation, the population represented is people who suffered a complete paraplegia, therefore the example is included.

If the number of values of d540 was 0, then it would be inferred from its children attributes from the ICF taxonomy, such as *putting on clothes* (d5400), *taking off clothes* (d5401), *putting on footwear* (d5402), *taking of footwear* (d5403), *choosing appropriate clothing* (d5404), *dressing, other specified* (d5408) or *dressing, unspecified* (d5403).

The *presentation system* enables the graphical representation of all normalized data of a patient's health as well as evolution throughout time and taking into account the different profiles. It also allows patients to compare themselves with other patients with their same profile. Patients and health practitioners can select a particular time frame for which they want the data to be analyzed, for instance, a number of years that have passed since a particular lesion or discovery of the disease. The presentation system allows the visualization of aggregated data from different health institutions.

The *presentation system* is centralized and optimizes the management of patients and clinicians which could be located at different geographical locations worldwide.

Figure 3:3 shows the interaction of healthcare institutions if information comes from more than one MHR.

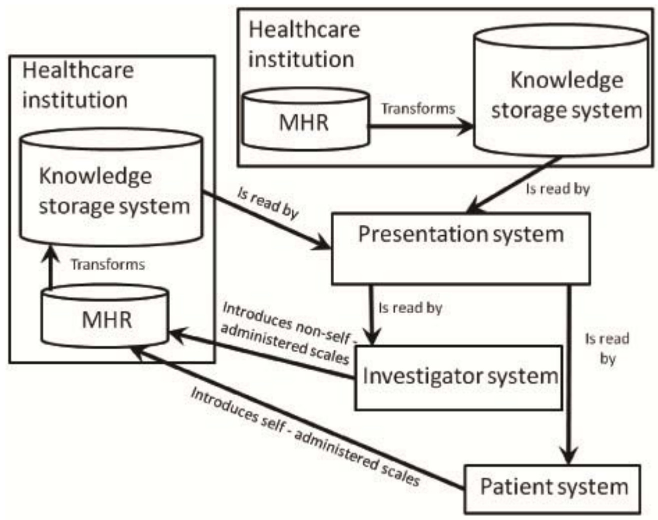


Figure 3:3 Interaction of healthcare institutions in the automatic generation of multidimensional indicators.

The obtained indicators of clinical data described in chapter 2 standardized to the ICF standard following this methodology are the following:

- *Body functions*: mental stability (b1263), optimism (b1265), trust (b1266), functions of energy and drive (b130), power level (b1300), motivation (b1301), impulse control (b1304), sleep function (b134), functions of care (b140), memory functions (b144), emotional functions (b152), appropriateness of emotion (b1520), range of emotion (b1522), emotional functions, other specified (b1528), thought content (b1602), troubleshooting (b1646), body image (b1801), feeling pain (b280), functions of blood pressure (b420), respiratory functions (b440), functions of ingestion (b510), functions of assimilation (b520), functions of defecation (b525), frequency of defecation (b5252), fecal continence (b5253), functions of the urinary excretion (b610), urination functions (b620), urinary continence (b6202) and functions of the gait pattern (b770).
- *Activities and participation*: carry out a single task (d210), conduct a simple task (d2100), undertaking multiple tasks (d220), carrying out daily routine (d230), communication (d3), communication - receiving spoken messages (d310), communication -receive text messaging (d325), talk (d330), producing nonverbal messages (d335), mobility (d4), sit (d4103), remain lying (d4150), remain seated (d4153), transfer the body (d420), transfer the body while sitting (d4200), transfer the body. while lying (d4201), lifting and carrying objects (d430), lifting objects (d4300), carrying in the arms (d4302), handle (d4402), rotate or twist your hands or arms (d4453), walking (d450), walking short distances (d4500), climb (d4551), move inside the house (d4600), move out of the home and other buildings (d4602), navigate using some kind of equipment (d465), walking and moving , other specified and unspecified (d469), using transportation (d470), drive (d475), self-care (d5), wash (d510), wash individual body parts (d5100), wash the whole body (d5101), caring for body parts (d520), care of the teeth (d5201), hair care (d5202), personal hygiene related to the processes of excretion (d530), dressing (d540), put on your clothes (d5400), take off your clothes (d5401), wear shoes (d5402), eating (d550), drink (d560), acquiring a place to live (d610), interpersonal interactions (d7), basic interpersonal interactions (d710), sex (d7702), no formal education (d810), paid work (d850), rights on public economics (d8701), recreation and leisure (d920), sport (d9201) and leisure time, other specified (d9208).

- *Environmental functions*: food (e1100), drugs (e1101), products and assistive technology for personal use in daily life (e1151), products and assistive technology for mobility and personal transportation in closed and open spaces (e1201), design, construction , building materials and architectural technology of buildings for private use (e155), design , construction, building materials and architectural technology to get access to the facilities within private buildings (e1551), financial belongings (e1650), natural environment and environmental changes resulting from human activity (e2), support and relationships (e3), close relatives (e310), friends (e320), acquaintances, peers , colleagues, neighbors and community members (e325), health professionals (e355), transport (e5400), services of general social support (e5750), health services (e5800), muscles of respiration (s4303), structure of the intestine (s540), structure of urinary system (s610) and structure of areas of skin (s810).

3.2 Knowledge representation

Here we carry out an ontology-based exploration of the concepts and relationships in the rehabilitation domain, integrating clinical practice, the ICD (specifically its 11th revision), the clinical investigator record ontology, the ICF and SNOMED CT. And we discuss a mapping of the relationships used in rehabilitation to existing standard concepts. The aim of the analysis is to identify potential logical problems with the use of existing models and international standards in representing and reasoning with real clinical data, and to understand whether and how these data might be defined more formally than in the current practice. The mapping of the information observed and recorded during the rehabilitation process to international standards proved partially feasible. Our analysis of the relationships among rehabilitation concepts revealed issues related to confusion among classes and their properties, incorrect classifications, overlaps and loss of information. It also suggested properties that should be included in a formal model suitable to be used by decision support systems.

Healthcare organizations use several tools to capture information. These tools make use of specific terms, which are sometimes ambiguous: descriptor, grade, index, indicator, parameter, questionnaire, scale, score and test. The terminology used in this paper is defined as follows and is part of an ontology, which we defined (and encoded in OWL 2) based on standard nomenclatures and ontologies:

- *Index*: a combination of indicators, questionnaires and possibly other indexes. The function representing this combination gives as summarizing result a score.
- *Indicator*: a (subjective or objective) parameter or descriptor used to measure or compare activities and participation, body functions, body structures, environment factors, processes, and results.
- *Questionnaire* (or test): a set of questions answered using a scale.
- *Scale*: a mapping between some ordered (qualitative or quantitative) values (or grades) and their description. These values are used to answer questionnaires.

The main metrics of the ontology (which can be accessed at https://github.com/laiasubirats/circlesofhealth are as follows: 1597 classes, 52 object properties, 945 individuals, ALCHI DL expressivity. There are not multiple primitive representations of the same concepts as it is specified in Cimino's guidelines (Cimino, 1998)).

3.2.1 Classes

Most top-level concepts come from SNOMED CT, e.g.: *Assessment scales* (subclass of *Staging and scales*), *Physical object*, *Process* (subclass of *Observable entity*). Some of them come from ICF, e.g.: *Environmental factors* (e), *Activities and participation* (d). ICF codes have a letter b, d, e or s possibly followed by a number while SNOMED CT codes are composed by a 9-digit number. In selecting and reusing concepts, we always maintained the hierarchical organization and consistency of the original standard ontologies. In any case, top-level categories are overly general to characterize the ontology, which can be better framed, conceptually, through other categories, more related to rehabilitation scenarios, such as: *Diagnosis*, *Functional independence measure* (FIM), *Orthotic devices* or *Rehabilitation therapy*.

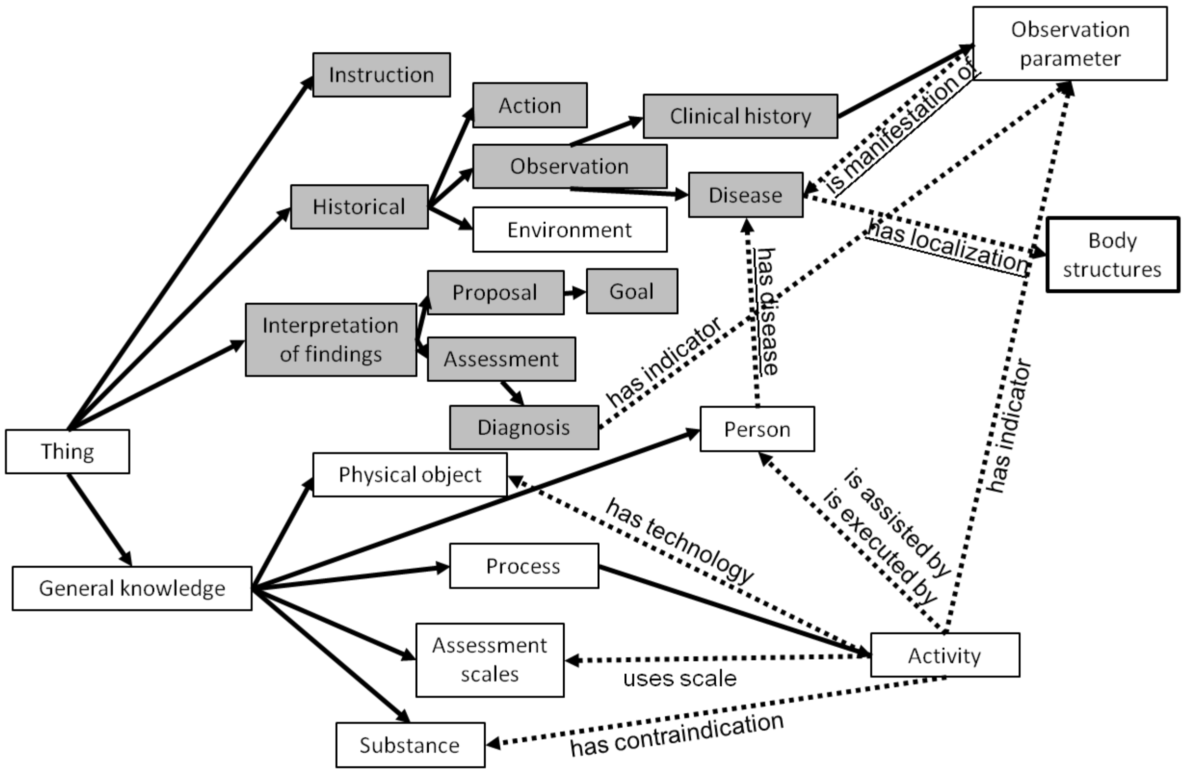


Figure 3:4 Summary of ontology's classes and relationships.

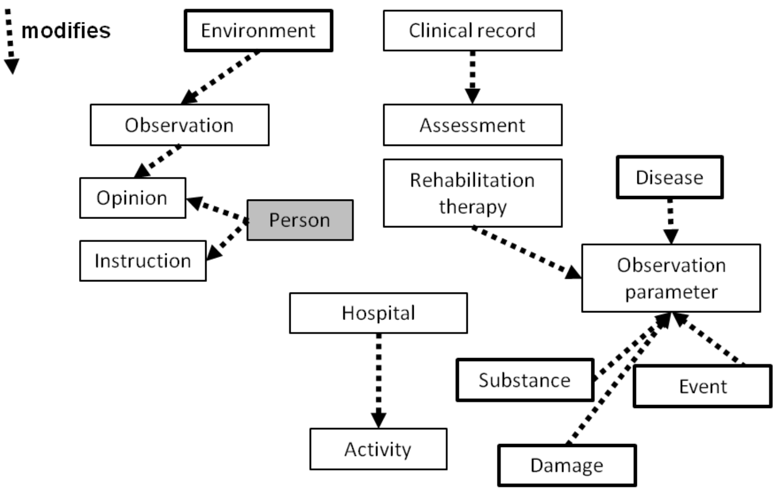


Figure 3:5 Ontology's modifies relationship.

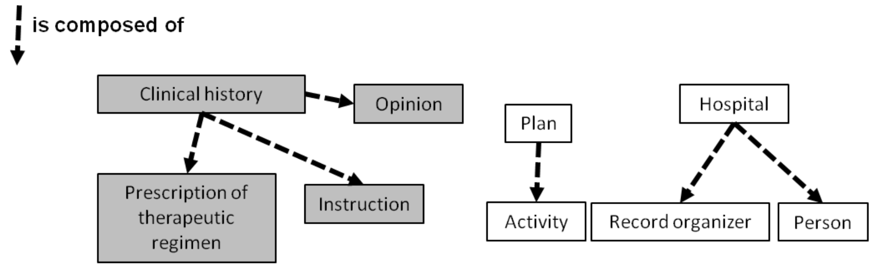


Figure 3:6 Ontology's is composed of relationship.

In Figure 3:4, Figure 3:5 and Figure 3:6 we show how the ontology is related to and integrate the state of the art. Concepts are encoded as SNOMED CT classes if not otherwise specified. Concepts in bold-frame boxes are encoded as ICF classes (e.g., Body structures). Concepts in grey boxes are reused from Beale and Heard (2007) *clinical investigator record* ontology (e.g., Instruction). Underlined relations are reused from the ICD (specifically its 11th revision, ICD-11[1]) (e.g., *has localization*). Concepts in white boxes and non-underlined relations represent an extension by the authors of existing approaches based on clinical practice (e.g., *has indicator*). Relations in the figure, if not otherwise specified, are "*is a*".

At the top-level of the ontology, there are the following concepts, as shown in Figure 3:4.

- *Historical* (following Sowa (2000)'s top-level categories), which subsumes:
    - Observation: information created by an act of observation, measurement, questioning, or testing of the patient or related substance;
    - Action: a record of intervention actions that have occurred, due to instructions or otherwise;
    - Environment: information on the context of the patient;
- *Interpretation of findings*: inferences of the investigator using the personal and published knowledge bases about what the observations mean, and what to do about them; includes all diagnoses, assessments, plans, goals;
- *Instruction*: instructions, based on interpretations of findings, sufficiently detailed so as to be directly executable by investigator agents (people or machines), in order to accomplish a desired intervention.

The *Interpretation of findings* category corresponds to Sowa (2000)'s *Description* category, to the notion of hypothesis in general science and to Rector (2011)'s *Meta-observations*. The *Instruction* category corresponds to Sowa (2000)'s *Script* category. *Assessment* (see Figure 3:4, top) relates to past, current or projected states of affairs. *Proposal* relates to desired ones. A *Diagnosis* is the attachment of a label to a group of observed signs and symptoms, which designates it (in the understanding of the investigator) as being a particular phenomenon. A *Differential diagnosis* allows for multiple possibilities, due to the lack of sufficient information or understanding to attach one label. A *Goal*, such as monitoring the upper limb while performing the *Activity of daily living* (ADL) *Dressing*, is a statement about what the desired state of the patient system should be, while a *Plan* is a statement about how to get there.

[1] Available online at http://apps.who.int/classifications/icd11.

3.2.2 Object properties

Object properties represent relationships between two classes or instances. Apart from properties *is a*, *is composed of* and *modifies*, object properties of the proposed ontology are described in Table 3:1. Some of them are based on ICD-11.

Table 3:1 Object properties of the ontology.

Object property	Domain	Range
Has contraindication	ADL	Substance
Has disease (ICD-11)	Person	Disease
Has goal	Person, ADL	Participation, Body functions, Body structures
Has indicator	Diagnosis, ADL	Observation parameter
Has interpretation	Diagnosis	Indicator
Has location (ICD-11)	Observation	Body structures
Has occupation	Person	Occupation
Has technology	ADL	Physical object
Is assisted by, is assisted as needed by	ADL	Person, Physical object
Is executed by	ADL	Person
Is manifestation of (ICD-11)	Disease	Observation parameter
Is programmed by	ADL	Person
uses	ADL	Observation parameter

Figure 3:4, Figure 3:5 and Figure 3:6 show the most used relationships among concepts, and the context in which they are used in a typical scenario. Continuous lines represent the relation *has subclass*. Discontinuous lines represent: the object properties defined in Table 3:1, the relation *composed of* or the relation *modifies*.

3.2.3 Difficulties of mapping clinical questionnaires into standard terminologies and ontologies

Indicators are the main classes used for representing the status of a patient and potentially number in the thousands. The two main classes used in rehabilitation are *process indicators* and *result indicators*. A process indicator is used to assess whether a task is being performed correctly. A result indicator is used to assess the performance in carrying out an activity or whether the objectives of the activity have been achieved.

To facilitate human practice, only a selection of indicators, grouped into *core sets*, are used. Core sets can be formed according to *functionality*, *pathology* or *rehabilitation process* (Laxe et al., 2012). Core sets are useful because human investigators can process only a fraction of the categories found in relevant terminologies such as ICF and SNOMED CT. Core sets already exist for several pathologies, such as multiple sclerosis, (SCI) or ABI, though finding the core categories for rehabilitation processes and moving from a pathology-based approach to one based on functionality and rehabilitation is needed.

Several problems exist with the standardization of observation parameters (questionnaires' indicator names and values) into SNOMED CT and ICF. One main difficulty is that there is a mismatch between the way SNOMED CT categorizes its concepts and the questionnaire items. SNOMED CT makes a distinction between *Clinical finding* and *Observable entity*. *Findings* are observations that are meaningful by themselves whereas *observables* need to have values to complete their meaning. A questionnaire item should be mapped only to SNOMED CT observables (see Table 3:2). Combinations of a questionnaire item and its values map to findings. If we take, for example, the *Transfers: toilet* item, the complete mapping to SNOMED CT and ICF is shown in Table 3:3. (Here we see that the extremes of the scale can be mapped to findings, as an alternative to a mapping using only observables.) The mapping to ICF doesn't suffer from the problem of category mismatch, as ICF categories are neutral with respect to functional diversity.

Table 3:2 Encoding of indicator values into SNOMED CT and ICF.

Questionnaire (FIM) item	SNOMED T	ICF
Dressing upper body	Ability to dress	Dressing Putting on clothes
Dressing lower body	Ability to dress Ability to put on footwear	Dressing Putting on footwear
Toileting	Toileting	Toileting
Transfers: bed/chair/wheelchair	Chair/bed transfer between wheelchair and toilet	Transferring oneself
Transfers: toilet	Ability to transfer between wheelchair and toilet	Transferring oneself
Locomotion: stairs	Climbing stairs	Climbing

Sections 3.2.3.1 and 3.2.3.2 deal with specific issues found in the standardization of indicators to ICF and SNOMED CT, respectively. Observation parameters are obtained from questionnaires FIM, SCIM, PCRS, GOSE, HAD from the *Historical* class.

3.2.3.1 Problems found in the standardization of observation parameters to ICF

Overlaps and loss of information are found when implementing the methodology of ICF-encoding of Cieza et al. (2002 and 2005).

Table 3:3 Encoding of indicator values into SNOMED CT and ICF.

Questionnaire item's value	SNOMED CT	ICF
Transfers: toilet	Able to transfer between wheelchair and toilet	Transferring one-self 0
Transfers toilet 6/5/4/3/2	Ability to transfer between wheelchair and toilet with modified independence / supervision / minimal assistance / moderate assistance / maximal assistance	Transferring oneself 1/1/2/2/3
Transfers: toilet 1	Unable to transfer between wheelchair and toilet	Transferring oneself 4

When the content of an item is not explicitly named in the corresponding ICF category, but at the same time is included, then the item is linked to this ICF category and the additional information not explicitly named by the ICF is documented.

The problem here is that when, e.g., functions of body structures are linked only to the activity or body function, body structures cannot be mapped and distinguished. For example, the ICF standardizations of *Dressing the upper body* and *Dressing the lower body* of FIM are the same (*Dressing*, d540) or they are not based on the original body structures (*Putting on clothes*, d5400, and *Putting on footwear*, d5402).

A solution would be to add relations representing the additional information not explicitly named by the ICF, e.g., the relation *has localization* to link to *Body structures*.

<u>*The response options of an item are linked if they refer to additional constructs.*</u>

Depending on the answer-option chosen, the item is standardized to some indicators or others. For example, the standardization of *Dressing the upper body* of SCIM is: "Support and relationships (e3)" if the answer is *Requires total assistance* or *Requires partial assistance*; "Dressing (d540)" if it is Independent; "Assistive products and technology for personal use in daily living (e1151)" if it is Independent but requires adaptive devices. Items, independently of their nature, are standardized to "e3" if the answer is *Requires total assistance*. This causes an aggregation of diverse content into one indicator, with potential loss of semantics.

A solution would be to keep the information found in the question and answer, and add a relation representing the degree of assistance, as shown for SNOMED CT in Table 3:3.

<u>*Items are linked into high level categories.*</u>

If the content of an item is standardized into high-level ICF categories, it can cause a loss of information as higher categories are a generalization of the concept and also do not usually appear in core sets. For example, the option *Requires total assistance* of nearly all items of SCIM is standardized to "e3". In the ICF core sets of SCI, there is a lower-level category *Health professionals* (e355) (which is not a completely satisfactory standardization), but the category "e3" is not there.

A solution would be to form new core sets according to rehabilitation processes (an initiative which is already ongoing) and taking into account this generality issue both in core-set definitions and in mappings of questionnaires.

<u>*If the information provided by the meaningful concept is not sufficient for making a decision about the most precise ICF category it should be linked to, the meaningful concept is assigned not definable (nd), personal factor (pf), not covered by ICF (nc) or health condition (hc).*</u>

For example, in PCRS, the item "Problem in accepting criticism" is standardized to "nd".

A solution would be to use SNOMED CT to complement the ICF, e.g., in this case, with the class *Tends to be sensitive to criticism*.

3.2.3.2 Problems found in the standardization of observation parameters to SNOMED CT

In the mapping of MHR observation parameters to SNOMED CT, the main difficulty is the standardization of the values of the indicators. Specifically, problems appear when different types of items are mapped to the same standard concept and their values need to be combined. Combining values, and changing data types, has the disadvantage that information is (potentially) lost.

For example, the observable entity *Climbing stairs* is the standardization of SCIM and FIM concepts. The range of values of FIM's *Locomotion: stairs* is 1 to 7, while the range of SCIM's *Stair management* is 0 to 3. Another example is the concept *Anxiety*, which is measured by GOSE as a Boolean, while in HAD it is presented as a subscale with a range of values between 0 and 21.

3.2.4 Development

As in typical, modern knowledge-based systems, an ontology is used to represent patient's information (Ceccaroni and Subirats, 2012) and it is updated using *SPARQL Protocol and RDF Query Language* (SPARQL) and Jena. Both the ontology and *OntologyInteraction* Java library are available under the LGPL at https://github.com/laiasubirats/circlesofhealth. If health professionals need to update the ontology, they can do it collaboratively using WebProtégé, which is integrated in the monitoring system. Figure 3:7 depicts the ontological representation of people who suffer from a SCI, which describes their administered questionnaires (summarized at the questionnaires and their ICF core sets (Bickenbach et al., 2012). All this information is extracted through SPARQL queries. An example of a SPARQL query to extract indicators obtained from questionnaires administered to people who has SCI is the following:

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

select * where { ?indicator

<http://purl.bioontology.org/ontology/PMR.owl#is_indicator_of>

<http://purl.bioontology.org/ontology/PMR.owl#Paraplegia_and_tetraplegia>

 .?indicator

<http://purl.bioontology.org/ontology/PMR.owl#code> ?code .?indicator

<http://purl.bioontology.org/ontology/PMR.owl#name> ?name .?indicator

<http://purl.bioontology.org/ontology/PMR.owl#description> ?description}
```

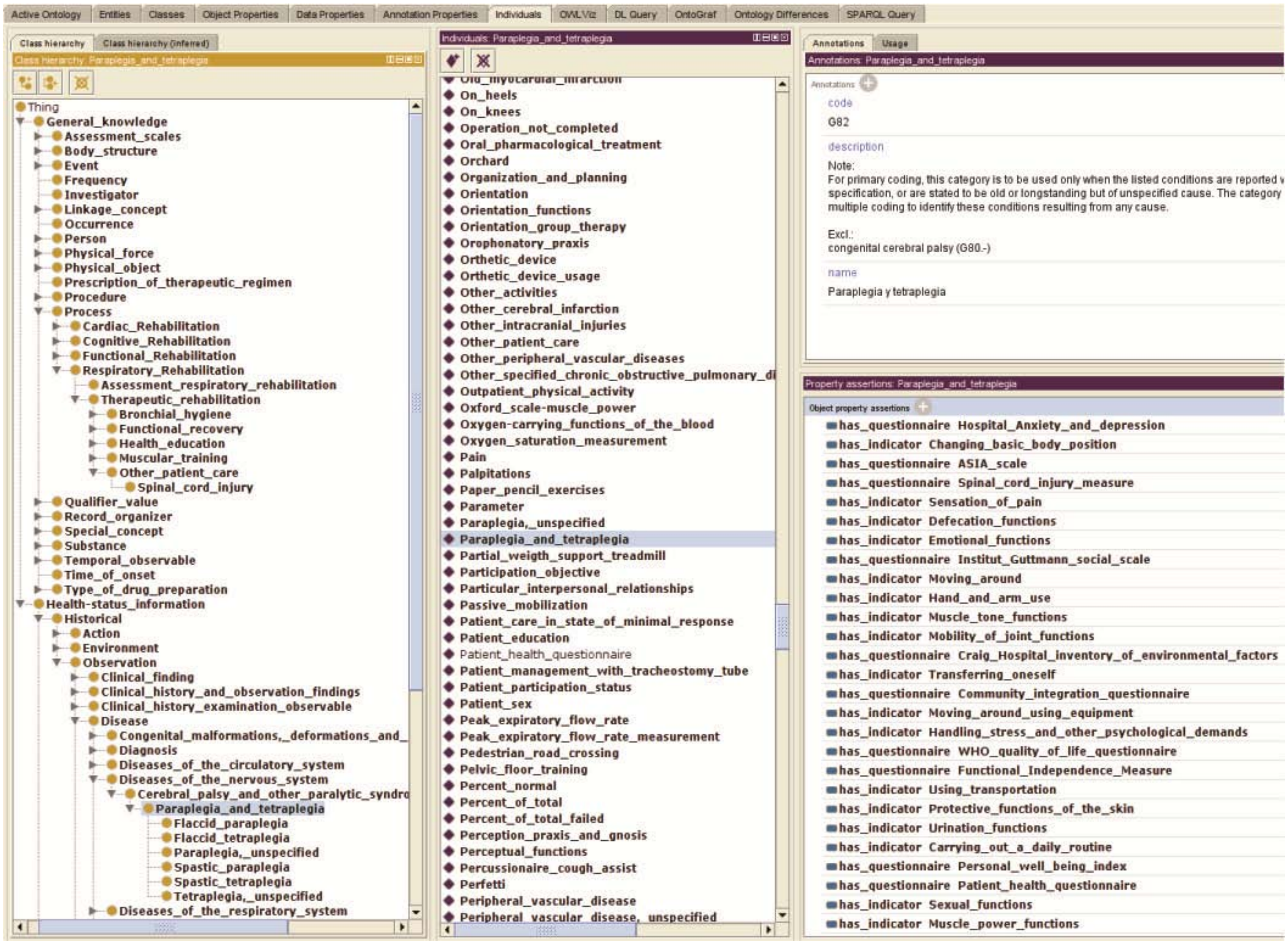


Figure 3:7 Ontological representation of people who has SCI.

The ontology and the MySQL database can be easily connected as they are both accessed through Java libraries, and concepts are represented using the same ICF and SNOMED CT standard codes. 57 new tables are created:

- 4 tables to represent user's ICF values, for example, in the case of *body functions*, in *icftableb* table;
- Table *nodenames* to help users to understand ICF categories, to which disease it can be applied, and from which scales it has been extracted;
- 5 tables to contain information of the ASIA scale;
- 3 tables to contain information of each questionnaire described in chapter 2 (except for ASIA scale that has more tables due to its subscales), for example, in the case of WHOQOL, in *whoqolquestions*, *whoqoloptions* and *whoqolanswers* tables;
- Table *profile*, to specify medical information; and
- Table *punctuationslink*, to normalize values of questionnaires.

In Figure 3:8, there is the entity relationship modeling of a sample of 7 of these tables and the table *user* (which contains user identification).

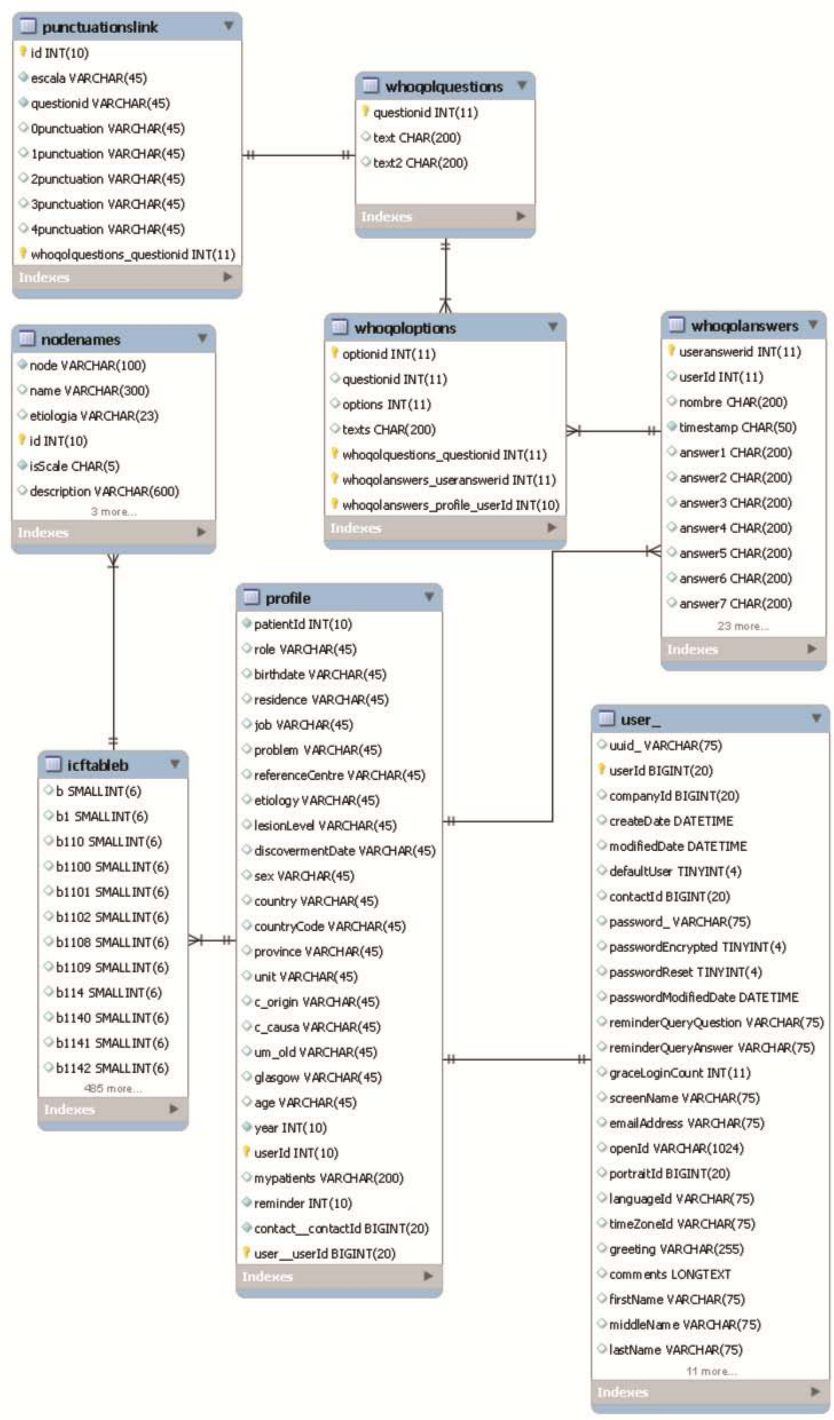


Figure 3:8 Entity Relationship Modeling.

3.3 Interoperability and information models

With respect to interoperability, to exchange EHRs in a standard way, we have taken into account the *Smart Open Services for European Patients* (epSOS), the *vMR* and the *EN/ISO 13606* initiatives. The *patient summary* is part of epSOS and includes evaluated persons' relevant health information. *Relevant information* is understood as the minimal set of personal health data (160 descriptors) that are of interest to health professionals to assist citizens, and the ignorance of which could pose a risk to the health of the evaluated person. The epSOS' main objective is to assist practitioners in unscheduled care and its general policy is to adopt international standards, such as HL7 v2.

vMR, a simplified view of HL7 v3, is still in the process of improvement and evolution by the vMR Project Team[2], and there are several CDSSs that use it (Kashfi (2011)). vMR is designed to reduce development costs and time responses in CDSSs. As a consequence, although it is not widely adopted in hospitals and there are not many tools available to facilitate its implementation (unlike the information model EN/ISO 13606), we considered it the most appropriate information model for CDSSs today. Figure 3:4 shows the mapping between the ontology and vMR objects, where the data types used are:

- *concept descriptor* (CD), a reference to a concept defined in an external code system, terminology, or ontology;
- *entity name* (EN), a name for a person, organization, place or thing;
- *timestamp* (TS), a quantity specifying a point on the axis of natural time; a point in time is most often represented as a calendar expression; and
- *interval timestamp* (IVLTS), a set of consecutive values of an ordered base data type.

These data types are a simplified version of ISO 21090 data types, which is an implementable specification based on the abstract HL7 v3 data types specification.

Table 3:4 Relationships between ontology and vMR objects.

Ontology object	Reference information model	Attribute	Data type
Name	Person	Name	EN [0..*]
Date of birth	Evaluated person	Birth time	TS[0..1]
Problem	Problem base	Problem code	CD
Event	Adverse event base	Adverse event code	CD
Date of diagnosis	Problem base	Diagnostic event time	IVLTS
Body functions, activities and participation, environmental factors and body structures	Observation base	Observation focus	CD

3.4 Summary

Interoperability problems in healthcare systems cover three issues: terminologies, ontologies and information models. This Chapter was divided in three parts: how to automatically generate interoperable knowledge using standard terminologies, how to build an ontology that establishes relationships in the disability domain, and how to relate information models with the ontology.

The generated standard and multidimensional indicators will feed both the monitoring system and the prognosis system that will be explained in the next two chapters. As a consequence, both the monitoring

[2] Available online at http://wiki.hl7.org/index.php?title=Virtual_Medical_Record_(vMR).

and prognosis systems are interoperable and they include main aspects of a person's life (development, participation, and environment) using the ICF instead of solely focusing on his or her diagnosis.

Chapter 4 Medical domain analysis

This chapter describes a monitoring system and a prognosis analysis using standardized data obtained using the methodology explained in the previous chapter. The monitoring system was tested with professionals and patients that gave their opinion. Therefore, the chapter is organized in two sections: in the first one, there is a description of the monitoring system implementation and functionality, while in the second section, there is an experimental analysis of using available standard machine learning methods for prognosis.

4.1 Monitoring system

4.1.1 Description of the monitoring system

The monitoring system is developed in Liferay Portal CE which is available under the Lesser GNU Public License (LGPL) v2.1 at absolutely no cost. Liferay is a content management system (CMS) written in Java with certain portlets preinstalled which cover most of the functionalities required in the monitoring system (see Figure 4:1, Figure 4:2 and Figure 4:3). New portlets for the automatic generation of interoperable, multidimensional and holistic indicators of MHRs, developed using Matlab. Nowadays, the monitoring system is not publicly available because the business model to ensure sustainability of the monitoring system is still under study and a patent has been published Miralles et al. (2013).

The monitoring system represents individually and comparatively people with disabilities or neurological origin. Figure 4:4 depicts the information related to the evolution of the person's health and a set of indicators as a function of the time elapsed since the detection of the lesion. Regarding environmental indicators, facilitators are represented by circles while barriers are represented by stars. Looking at Figure 4:4, users can quickly notice that this person has a mild facilitator in drugs, and no facilitator in assistive products and technology for personal indoor and outdoor mobility and transportation. Figure 4:4 depicts *person's evolution* using standardized holistic indicators. So it helps him to improve his knowledge of his health indicators. For example, he can see that he has a lack of assistive products and technology.

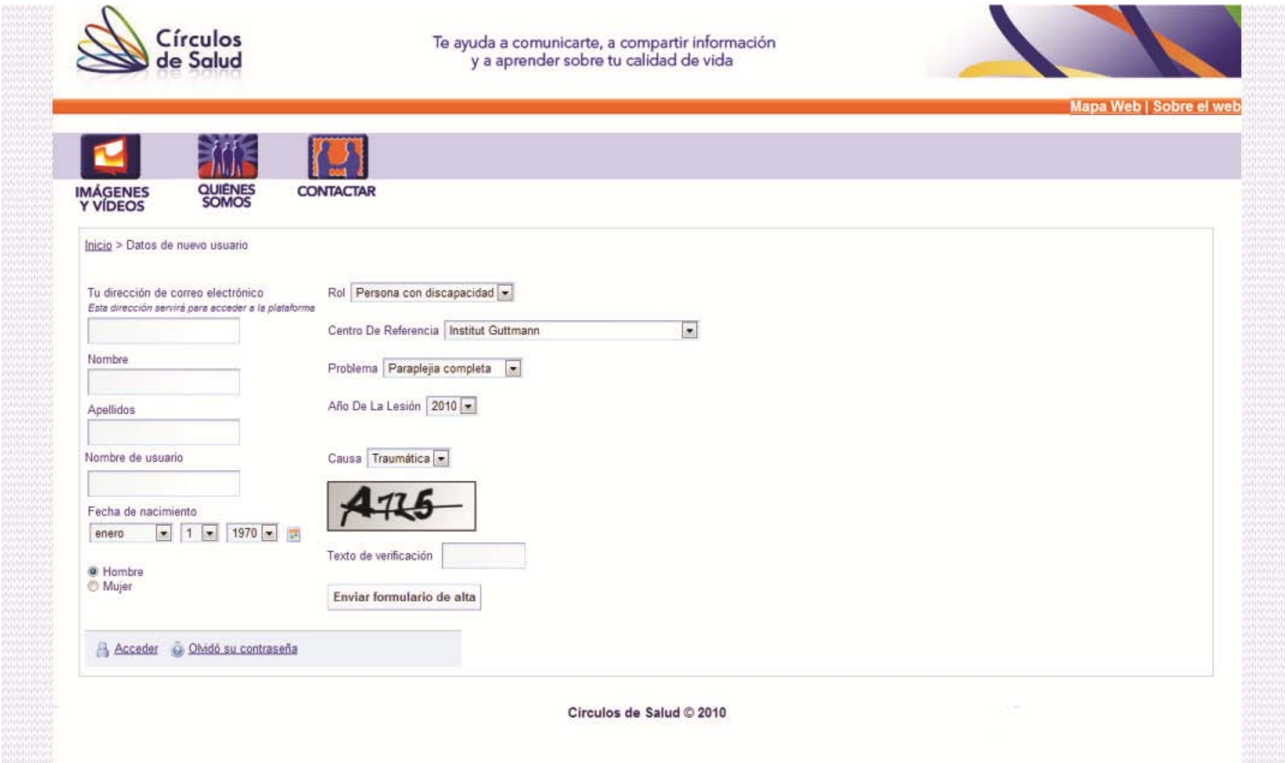


Figure 4:1 Identity management, creation of an account.

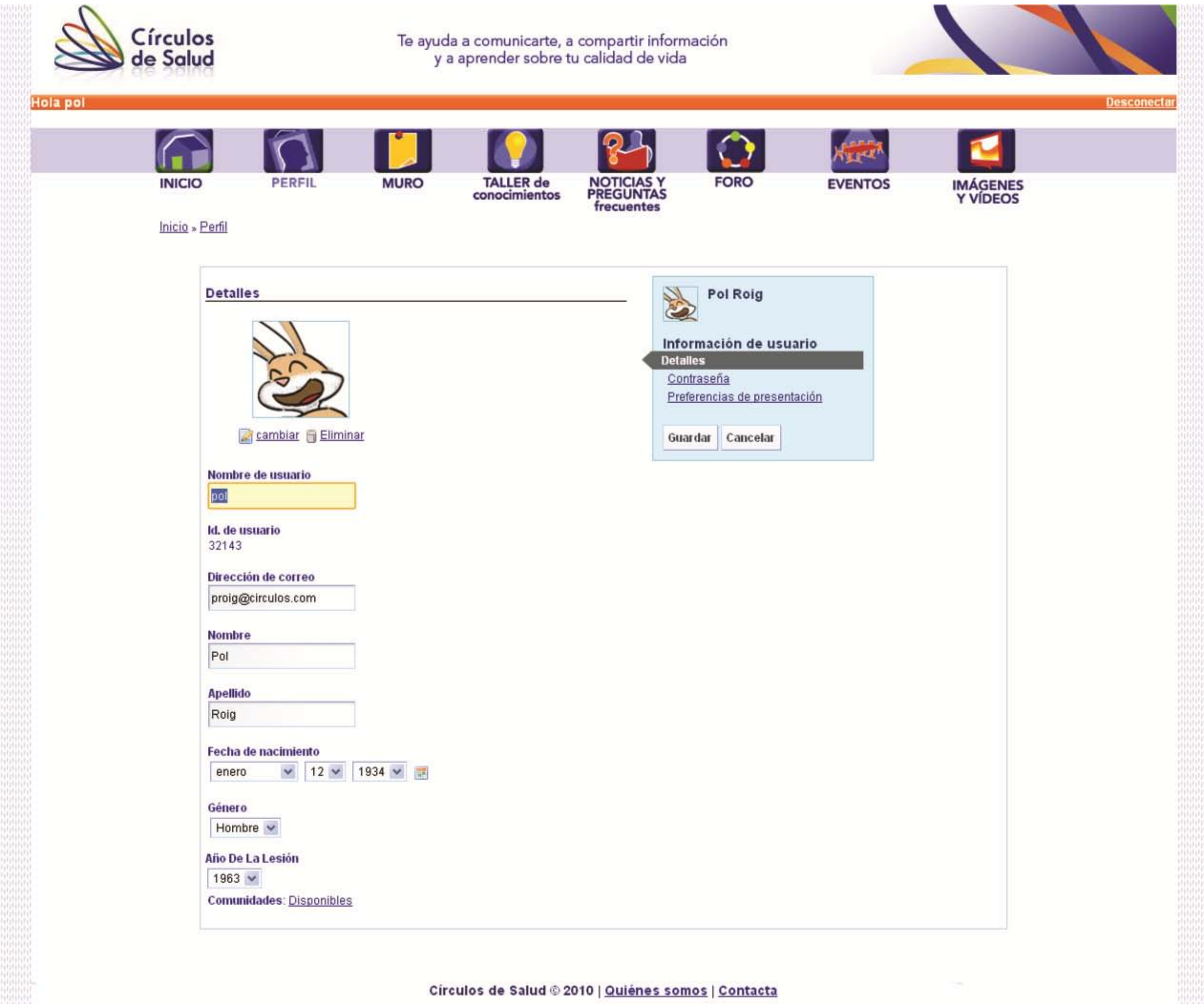


Figure 4:2 Identity management, profile.

Figure 4:3 Information exchange, management of questionnaires.

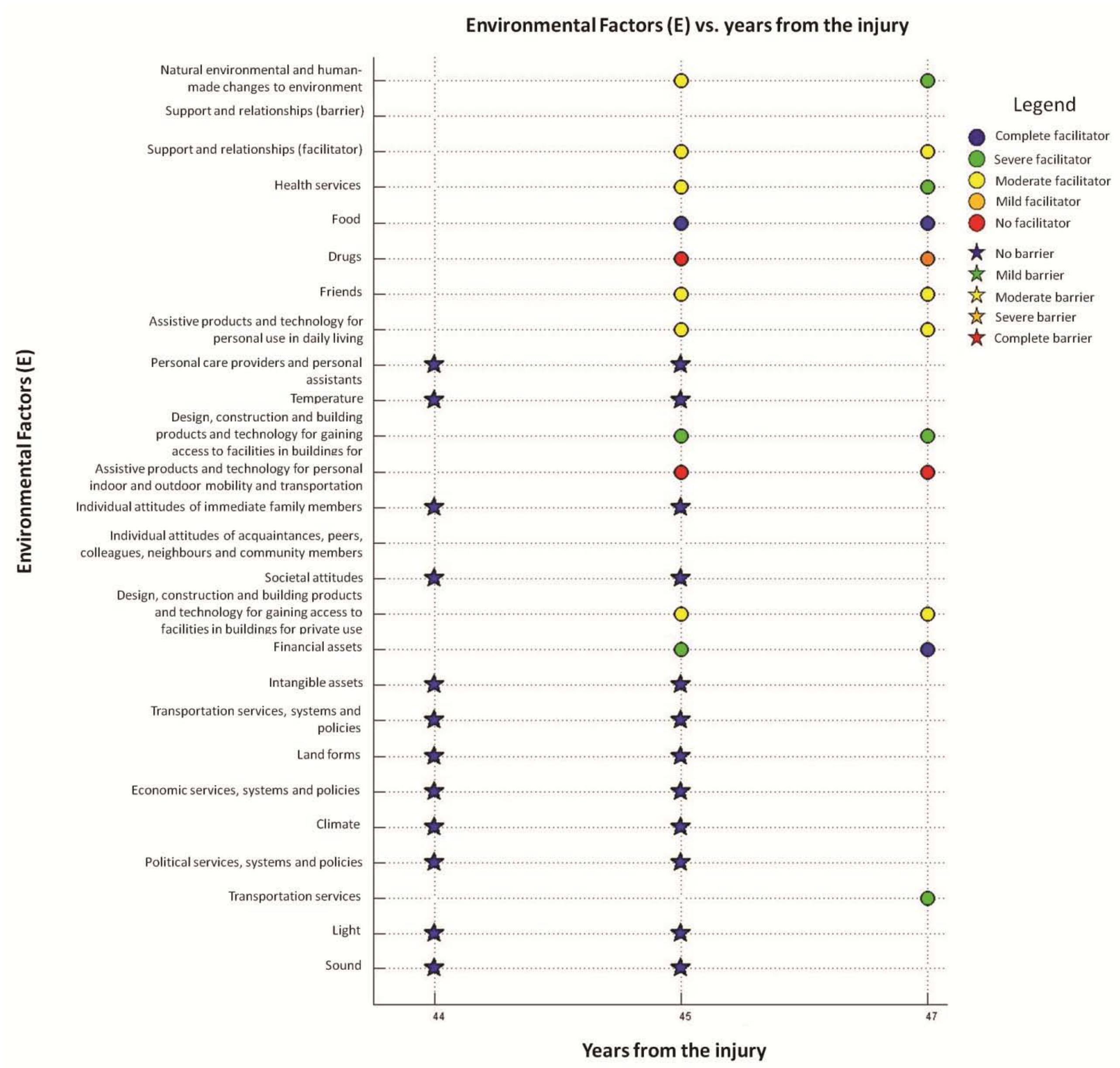


Figure 4:4 Graphical representation of the evolution of an individual across categories of the ICF. ICF values of difficulty, deficiency or barrier are represented with red/4 in complete levels, orange/3 in severe levels, yellow/2 in moderate levels, green/1 in mild levels and blue/0 in no difficulty, deficiency or environmental factors.

In Figure 4:5 there are several individual and comparative representations of users. Figure 4:6 is a *population's evolution of an ICF category*. It is yet another representation wherein a particular disease has been selected as well as a particular time span, and the MHR is depicted, in the top graph, in relation to a variety of body functions, such as urination and emotional functions. This graph shows the percentage of population with a certain value of deficiency of different normalized indicators, in a particular instant of time. Different colors indicate different percentage of patients with that deficiency. In his last periodic evaluation, his doctor recommended that he use a collector instead, but the patient argued that he prefers to perform intermittent self-catheterization. Doctors use the graphical representations to take a joint decision with Patients, showing them the percentage of patients that follow her recommendation of using a collector.

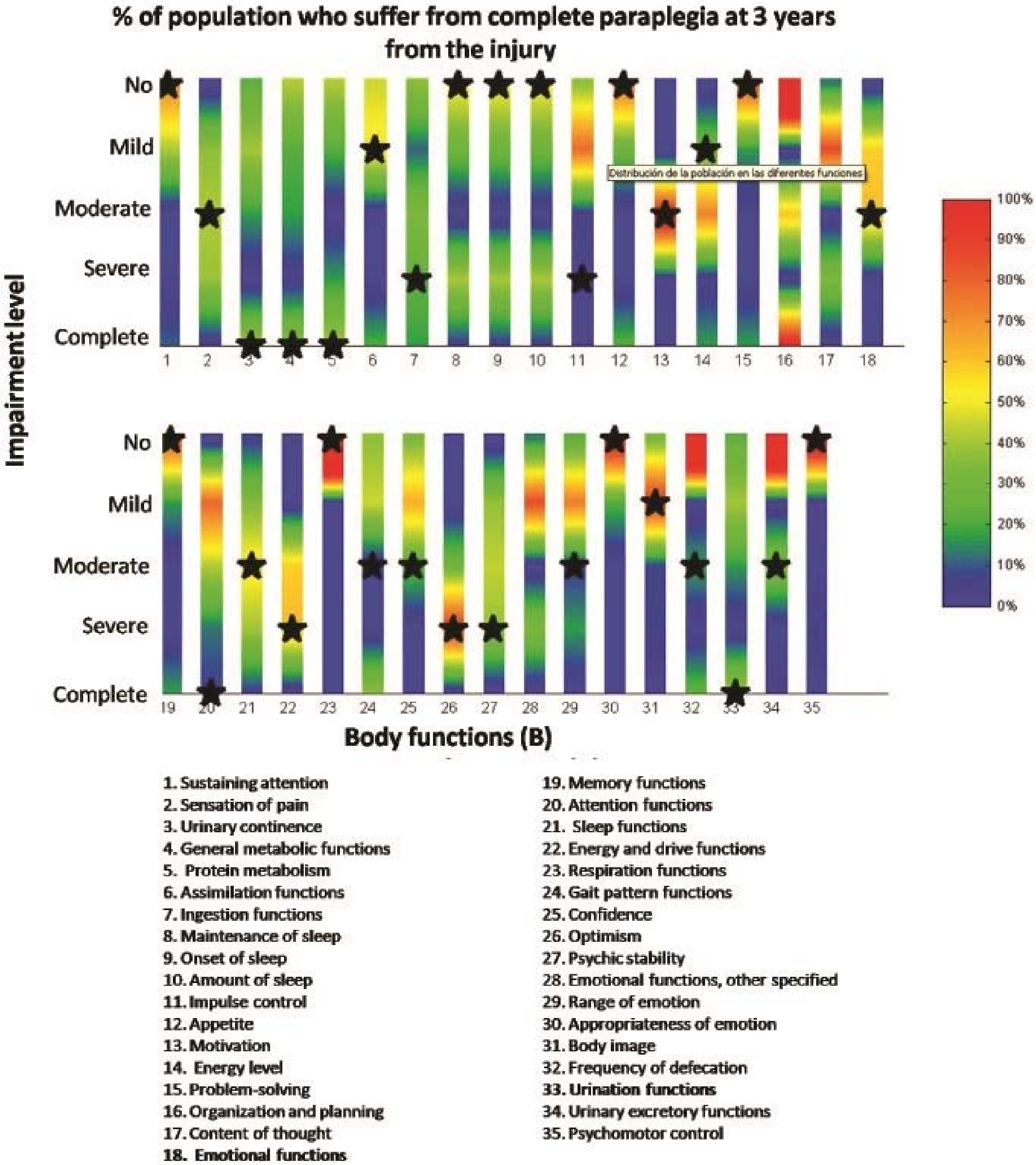


Figure 4:5 Graphical representation of the status of populations and individuals across categories of the ICF.

In Figure 4:5, person's state, represented by a star, is compared to a particular *population's state* exhibiting similar profiles. The level of functional diversity is graded in relation to the number of years from the lesion. It shows the percentage of population with a level of functional diversity of one normalized indicator in time. Different colors indicate different percentages of patients with that deficiency. When the user of presentation system is a patient or a clinician who wants to monitor the evolution of one of his patients, the patient is represented by a star. Regarding the evolution of person's emotional functions, in Figure 4:6, it gets worse (his impairment level goes from moderate to severe). This representation enables the comparison between him and population who suffer similar problems. The population is the set of people with an SCI, who had a traumatic injury and with the same time elapsed since the lesion. His psychologist sees then a graphical representation of person's emotional functions; as it is an abnormally severe deficiency, he derives him to an external psychological consultation.

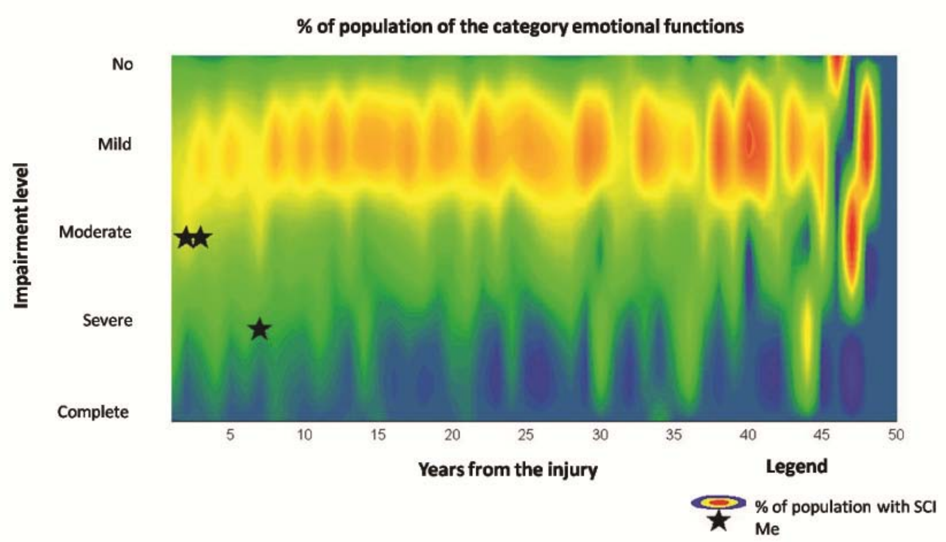


Figure 4:6 Graphical representation of the evolution of a population across categories of the ICF.

The social worker, wants to have more information about anonymous population's evolution at 5 years from the lesion. Figure 4:7 is a static representation of *population's state in SCIM*. It depicts an overview of a whole population suffering a single disease after a certain time span after the lesion, in this case 5 years. The data for this representation are extracted from the MHR and forwarded to the filter without intervention from the attribute normalizer, aggregator, or the inference engine. The level of deficiency is depicted in relation to a variety of symptoms suffered by the patient, divided into categories such as eating, hygiene, dressing, self-care, breathing, body control, physical mobility (internal/external, below/above 100m, etc.). It also shows the percentage of population with a certain value of deficiency of different indicators, in a particular instant of time. Different colors indicate different percentages of patients with that deficiency. In Figure 4:7, she can see that most people have mild facilitators for assistive products and technology for personal indoor and outdoor mobility and transportation, and moderate barriers to transportation services systems and policies and intangible assets. All this information helps her to learn that mobility is a problem for most people with this profile and that it would make a great impact to prevent it through public policies. Therefore the solution enables the social worker a straightforward, cost efficient and universal processing of health information and data, thereby allowing health practitioners and patients to ascertain their health data using common benchmarks from a wider population in a normalized and harmonized manner.

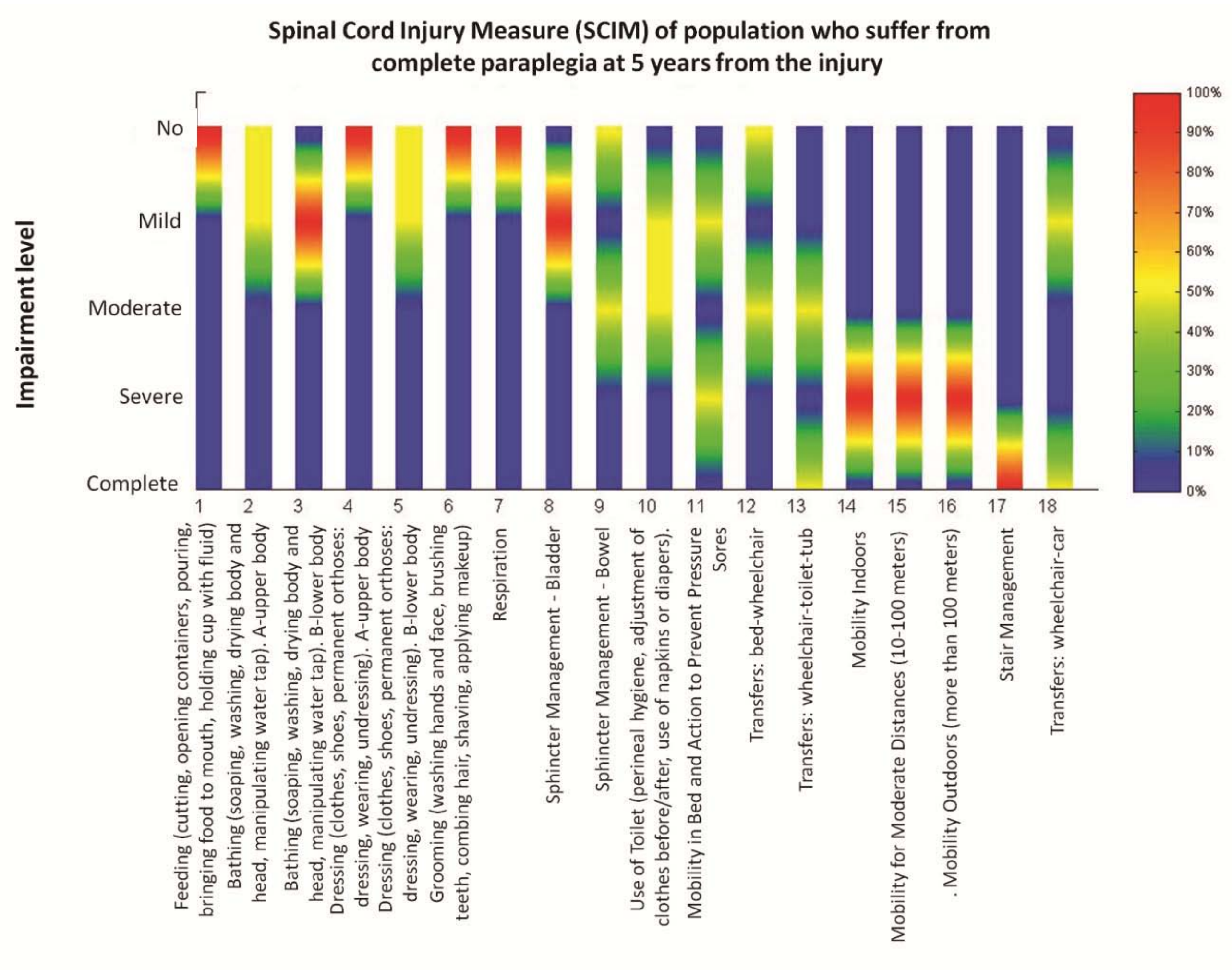


Figure 4:7 Graphical representation of the status of populations and individuals through the spinal cord injury measure (SCIM).

Recommendations of two usability and user experience experts were followed and helped to improve the usability, accessibility and user experience of the monitoring system before the analysis with users. The recommendations of the two experts were:

- Benefits of belonging to a community, which are interactions and access to useful information, should be explained.
- The monitoring system should carefully use vocabulary and help users in their daily life. Avoid in all profiles technicalities such as *Open Id*. In the profiles of person with disability and family, avoid medical jargon.
- The design should focus on user needs, and icons should be more serious. Excess of links and sections should be removed.
- Users should be able to navigate from static to evolution graphical representations to better understand QoL.
- Links to other social networks should be removed. The user does not need other social networks to interact with the monitoring system.

4.1.2 Evaluation of the monitoring system

The evaluation of the monitoring system is divided into two parts: an expert and a user experience analysis with eleven users. These users are: five health professionals (a neuropsychologist, a medical doctor, a social worker, a psychologist and a psychotherapist) and six people with a disability (SCI, non-traumatic ABI and TBI). Neither more professionals nor people with a disability have been included due to time limitations.

The protocol of verbal reports (see Ericsson and Simon (1980)) allowed generating both quantitative and qualitative results. This method is chosen because qualitative results include behavior (actions taken by the task) and literals (subjective views on the experience and interface). The main benefit of this method is to identify the user's mental model and their interaction with the monitoring system. In addition, an answer-response protocol was used to analyze user's activity and identify which parts of the interface or the system are obvious or confusing. These are the tasks of the test:

- Getting acquainted with the monitoring system. Before registering, what kind of benefits did you expect to derive from joining the community? What do you think about this page? After registering, did you find the process easy? What would you change?

- Interpretation of the content. Visit the Knowledge section. What do you think of the information provided in this section? Do you think it is appropriate? What would you improve? Visit the Scales section. Can you describe the type of content shown? Do you find it useful? Would you trust the data? Why?

- *Identity management*. It allows managing the availability of identity information (i.e. filling in information and setting access rights). Examples for functions enabling identity management are creation of an account (Figure 4:1) and profile and community memberships to accredited professionals (Figure 4:2). Regarding the last one, once professionals have already been accepted in the community of accredited professionals, they get a new menu of people management.

- *Information exchange*. Users can manage in their private pages questionnaires and see their graphical representation in a standardized format (Figure 4:3, Figure 4:4, Figure 4:5, Figure 4:6 and Figure 4:7).

During the interviews, professionals indicated the graphics provided to them a lot of interesting information, and the usefulness for the professional is clear to do multidimensional studies or articles. The monitoring system helps them to compare for example, body functions of people who had non-traumatic ABI and TBI. They found it useful to make comparisons with someone from another country that uses a different scale and/or language, to know which the most important problems are, and to organize rehabilitation objectives. Regarding to users with disabilities, they said the monitoring system should be optional; users can look at it when they want and that it should be taken into account that some old people have problems using the Internet and new technologies. Some professionals indicated that although it is useful for their daily practice, sometimes they would need to have more information (for example, if someone has urination problems, it is important to know how often the leaks happen).

During the interviews, people with disabilities said it is very interesting to be able to bring together people with a very similar profile, is always good because you feel supported. Some of them suggested the monitoring system is useful to people who had their lesion from 1 to 3 years ago. They liked the use of use of color-coding because it is very visual. Some of them liked the comparative graphic, because it helps to know where you are in relation to the rest of the population. They said it would be interesting to have these graphs in the annual review because it is very visual and it could be a standard way of representing knowledge. This can be helpful when they travel abroad because sometimes the information is in Catalan or Spanish. However, some of them indicated that it worries them to know what their deficiency level would be in some years although there were some things that at some point they would like to see. Some of them see these graphics more professionally oriented, there is too much information and ICF concepts are difficult to understand. They suggested seeing the whole individual graphical representation in the screen, and do it larger and smaller.

To summarize, results obtained are:

- All professionals have noted that graphical representations of the evolution of individuals provide valuable visual feedback, e.g., for identifying categories with a high level of deficiency with the color-codification.
- 83% of the users with disabilities gave special importance to the possibility of accessing the information at any time. The type of information exchanged by people with a disability is not necessarily medical but rather about aspects of everyday life, organizations and activities of interest.
- 66% of the users with disabilities indicated that participating in a community brings many benefits and enriches personal knowledge based on the experiences of other users. They gave special importance to knowing the people or institution who is behind the management of the content and who certifies the validity of what is shown, and to access the information when they want to.
- 33% of the users with disabilities found the monitoring system useful to perform online follow up questionnaires after clinical discharge.
- 50% of the users with disabilities had problems interpreting the standardized values and indicated that graphical representations had too much information.
- 50% of the users with disabilities indicated that graphical representations are too difficult to be interpreted, and they did not want to spend time to understand them.
- 18% of the users highlighted that the monitoring system is useful to interchange information with non-Spanish speaking countries.

Findings of usability and user experience are:

- To strengthen the explanatory text by giving priority to the need to register and privilege to participate in the monitoring system. To include who is responsible for managing the monitoring system.
- To explain in detail the purpose of filling in the register and the role of the user when checking the contents of the monitoring system.
- To add contextual help text to know the content of each section before accessing it.
- To provide information about content privacy all over the network, to inform users who is able to see their information.

- To review the format of the monitoring system with emphasis on adding the title in all sections and to add different colors in the text.

Regarding efficiency, users were satisfied with time response of the monitoring system when generating graphical representations, and introducing and standardizing questionnaires. We show the response times of the integration of Matlab and Liferay in a X3430 (2.39GHz, 3.99 RAM) processor was acceptable by users.

All professionals noted that graphical representations of the evolution of individuals provide valuable visual feedback. They also highlighted the monitoring system integrates understanding of health forming a comprehensive profile of an individual and population. Therefore, results indicate the monitoring system is useful to support multicenter studies thanks to the use of standardized knowledge.

83% of users with disabilities gave special importance to the possibility of accessing the information at any time, 66% indicated that participating in a community brings many benefits and enriches personal knowledge based on the experiences of other users, and 33% of people with disabilities indicated the monitoring system is useful to perform online follow up questionnaires. As a consequence, the monitoring system has potential to promote user empowerment and making decisions with a more informed opinion.

However, 50% of users with disabilities remarked ICF categories obtained from PCEs provide too much information to people with disabilities. In addition, 50% of them found some categories of the ICF standard difficult to understand. Therefore, it is recommended to limit the number of ICF categories for each user's interface to no more than nine. The number nine was set due to human limits in the ability of processing information (Miller, 1956).

4.2 Prognosis analysis using available standard machine learning methods

4.2.1 Description of the prognosis analysis

This section presents an experimental analysis of using available standard machine learning methods for prognosis in our four prediction domains. We chose Weka to perform the experimental analysis because it supports several learning methods we are interested in. Moreover, Weka is an open-source tool that provides a user-friendly GUI that allows evaluating the solutions. The chosen algorithms to perform the benchmark are: linear regression, instance-based learning IBk, Naïve Bayes (NB), support vector machine (SVM) and decision trees (J48). The instance-based learning k (IBk) method is a k-Nearest Neighbours (kNN) classifier.

In the tables from 4:3 to 4:6, the best values of accuracy, precision, recall, specificity and MAE are highlighted in bold. Given that in a *confusion matrix* we have true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN), we define the usual way the following measures:

$$accuracy = \frac{TP+TN}{TP+FP+FN+TN}\,;\; precision = \frac{TP}{TP+FP}\,;\; recall = \frac{TP}{TP+FN}\,;\; specificity = \frac{TN}{TN+FP}$$

We need to adapt the original representation of an instance as a vector of attribute value pairs used in kNN, NB, SVM and decision trees to our time-dependent data. Given an indicator IND, in a time T with a value V, each indicator has a time series of values: IND, (T_i,V_i) i=1,…,n. We adapt the domain data in the following way: for each triple (IND, T_i, V_i) we create one attribute for each pair (IND, T_i) with value V_i. We define a *cost matrix*

0.0	0.2	1.0	1.0	1.0
0.2	0.0	0.2	0.2	0.2
1.0	0.2	0.0	1.0	1.0
1.0	0.2	1.0	0.0	1.0
1.0	0.2	1.0	1.0	0.0

for the emotional functions prediction domain. This cost matrix is necessary to solve the class imbalance problem in this domain, where 66% of instances have one solution while the other 34% of instances have the other 4 solutions. This cost matrix was designed to give a lower weight to both FP and FN in the most frequent class. We took 0.2 because is the maximum cost weight that solved the class imbalance problem in decision trees.

4.2.2 Prognosis analysis

In this section we will present evaluation tables for each one of the four prediction domains. For each prediction domain, a table will show for each ML method its accuracy and its recall value for each Qualitative Value (0, 1, 2, 3, 4) of the prediction. Since ML methods are based on classification, the Qualitative Value (0, 1, 2, 3, 4) will be called henceforth *solution classes*, or simply *classes*.

In the next two tables, accuracy and recall values of the classes are shown without (Table 4:1) and with (Table 4:2) the cost matrix. However, no cost matrix was needed to be used neither in NB nor SVM so they are not shown in Table 4:2. Class 4 does not appear neither in Table 4:1 nor in Table 4:2 because there are not instances of class 4 in the prediction of emotional functions (see the number of instances of complete deficiency in Table 2:1)

Table 4:1 Prediction of emotional functions of people with SCI without cost matrix.

Learning	Accuracy	Recall class 0	Recall class 1	Recall class 2	Recall class 3
Linear regression	**0.66**	0	**1**	0	0
IBk (k=7)	0.65	0	0.98	0.03	0
NB	0.49	**0.43**	0.56	**0.27**	**0.40**
SVM	0.54	0.01	0.72	0.17	0.19
J48	**0.66**	0	**1**	0	0

Notice that if the cost matrix is not used, as shown in Table 4:1, methods IBk and J48 only learn to predict correctly one class (as shown by having a recall close to 1 in that class and close to 0 in the other classes). The reason is that there is a problem of class imbalance, where Class 1 has number of instances much greater than the other classes.

Table 4:2 Prediction of emotional functions of people with SCI with cost matrix.

Learning	Accuracy	Recall class 0	Recall class 1	Recall class 2	Recall class 3
IBk (k=7)	0.38	0.06	0.39	0.59	0.15
J48	0.42	0.17	0.49	0.34	0.17

Using the cost matrix, Table 4:2 shows that the methods IBk and J48 improve recall, but accuracy worsens with respect to Table 4:1.

Table 4:3 combines the two previous tables by eliminating IBk and and J48 rows in Table 4:1 by the respective rows in Table 4:2, eliminating the methods that had the class imbalance problem. Moreover, the Table 4:3 now includes those values for precision, recall, specificity and MAE for the 5 selected ML methods. Table 4:3 shows in bold the best values of accuracy, precision, recall, specificity and MAE in the prediction of emotional functions of people with SCI.

Table 4:4 shows the prediction of urination functions in SCI. Table 4:5 shows the prediction of emotional functions of people with ABI. Finally, Table 4:6 shows an example of the prediction of executive functions of people with ABI.

Regression cannot be compared directly with the other methods, because its output is a number instead of a solution class. In other words, there is no confusion matrix for regression we cannot calculate accuracy, precision, recall and specificity directly. In order to be able to compare regression and the other methods in a uniform way, we create a confusion matrix for regression by rounding the predicted value to the closest QV in (0, 1, 2, 3, 4).

Now, the tables show that regression has a class imbalance problem in all prediction domains and accuracy has very low values. Class imbalance can be observed in Table 4.1 where regression predicts all examples to be in one class (QV 1). A similar class imbalance with respect to one of the solution classes occurs in the rest of the domains, and although we do not show these in detail, the effect over accuracy is clear, since it is always very low.

On the other hand, if we compare regression using the MAE measure, we see the error is lower than the other methods. This is because regression output is a real number, not one of the five qualitative values we are interested in predicting. However, as we have explained, when the output is approximated to the qualitative values the result is worse, as we can see with the accuracy, recall and precision measures in which have worse results. In conclusion, although regression takes into account the fact that our solutions are ordered (while the other methods do not), it assumes the intended prediction to be a real value and when applied, as in this case, to five qualitative values, is not better than the alternatives, even when it has better evaluation in some data with the MAE measure.

For this reason, in the next chapter we will compare our proposed method with the classification methods (using accuracy, precision, and recall) but not regression (since the approximation to qualitative values needed to use accuracy, precision, and recall gives poor results).

Table 4:3 Prediction of emotional functions of people with SCI.

Learning	Accuracy	Precision	Recall (or sensitivity)	Specificity	MAE
Linear regression	*0.66*	*0.44*	*0.66*	*0.35*	*0.59*
IBk (k=7)	0.38	0.54	0.38	**0.70**	0.76
NB	0.48	**0.57**	0.48	**0.70**	0.81
SVM	**0.54**	0.56	**0.54**	0.60	**0.72**
J48	0.42	0.53	0.42	0.62	0.76

Table 4:4 Prediction of urination functions in SCI.

Learning	Accuracy	Precision	Recall (or sensitivity)	Specificity	MAE
Linear regression	*0.04*	*0.001*	*0.04*	*0.96*	*1.19*
IBk (k=7)	0.70	0.67	0.70	0.64	**1.14**
NB	0.77	0.78	0.77	**0.88**	1.51
SVM	0.78	0.77	0.79	0.83	1.40
J48	**0.84**	**0.79**	**0.84**	0.84	1.36

Table 4:5 Prediction of emotional functions of people with ABI.

Learning	Accuracy	Precision	Recall (or sensitivity)	Specificity	MAE
Linear regression	*0.27*	*0.01*	*0.27*	*0.73*	*1.25*
IBk (k=7)	0.33	0.35	0.33	0.82	1.12
NB	0.37	0.38	0.37	0.79	1.21
SVM	**0.41**	**0.41**	**0.41**	**0.85**	1.11
J48	0.38	0.34	0.37	0.84	**1.07**

Table 4:6 Prediction of executive functions of people with ABI.

Learning	Accuracy	Precision	Recall (or sensi-tivity)	Specificity	MAE
Linear regression	*0.22*	*0.05*	*0.22*	*0.79*	*1.25*
IBk (k=7)	0.32	0.30	0.32	0.82	**1.63**
NB	**0.43**	**0.42**	**0.43**	**0.86**	1.70
SVM	0.40	0.41	0.40	0.85	1.68
J48	0.42	0.34	0.42	0.84	1.87

In conclusion, we can see that there is no method better than all others, neither in accuracy nor in MAE, for all 4 prediction domains. In the next chapter, we will compare the best results of accuracy and MAE with our proposed method.

4.3 Summary

In this chapter we have presented a monitoring system that uses a new methodology that automatically collects, transforms, shares and graphically represents standardized and multidimensional indicators. The monitoring system includes the main aspects of a person's life (development, participation, and environment) using the ICF instead of solely focusing on his or her diagnosis. Results indicate the monitoring system has great potential for users. Professionals found the monitoring system useful to generate new knowledge. People with disabilities found helpful to enrich personal knowledge with the experiences of other users and to perform online follow up questionnaires after clinical discharge. As a consequence, the monitoring system has potential to promote user empowerment and making decisions with a more informed opinion. For future work, it would be interesting to establish no more than nine ICF categories for each user's interface and to involve other institutions.

In this chapter we have also evaluated a prognosis analysis using Weka. These results will be a baseline to compare our temporal case-based method in the next chapter.

Chapter 5 Temporal case-based prognosis

Time plays an important role in medicine, therefore health professionals want to see the evolution of attributes as a function of time elapsed from the injury, as described in previous chapters. However, the best way to make use of temporal data may depend significantly on the domain characteristics. In some domains, taking into account the data of previous year for a patient can provide enough information for an accurate prediction on the current year. In other domains, the long-term evolution (from the year of the patient's injury up to the current year) may be necessary for an accurate prediction on the current year.

According to the survey of Esling and Agon (2012) on time-series data mining, there are three major design choices to make: 1) a data representation of time-series, 2) a similarity measure between time series, and 3) an indexing method. In this chapter we describe and evaluate techniques to address the first two issues[3] in the context of our prognosis system.

This chapter studies the role played by time in our four different prediction domains from a case-based reasoning point of view. In previous chapters, the evolution over time of the population was visualized and the relation of a target patient with the data of the whole population was also visualized.

This chapter changes this perspective, following the paradigm of CBR. In this perspective the "target patient" is related to a small set of similar patients (instead of the whole population of patients). We will introduce the concept of similar patients of a target patient taking into account the key temporal patterns of patient's attributes.

This chapter presents our approach, implemented in the CAse-based Prognosis using Temporal Abstractions (CAPTA) system, and is structured as follows. First we give an overview of our approach, detailing the case description and the temporal representation that we will use. Then we propose a similarity measure for cases with our temporal representation. This similarity is used by the CAPTA system to find the most similar patients to the current patient. Then we describe the experimental evaluation of our approach in our 4 domains. Finally, we report our results and compare the evaluation of our prognosis system with a baseline of existing methods.

This chapter also describes a visualization of CAPTA's temporal case-based prognosis that can be used to support clinical decision making by capitalizing on the main features provided by CAPTA: the use of similar cases and their temporal behavior. Nowadays several visualizations of CDSS exist that take into account time as explained in the state of the art chapter. Our visualization approach focuses on how to capitalize the main features provided by our temporal case-based approach to prognosis. This chapter describes an interface based on CAPTA that helps clinicians to visualize the most similar person (or persons) given a target person and the temporal behavior of different attributes. We provide some screenshots of the visualization of real patients.

[3] The third issue, the indexing method, is not studied here because our time-series are not very long, alleviating the need to optimize for space consumption and computational efficiency.

5.1 Approach

5.1.1 Description of a case

A case is classically represented in CBR as a pair (*p,s*) composed of a problem (*p*) and its solution (*s*). In our approach a case has the following information:

- The problem (*p*) is composed of both temporal and static data. The notation we use for *temporal* or *clinical data* is $Z^A_{a_i,t}$ where *A* is the patient, a_i is an indicator (or attribute) and *t* is the instant of time. The notation for *static* or *demographic data* is $Z^A_{a_i}$, and is independent of time. A *problem p* is a collection of temporal data organized in *states*, where a *state* at a time *t* is the collection $Z^A_t = (Z^A_{a_1,t},...,Z^A_{a_m,t})$ and where the static indicators have the same value in each state.
- The solution *s* is a predicted value for specific attribute in each domain, which in our domains correspond to the following attributes or indicators: emotional functions, urination functions, and executive functions. The prognosis in a domain is the prediction of the value of that attribute *s* in the next state of the patient. The notation $Z^A_{s,n}$ indicates the prediction for the solution indicator with time series $Z^A_{s,n-1}$ where n-1 is the length of the time-series.

5.1.2 Temporal representation

We introduce four temporal representations used for the prognosis of neurological functions. Temporal abstractions (TA) are used to interpret intervals of temporal sequences of past and present indicators in order to predict the value of emotional and urination functions. Intervals of temporal sequences are interpreted as series of changes, where the set of changes are: I^g_n, D^g_n, and S^g, where *I* means *increase*, *D* means *decrease*, *S* means *stationary*, *n* is the amount of change, and *g* is the granularity of the change in an interval. Considering that ICF has 5 possible values 0, 1, 2, 3 or 4 (where 4 means complete deficiency and 0 means no deficiency) the “amount of change” can vary from 1 to 4.

Three different granularities are proposed: fine (f), medium (m) and coarse (c) (see Figure 5:1). In the *fine* granularity, all 4 levels of change are differentiated; in the *medium* granularity, 2 levels, and only 1 level of change can be differentiated in the *coarse* granularity. We have these granularities because values can be subdivided in 3 groups and smaller indistinguishable entities are joined together to become larger distinguishable entities. In the tree of Figure 5:1, leaves are the smaller indistinguishable entities and belong to the fine granularity. There are 9 leaves because these are the possible change values from a (sub)sequence of states. Therefore, for an attribute a_i, the change at granularity *g*, denoted by $C^g(Z_{a_i,t}, Z_{a_i,t'}) = C^g_n$, is computed as the subtraction of a preceding state value from a subsequent state value. If the difference is negative then the change is denoted by D_n (decrease), with subindex n being the amount of the difference within granularity *g*. If the difference is positive then the change is denoted as I (increase), with subindex n being the amount of the difference within granularity *g*. Finally, if the difference is 0, it is denoted as S (stationary).

Given an attribute a_i time series, we can then calculate the changes, obtaining a composed of the values D, I, and S at a granularity *g*. When granularity is fine, as show in Fig. 5:1, the change symbols C_n^g have subindex n={1,..., 4} corresponding to the difference $Z_{a_i,t} - Z_{a_i,t'}$ (when difference is 0 then change is denoted by S).

Each two consecutive nodes in the tree labeled *D* or *I* have a parent in the level above. Therefore, when there is an absolute change of 1 or 2 we will use the notation C_1^m and when the absolute change is greater than 2 we use the notation C_2^m . Finally, in the upper level the amount of change is not specified and only it is indicated if the difference is negative D^c , stationary S^c or positive I^c .

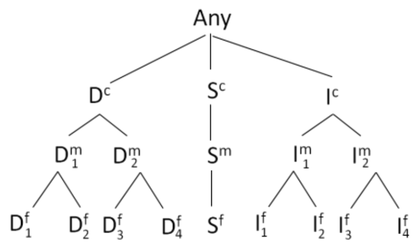


Figure 5:1 TA granularities.

CASE ABSTRACT TEMPORAL REPRESENTATIONS

Recall that a *case* with *n* states is *originally* represented as a sequence (Z_0, Z_1, … , Z_n), where the *case problem* is the sequence Z_0, Z_1, … , Z_{n-1} and the *case solution* is the state Z_n. However, there a several ways in which temporal information can be represented, and hereby we propose four temporal representations:

1. *Original*. A case problem is a collection of temporal indicators and states; where the temporal indicators are represented by the states in the series Z_0, Z_1, … ,Z_{n-1} and where the static indicators have the same value in each state.
2. *Previous state.* A case problem is a collection of temporal indicators and states; where a temporal indicator *a* is represented by the last state in the series Z_{n-1} alone (regardless of previous states and where the static indicators have the same value in each state).
3. *Long-term change.* A case problem is a collection of temporal indicators and states using TA; where the temporal indicators are represented by sequence: (Z_0, $C^g(Z_0, Z_{n-1})$), i.e. the initial state Z_0 and the overall change from beginning to end, i.e. from Z_0 to Z_{n-1}. Moreover, different granularities *g* can be chosen for C^g. Static indicators have the same value in each state.
4. *State changes*. A case problem is a collection of temporal indicators and states using TA; where the temporal indicators are represented by the sequence (Z_0, $C^g(Z_0, Z_1)$, … , $C^g(Z_{n-3}, Z_{n-2})$, $C^g(Z_{n-2}, Z_{n-1})$), i.e. the initial state Z_0 and a sequence of pairwise changes. Moreover, different granularities *g* can be chosen for C^g. Static indicators have the same value in each state.

We use these four types of case representation in CAPTA.

An example of the 4 temporal representations

The following examples will illustrate the four different temporal representations at the level of indicators or attributes whose values change along time.

An example of the *original* temporal representation, which we have obtained from PCEs, is the following two sequences of values for an attribute *a* in two cases *A* and *B*:

Year of measurement / Attribute *a* of Case	2009	2010	2011	2012	2013	2014
A	3	2	0	1	2	?
B	4	3	1	2	3	?

If we represent these two temporal sequences of the attribute *a* in our four temporal representations with granularity *fine* we have the following:

Temporal representation / Attribute *a* of Case	Original	Previous State	Long-Term Change	State Changes
A	3, 2, 0, 1, 2	2	3, D_1	3, D_1, D_2, I_1, I_1
B	4, 3, 1, 2, 3	3	4, D_1	4, D_1, D_2, I_1, I_1

The representation using Previous State in the example is simply the attribute's value of the year 2013, as 2014 is the year to be predicted. So, in patient A, the value of the previous year is 2, the value for year 2013.

The representation using Long Term Change has two parts, first the value of the first year where the data is available, 2009, and second the change from the first year (2009) to the last year before the prognosis (2013). So, in patient A, the representation is (3, D_1), since the attribute's values are 2 and 3, and the change 2-3=-1 corresponds to D_1.

The representation using State Changes has the initial value and all the pairwise changes. In patient *A,* we have the following changes: (2-3), (0-2), (1-0), and (2-1). Therefore the resulting State Changes representation is: (3, D_1, D_2, I_1, I_1).

5.1.3 Case temporal structure comparison

We later present similarity measure *sim(A, B)* between 2 cases A, B. Since these cases have a temporal structure, we will need to define some operations that are able to compare them.
Thus, to compare two TA sequences from two cases for an attribute *a*, we we will use the following

operations: (1) *Allen's interval relations* and (2) *longest common subsequence*.

Allen's interval relations

Allen's approach to reasoning about time is based on the notion of time intervals and binary relations among them. A time interval X is an ordered pair $[x^-, x^+]$ where x^- and x^+ are interpreted as points on the real line. Given two time intervals, their relative positions can be described by exactly one of the elements of a set of thirteen basic interval relations, where each basic relation can be defined in terms of its endpoint relations (see Table 5:1).

Table 5:1 Allen's temporal interval relations between intervals $X=[x^-,x^+]$ and $Y=[y^-,y^+]$.

Basic interval relation	Symbol	Endpoint Relations	Pictorical example
X before Y	b (X,Y)	$x^+<y^-$	xxx yyyy
Y after X	bi (X,Y)	b (Y,X)	
X meets Y	m (X,Y)	$a^+=b^-$	xxxx yyyy
Y met by X	mi (X,Y)	m (Y,X)	
X overlaps Y	o (X,Y)	$x^-<y^-$ and $x^-<y^+$ and $x^+<y^+$	xxxx yyyy
Y overlapped-by X	oi (X,Y)	o (Y,X)	
X during Y	d (X,Y)	$y^-<x^-$ and $x^+<y^+$	xxx yyyyyyyyy
Y includes X	di (X,Y)	d (Y,X)	
X starts Y	s (X,Y)	$x^-=y^-$ and $x^+<y^+$	xxx yyyyyyyyy
Y started-by X	si (X,Y)	s (Y,X)	
X finishes Y	f (X,Y)	$x^+=y^+$ and $y^-<x^-$	xxx yyyyyyyyy
Y finished by X	fi (X,Y)	f (Y,X)	
X equals Y	e (X,Y)	$x^-=y^-$ and $x^+=y^+$	xxxx yyyy

LCS in abstract time series

Let us consider two new operations between two time series, that we will need for describing case problems when using the State Changes temporal representation (recall that the initial value is not included in the sequence of changes). We will now use *longest common subsequence* and *temporal intersection* operations among time series (Hirschberg, 1977)[4].

Given two case problems represented as a collection of attribute-wise time series, we use a *longest common subsequence* (LCS) for each attribute that is time-dependent. Notice that the LCS problem is often defined to be finding *all* common subsequences of a maximum length; in our approach we will find just the *rightmost longest common subsequence* (RLCS). The reason to select the rightmost LCS is because it is clos-

[4] The following Java library has been used http://introcs.cs.princeton.edu/java/96optimization/LCS.java.html.

er to the predicted value. Therefore, the RLCS for an attribute *a* is an element of the set *LCS(A,B,a)* of longest common subsequences (LCSs) of *A* and *B*, chosen as follows: if there is more than one LCS, pick the one that starts at a later position in the timeline of *a* from the year of the injury; if there is more than one, the one in which the next character starts the latest, and so on. Since we are working in an abstraction of time series given by TA, we finally obtain $RLCS(A,B,a)=(C_1, \dots ,C_m)$ where each C_i is a symbol in TA.

Moreover, as we will explain later, we will associate a weight γ(C) to each C in TA .

Case temporal structure comparison with Allen's IR and LCS

For the comparison of cases we will use, for each temporal attribute *a*, the relations defined in Allen's interval algebra plus the LCS.

In a case problem, each time-dependent attribute *a* is represented by a sequence (or string) of Allen's intervals calculated from (a) the date of the patient injury, and (b) the change sequence $(C^g(Z_0,Z_1), \dots , C^g(Z_{n-3},Z_{n-2}), C^g(Z_{n-2},Z_{n-1}))$.

Suppose we need to compare cases *A* and *B.* We first calculate the RLCS of the change sequences for each attribute *a* of *A* and *B* (ignoring the initial state). Next, we compute the shortest time intervals (each measured from the start of the corresponding injury) that contain the RLCS sequence; let's call them *X* and *Y*. Finally, we compare *X* and *Y* to find which one of Allen's relations *p* holds in *p(X,Y)*. The specific relation *p* is the result of this comparison.

An example

In the next table below we illustrate how to use Allen's temporal interval relations to compare two case problems in the *State changes* representation.

Years from the diagnosis of the injury / Attribute *a* of Case	18	19	20	21	22	23
A	3	**D_1**	**D_2**	**I_1**	**I_1**	-
B	-	4	**D_1**	**D_2**	**I_1**	**I_1**

In this example, we first calculate the RLCS of the change sequences (disregarding the initial value) of A and B shown in the table, shown in bold in the table, which yields the RLCS as (D_1, D_2, I_1, I_1). Note that the RCLS is not necessarily made up of consecutive values (although this is indeed the case in this example). Next, we compute the rightmost time intervals X and Y for each case that contain the RLCS. In the example, these are X=[19,22] and Y=[20,23]. Finally we determine the Allen relationship between X and Y. In this example it's the *overlap(X,Y)* relation. From now on, for brevity's sake, we will say that "the Allen's relation *overlap* holds in the RLCS of attribute *a* for cases A and B", and we will use the notation *AllenRLCS(A,B,a)=overlap*.

5.2 Case-based reasoning

Given an input problem *A* and a solved case *B* from the case base, the proposed similarity over their problem descriptions is computed taking into account three aspects: the *similarity on static attributes*, the *similarity of the temporal intersection* and the *similarity of TA sequences*.

We will first define the temporal intersection of time series, then we will define the temporal distance (td) of an attribute *a*, and finally we will define the overall CAPTA similarity measure over cases.

Temporal intersection

For two time series we define their *temporal intersection* T as the two sub-series whose items have the same distance to the year of the injury. As a consequence, for each item of a time series we define a pair $(L_t, Z^A_{a_i,t})$ where a_i is an indicator and L_t is the temporal distance from *t* to the year of the injury. In the example below the two sub-series of the temporal intersection are the values shown in bold.

Years from the diagnosis of the injury / Case	18	19	20	21	22	23
A	3	**2**	**0**	**1**	**2**	-
B	-	**4**	**3**	**1**	**2**	3

Time-shift in temporal intersection

Given two time-series, we can calculate their temporal intersection with respect to different *time-shifts*. Given a time-series Z=(L_1, Z_1), ... ,(L_{n-1}, Z_{n-1}), a time-shift ∂ applied to Z yields a series W=($L_{1+\partial}$, Z_1),...,($L_{n-1+\partial}$, Z_{n-1}).

Therefore, a time-shift ∂ applied to a time series yields a new (shifted) series, where each state is shifted as follows: $W^A_{a_i,t} = Z^A_{a_i,t+\delta}$. In the example above, a time-shift ∂=1 would change patient A having values from years 18 to 23, to having values from years 19 until 24.

In our experiments, we will consider time-shifts ∂ from 1 to 10 years, and we will use the notation Δ= {1, … , 10}. Time-shifting allows us to compare the temporal intersection of two time series of an attribute in a more flexible way with respect to the year of the injury for each patient, and to find larger intersections that will be used in the similarity measure *sim*. Without time-shifting, the intersections considered would depend rather strictly from the year the injury for each patient, and we could only compare in practice *isochronous patients* (patients that have the same quantity of years since the injury), since non-isochronous patients tend to have much smaller intersections.

Temporal distance (td)

The temporal distance *td(A,B,a)* of on attribute *a* for two patients A and B measures the distance of two intervals with respect to *RLCS(A,B,a).* Let (a_i, a_f) and (b_i, b_f) be two intervals, where a_i, b_i are the points where the RLCS starts in each patient (these points are measured as the years from the injury), and a_f, b_f are the points where the RLCS ends in each patient (see Fig. 5:2). We define the temporal distance as $td(A,B,a)= |a_f-b_f|$.

Fig. 5:2 shows an example where $b_f>a_f$. Notice that the closer the final points are, the smaller the distance, regardless of the continuity is time of the RLCS sequence.

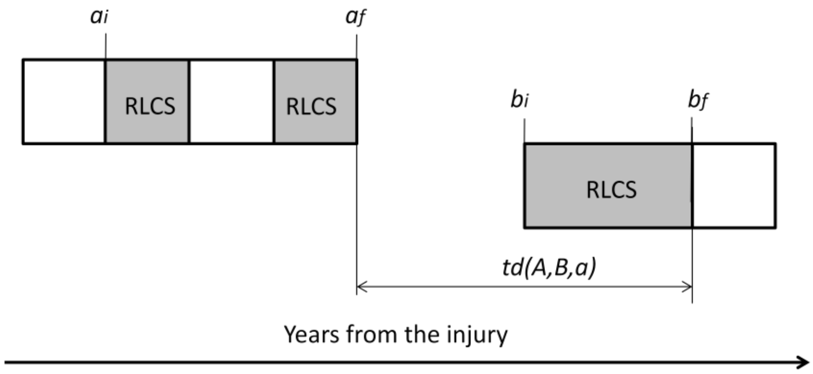


Figure 5:2 Temporal distance for attribute a with respect to RLCS.

The CAPTA similarity introduced below uses both concepts of temporal intersection and time distance.

CAPTA Similarity

We propose to have a weighted similarity, with weights α_i, that is the weight of the indicator; $\beta_{AllenRLCS(A,B,ai)}$, that is the weight of the Allen relation of the RCLS of *A* and *B* of the indicator a_i; and γ*(C)* that is the weight of the TA symbol contained in the RLCS (see below).

$$sim(A,B)=\sum_{a_i \in A}\alpha_i\left\lfloor \max_{\delta\in\Delta}\left[\sum_{t\in T}s(Z^A_{a_i,t},Z^B_{a_i,t+\delta})\right]+\left[1-\frac{td(a_i)}{R}\right]\beta_{AllenRLCS(A,B,a_i)}\left(\sum_{C\in RLCS(A,B,a_i)}\gamma(C)\right)\right\rfloor$$

Equation 5:1 – CAPTA similarity

where

1. a_i is an indicator in A={a_1, …, a_i, …a_m},
2. α_i is the weight of the indicator *i,*
3. ∂ is a time shift value in Δ, (in our approach 10 years is the maximum, i.e. Δ= 1, … ,10).
4. *t* is the number of years since the injury, and ranges over T=(ti,tj), where T is the temporal intersection of A and B
5. $Z^A_{i,t}$ is the state of indicator *i* of a case *A* at time *t,*
6. $s(Z^A_{a_i,t},Z^B_{a_i,t})$ is an attribute-wise normalized similarity at a time *t* This similarity is computed from a

distance measure $s(Z^{A}_{a_i,t}, Z^{B}_{a_i,t}) = 1 - d(Z^{A}_{a_i,t}, Z^{B}_{a_i,t})$. There are several distance measures that we use depending on the type of value in each attribute:

6.1. Numerical Attributes: absolute distance divided by the size of the value range (like age, years from diagnosis, etc.);

6.2. Categorical attributes: we use a Boolean match (e.g. for gender, cause, disease, completeness, ASIA scale or neurological level); and

7. *td* is the time distance of the indicator a_i between the time of the last value of the *RLCS(A,B,*a_i*)* and the time of the last value of the *RLCS(A,B,*a_i*)*,
8. *R* is the maximum of *td* among all possible pairs of patients and indicators,
9. $\beta_{AllenRLCS(A,B,a_i)}$ is the weight of the Allen relation that holds on the RLCS of *A* and *B* of the indicator a_i,
10. *γ(C)* is the weight of the temporal abstraction (TA) symbol that is contained in a subsequence *s* (the 9 temporal abstraction symbols in a fine granularity are shown in the leaves of the tree of Figure 5:1 plus the 5 possible values from 0 to 4). Except when saying the contrary, the fine granularity is used.

The solution is the next value of the RLCS of the attribute that is wanted to be predicted of the patient with higher similarity.

1. If the next value of the RLCS is not available, then the value that follows is used, and so on.
2. If there are not subsequent values available, then the next LCS is chosen.
3. If there are not LCSs left, then the algorithm gets the next patient with higher similarity, and so on.

Afterwards, let B(k) be the set of k more similar cases to A using *sim*. The solution of the CBR system for A is the mode of the solution values in B(k).

The computation of weights *α*, *β* and *γ* is performed using a genetic algorithm, and we performed the experiments for K = 5, 7 and 9. Finally, the accuracy, precision and recall are computed using 10-fold cross validation. We have experimented with K = 5, 7 and 9, where 7 gives the best results, so the experimental evaluation shown later in section 5.3 are those corresponding to k=7.

The parameters of the genetic algorithm are[5]:

- **Initialization**: 20 initial individuals are randomly generated.
- **Fitness**: We have experimented with two different fitness functions: (1) maximizing accuracy and (2) and minimizing mean absolute error (MAE). The two experimental results are explained below.
- **Selection**: The used method is tournament selection, where there are several tournaments among 5 individuals chosen at random from the population. The winner of each tournament is selected for crossover.
- **Crossover**: During crossover we create new individuals by combining aspects of 2 selected individuals with a uniform rate of 0.5.
- **Mutation**: Each entry has a probability 0.015 of being mutated. In the second step, the algorithm replaces each selected entry by a random number selected uniformly from the range [0-1] making very small changes at random to an individual's genome.

Weights *α*, *β* and *γ* are the weights of the fold with highest fitness.

[5]The following Java library has been used http://www.theprojectspot.com/tutorial-post/creating-a-genetic-algorithm-for-beginners/3.

Particular cases

Although CAPTA similarity uses both initial and state changes temporal representations, it can also be used with other temporal representations such as previous year and long-term change. In the CAPTA previous year approach, the similarity is computed with the values of the previous year *P*, and a weight *γ* is given to each of the five qualitative ordered values of attributes (*n* goes from 0 to 4) of the *RLCS (A,B,a_i)*.

$$Sim_{previous}(A,B)=\sum_{a_i \in A}\alpha_i\left[s(Z^{A}_{a_i,P},Z^{B}_{a_i,P})+\left[1-\frac{td(a_i)}{R}\right]\beta_{AllenRLCS(A,B,a_i)}\gamma(n)\right]$$

Equation 5:2 – Similarity in the CAPTA previous year approach

In the CAPTA long-term change approach, the similarity is computed with the initial *I* and final *F* states of the series.

$$Sim_{long\text{-}term}(A,B)=\sum_{a_i \in A}\alpha_i\left[\left[\sum_{t \in I,F}s(Z^{A}_{a_i,t},Z^{B}_{a_i,t})\right]+\left[1-\frac{td(a_i)}{R}\right]\beta_{AllenRLCS(A,B,a_i)}\gamma(C)\right]$$

Equation 5:3 – Similarity in the CAPTA long-term change approach

5.3 Evaluation

This section presents the evaluation of CAPTA compared with a baseline of existing methods. We performed experiments with the three values of granularity *g*, and using *fine* granularity was always better. Therefore, in the comparison of CAPTA with other techniques shown in this section we only show the results for *fine* granularity.

The evaluation will use the measures introduced in Chapter 4, namely accuracy, precision, recall and specificity. Moreover, we perform our evaluation with respect to two criteria: Mean Absolute Error (MAE) minimization and accuracy maximization.

Statistical significance has also been calculated. Statistical significance is an important parameter to consider when evaluating time series classification (Demšar 2006).

In our case, to compare the accuracy of two learning methods we use the *paired t-test* to compute their *p-value*. As usual in ML evaluation methodology, the *null hypothesis* means that the two learning methods being compared have the same accuracy (their values are not significantly different). If the *p-value* is sufficiently small (typically < 0.05) then the null hypothesis can be rejected, and in this case, there is sufficient evidence at the alfa level (tipically 0.05) of significance to show that their accuracy values are significantly different.

In the evaluations below we also compute the *p-value* for the Mean Absolute Error.

Statistical significance is computed using the paired t-test of the R software environment. R has been chosen because it is an open-source tool that supports statistical computing, including a function to compute statistical significance.

Moreover, now that we have introduced the four temporal representations for CBR, we also asked ourselves how the new temporal representations could affect the performance of ML methods (since in the previous chapter the ML methods were using the *original* temporal representation). For evaluation purposes, we selected the decision tree method J48 and created three new variants according to the types of temporal representation: J48 Previous state, J48 Long term change, and J48 State changes. Since our CBR method also uses these temporal representations we wanted to determine whether a particular temporal representation was better in a specific domain for two different methods.

5.3.1 Emotional functions of people with SCI

We will start by evaluating CAPTA in the domain of emotional functions of people with SCI; since SVM was the best method for this domain in Chapter 4 we will use it for comparison. We will first evaluate CAPTA and other techniques without using a cost matrix. Given that these results showed a class imbalance problem, we perform a second evaluation using the cost matrix (introduced in Chapter 4).

Tables 5:2 and 5:3 show the evaluation for emotional functions of people with SCI without using the cost matrix, Table 5:2 minimizing MAE (Mean Absolute Error) and Table 5:3 maximizing accuracy. It should be remarked that CAPTA does not need to use a cost matrix due to its similarity measure.

These two tables show the existence of a problem of class imbalance in the emotional functions data. For this purpose, we show the recall value for each solution class, and we can observe in Table 5:3 that J48, although having a higher value of accuracy than CAPTA, in fact only learns to predict correctly one class (Class1, that has number of instances much greater than the other classes). This is shown by J48 having a recall of 1 in Class1 and a recall of 0 in the other classes.

Table 5:2 Prediction of emotional functions of people with SCI without cost matrix minimizing MAE.

Temporal representation	**Learning**	**Accuracy**	**Recall class 0**	**Recall class 1**	**Recall class 2**	**Recall class 3**
Original	SVM	0.54	0.01	0.72	0.17	0.19
Original and state changes	CAPTA	0.62	0.11	0.91	0.16	0.04

Table 5:3 Prediction of emotional functions of people with SCI without cost matrix maximizing accuracy.

Temporal representation	**Learning**	**Accuracy**	**Recall class 0**	**Recall class 1**	**Recall class 2**	**Recall class 3**
Original	SVM	0.54	0.01	0.72	0.17	0.19
Previous state	J48	**0.66**	0	**1**	0	0
Long-term change	J48	**0.66**	0	**1**	0	0
State changes	J48	0.63	0	0.98	0.01	0.01
Original and state changes	CAPTA	0.56	0.14	0.80	0.09	0.04

The problem of class imbalance can be alleviated using a cost matrix, which we perform now in order to be fair in comparing CAPTA (that does not need to use a cost matrix) with other techniques that are affected by class imbalance.

Table 5:4 shows the prediction of emotional functions of people with SCI minimizing MAE. Here, CAPTA obtains a better (lesser) value of mean absolute error (MAE) than SVM (0.63<0.72). Moreover, accuracy is also higher, although CAPTA does not have better values of precision nor specificity. The column "MAE

Significance" of Table 5:4 indicates with the value "yes" that (since its p-value is $p < 0.05$ and therefore the null hypothesis is rejected) CAPTA's MAE is significantly different to than of SVM.

Table 5:4 Prediction of emotional functions of people with SCI minimizing MAE.

Temporal representation	Learning	Accuracy	Precision	Recall (or sensitivity)	Specificity	MAE	MAE Significance
Original	SVM	0.54	0.56	0.54	0.60	0.72	yes
Original and state changes	CAPTA	0.62	0.54	0.62	0.48	**0.63**	-

Table 5:5 shows the prediction (maximizing accuracy) of emotional functions of people with SCI. The method which obtained the best accuracy in the previous chapter (SVM), and the different representations using decision trees are compared with CAPTA. CAPTA is the one which obtains the best accuracy and precision. We calculated all the p-values for accuracy and they are < 0.05, so there is evidence that CAPTA's accuracy is significantly better.

Figure 5:3 visualizes the MAE and Accuracy values of both Tables 5:4 and 5:5; in this domain CAPTA is better with respect to both measures, MAE and Accuracy.

Table 5:5 Prediction of emotional functions of people with SCI maximizing accuracy.

Temporal representation	Learning	Accuracy	Accuracy Significance	Precision	Recall (or sensitivity)	Specificity
Original	SVM	0.54	yes	0.56	0.54	0.60
Previous state	J48	0.4	yes	0.52	0.41	0.70
Long-term change	J48	0.37	yes	0.54	0.37	0.73
State changes	J48	0.40	yes	0.53	0.41	0.72
Original and state changes	CAPTA	**0.56**	-	0.57	0.56	0.49

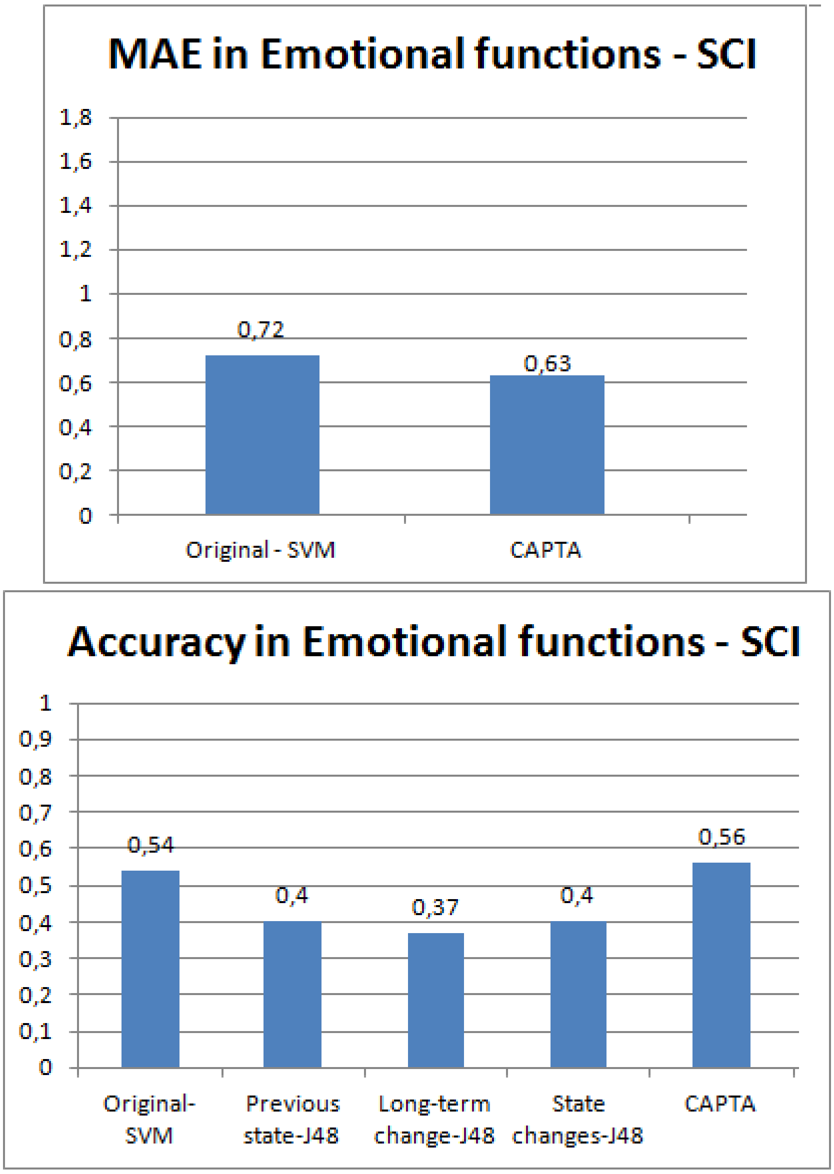


Figure 5:3 MAE and accuracy comparison in the prediction of emotional functions in SCI.

5.3.2 Urination functions of people with SCI

We will now evaluate CAPTA in the domain of urination functions of people with SCI.

Table 5:6 shows the prediction (while minimizing MAE) of urination functions of people with SCI, where CAPTA obtains a lower MAE value than that of IBk (the method which obtained the best MAE in the previous chapter). Moreover, CAPTA's accuracy is also higher than IBk. Since the p-value of the MAE with respect to CAPTA is < 0.05 there is evidence that CAPTA's MAE is significantly better.

Table 5:6 Prediction of urination functions in SCI minimizing MAE.

Temporal representation	Learning	Accuracy	Precision	Recall (or sensitivity)	Specificity	MAE	MAE Significance
Original	IBk (k=7)	0.70	0.67	0.70	0.64	1.14	yes
Original and state changes	CAPTA	0.74	0.70	0.74	0.74	**0.56**	-

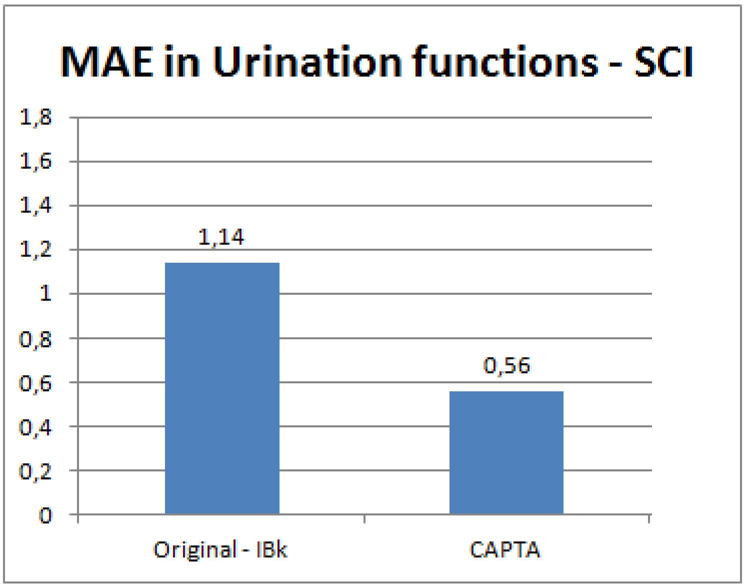
Table 5:7 shows the prediction of urination functions of people with SCI maximizing accuracy. The method that obtained the best accuracy in the previous chapter was J48. In our evaluation we are now comparing CAPTA with J48, not only with the original representation, but also J48 with our three temporal representations. In this prediction domain, the *long-term change* in J48. CAPTA does not obtain competitive values of accuracy, precision, recall nor specificity. In this prediction domain, as the long-term change temporal representation has the best accuracy results using decision trees, we have also used the *long-term change* representation for CAPTA in addition to our normal State Changes representation. CAPTA using the *long-term change* representation improves in accuracy, precision, recall and specificity with respect to using the *Original and state changes* representation; however the difference on accuracy is not statistically significant.

Moreover, the p-values of the accuracy with respect to CAPTA > 0.05, so there is no evidence against the null hypothesis with these methods. So, J48 with *Previous state* representation has the best accuracy but it is not statistically significant with respect to the accuracy of CAPTA with *Original and state changes* representation.

Figure 5:4 shows the results of Table 5:6 and Table 5:7; in this domain CAPTA is clearly better with respect to the MAE measure but is at the same level of performance to the other methods with respect to the accuracy measure.

Table 5:7 Prediction of urination functions in SCI maximizing accuracy.

Temporal representation	Learning	Accuracy	Accuracy Significance	Precision	Recall (or sensitivity)	Specificity
Original	J48	0.84	no	0.79	0.84	0.84
Previous state	J48	0.67	no	0.67	0.67	0.58
Long-term change	J48	**0.87**	no	0.87	0.87	0.89
State changes	J48	0.84	no	0.84	0.85	0.86
Original and state changes	CAPTA	0.61	-	0.33	0.34	0.69
Long-term change	CAPTA	0.70	no	0.57	0.44	0.73



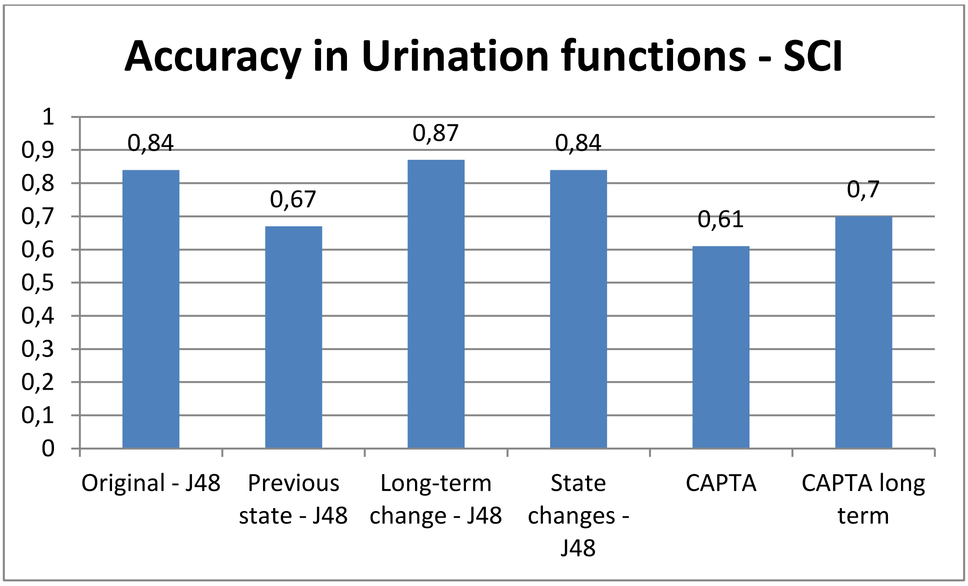


Figure 5:4 MAE and accuracy comparison in the prediction of urination functions in SCI.

5.3.3 Emotional functions of people with ABI

We will now evaluate CAPTA in the domain of emotional functions of people with ABI.

Table 5:8 shows the prediction (while minimizing MAE) of emotional functions of people with ABI. The method which obtained the best MAE in the previous chapter was J48. Here comparing CAPTA with *J48 with the original temporal representation* we see CAPTA has lower accuracy. However, the p-value of the MAE with respect to CAPTA > 0.05 so we can say than both methods have a similar performance.

Table 5:8 Prediction of emotional functions of people with ABI minimizing MAE.

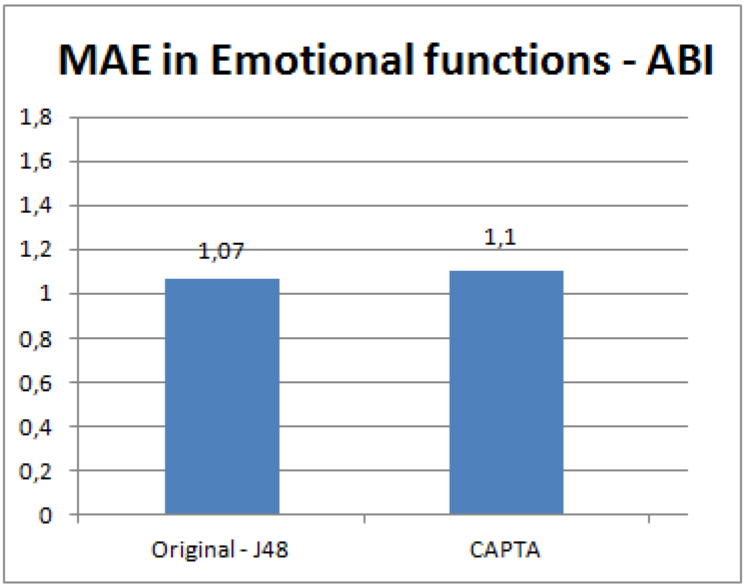
Temporal representation	Learning	Accuracy	Precision	Recall (or sensitivity)	Specificity	MAE	MAE Significance
Original	J48	0.38	0.34	0.37	0.84	**1.07**	no
Original and state changes	CAPTA	0.32	0.39	0.32	0.72	1.10	-

Table 5:9 shows the prediction (while maximizing accuracy) of emotional functions of people with ABI. We compare the accuracy of CAPTA with those of SVM (the method that obtained the best accuracy in the previous chapter) and decision trees using several temporal representations. In this prediction domain CAPTA obtains a better value of accuracy although it does not have the better values of precision nor recall nor specificity. We can see that there are some p-values of the accuracy with respect to CAPTA > 0.05 so there is no evidence against the null hypothesis.

Figure 5:5 visualizes the Table 5:8 and Table 5:9; in this domain CAPTA performance is very similar to the other methods for moth measures, MAE and Accuracy.

Table 5:9 Prediction of emotional functions of people with ABI maximizing accuracy.

Temporal representation	Learning	Accuracy	Accuracy Significance	Precision	Recall (or sensitivity)	Specificity
Original	SVM	0.41	no	0.41	0.41	0.85
Previous state	J48	0.37	no	0.32	0.35	0.77
Long-term change	J48	0.31	no	0.23	0.31	0.69
State changes	J48	0.30	no	0.27	0.30	0.70
Original and state changes	CAPTA	**0.42**	-	0.25	0.24	0.80



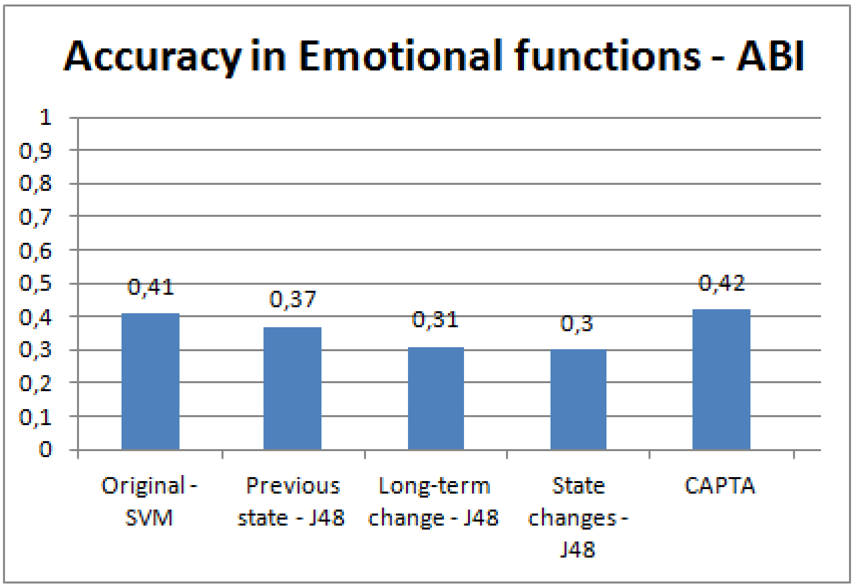


Figure 5:5 MAE and accuracy comparison in the prediction of emotional functions in ABI.

5.3.4 Executive functions of people with ABI

We will now evaluate CAPTA in the domain of executive functions of people with ABI.

Table 5:10 shows the prediction (while minimizing MAE) of executive functions of people with ABI. The method that obtained the best MAE in the previous chapter was IBk. In this prediction domain, CAPTA has a lower MAE value than IBk. Since the p-value of the MAE with respect to CAPTA is < 0.05 there is evidence that CAPTA's MAE is significantly better.

Table 5:10 Prediction (while minimizing MAE) of executive functions of people with ABI.

Temporal representation	Learning	Accuracy	Precision	Recall (or sensitivity)	Specificity	MAE	MAE Significance
Original	IBk (k=7)	0.32	0.30	0.32	0.82	1.63	yes
Original and state changes	CAPTA	0.28	0.28	0.30	0.81	**0.30**	-

Table 5:11 shows the prediction (while maximizing accuracy) of executive functions of people with ABI. We compare the accuracy of CAPTA with NB (the method which obtained the best accuracy in the previous chapter), and the decision trees using several temporal representations. Here, *J48 with Previous state representation* obtains the best values of accuracy, precision, recall and specificity. In this prediction domain, as the *Previous state* temporal representation has the best accuracy results using decision trees, we have also included an evaluation of CAPTA with the *Previous state* temporal representation. However, there is no significant difference on accuracy for CAPTA when using either temporal representation. The p-values of the accuracy with respect to CAPTA are higher 0.05, so there is no evidence that *J48* has significantly better accuracy than CAPTA.

Figure 5:6 shown the results of Table 5:10 and Table 5:11; in this domain CAPTA is better with respect to the MAE measure but is similar to the other methods with respect to accuracy.

Table 5:11 Prediction of executive functions of people with ABI maximizing accuracy.

Temporal representation	Learning	Accuracy	Accuracy Significance	Precision	Recall (or sensitivity)	Specificity
Original	NB	0.43	no	0.42	0.43	0.86
Previous state	J48	**0.48**	no	0.47	0.48	0.87
Long-term change	J48	0.34	no	0.34	0.34	0.83
State changes	J48	0.39	no	0.39	0.39	0.84
Original and state changes	CAPTA	0.39	-	0.34	0.31	0.83
Previous state	CAPTA	0.41	no	0.31	0.34	0.84

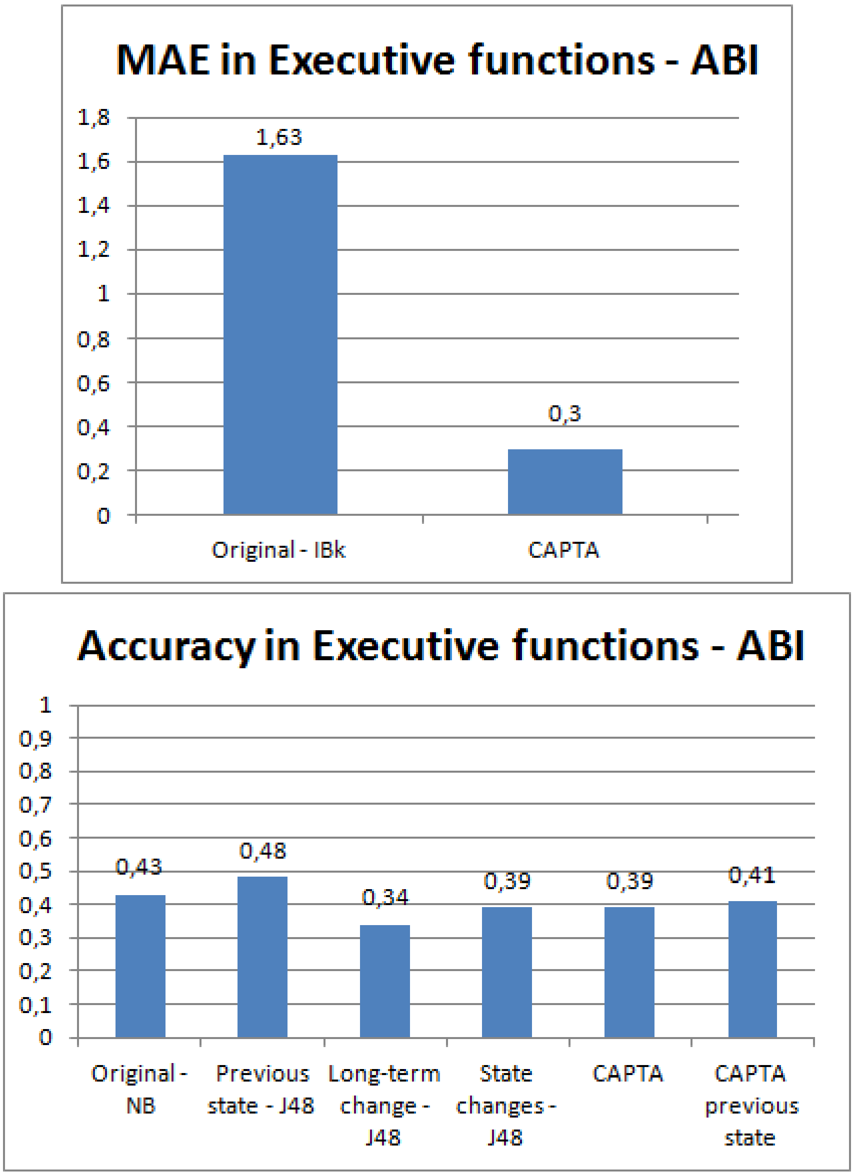


Figure 5:6 MAE and accuracy comparison in the prediction of executive functions in ABI.

5.4 Visualization of the prognosis

The visualization interface was developed using the LineChart library of Google charts[6] and the html pages where generated using the library strsubstitutor[7]. A Java library generated the static htmls for each patient.

In Figure 5:7 there is an example that shows the evolution in time of the deficiency level (Emotional functions) of a target person and the 3 cases that are more similar to that target. At the extreme right-hand side of the evolution the predicted value for the deficiency level of the target person is shown as a star (★).

[6] Available at: https://developers.google.com/chart/interactive/docs/gallery/linechart
[7] Available at : https://commons.apache.org/proper/commons-lang/javadocs/api-release/org/apache/commons/lang3/text/StrSubstitutor.html

Additionally, at the right-hand side of the three more similar cases the interface shows (under the "Comparison" title) two subsets of attributes for each case: the subset of more similar attributes to the target and the subset of attributes more dissimilar to the target.

When clicking at one of attribute in the Comparison table (be it similar or dissimilar) of a similar case, the interface plots the evolution of values of that attribute in two graphics: one for the target person and one for the similar case the user is interested in (see Figure 5:8 and Figure 5:9). This allows the user to explore the similar/dissimilar attributes between a similar case and the target person and observe the evolution in time of a specific attribute.

To compare two time-series of one attribute of two patients, we have used the dynamic time warping (DTW) distance[8], a method often used in time-series classification. If the DTW distance is below 0.7, the two attributes are considered similar, otherwise they are considered dissimilar. However, in order not to overwhelm the user, not all defined attributes are used: only a subset of attributes deemed relevant according to a psychologist and doctor is used.

In Figure 5:7 we can see the target case that is a 48-year-old man with traumatic injury, ASIA level A and neurological level L3. Currently that person has a mild deficiency (value 1) in emotional functions, and the system predicts that next year, that will be 31 years since his injury, the deficiency will continue being *mild* (value 1). The CAPTA system then provides three similar cases with their similar and dissimilar attributes. The first case is a 68-year-old man who had a medical injury, his ASIA level is A, and his neurological level T4. This case has in common with the target person six attributes: energy and drive functions, urination functions, defecation functions, immediate family, sports and education. On the other hand they are dissimilar in three attributes: gait pattern functions, remunerative employment and sensation of pain. In Figure 5:8 we can see the result of the user clicking on the attribute immediate family: the interface shows the evolution of the values on time of the immediate family in two graphics, one for each case. In this example the plots show that the values are 0 over their time periods, i.e. there is no deficiency of immediate family for both cases. Notice that the period for for the target is 29-30 years, while for the similar case under consideration the time period is 50-55 years.

The second most similar case (shown in Fig. 6.1) is a 61-year-old man with traumatic injury, his ASIA level is A, and his neurological level is C7. Now we can observe that sensation of pain is a new similar attribute (not present in the previous similar case) and education is also a new dissimilar attribute.

In Figure 5:9 we can see the result of the user clicking on the *dissimilar* attribute remunerative employment: the interface shows the evolution of the values on time of the remunerative employment in two graphics, one for each case. In this example the plots show that the target person has no difficulty (has value zero over the period of 29-30 years) while the second most similar case has values that alternates between complete difficulty (4) and severe difficulty (3).

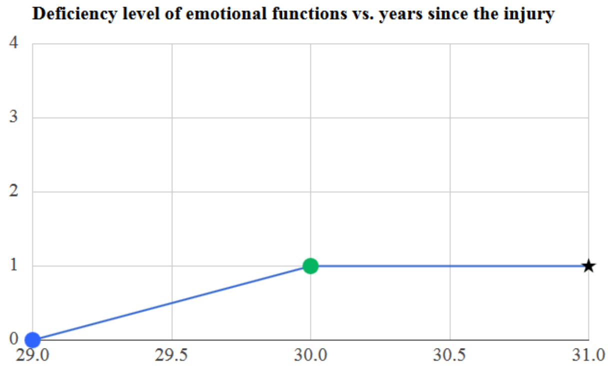
Finally, the third case is a 68-year-old man, with a traumatic injury, his ASIA level is A, and his neurological level is T8. His similar attributes with the target case are urination functions, immediate family, sports and

[8] Available at:
http://trac.research.cc.gatech.edu/GART/browser/GART/weka/edu/gatech/gart/ml/weka/DTW.java?rev=9

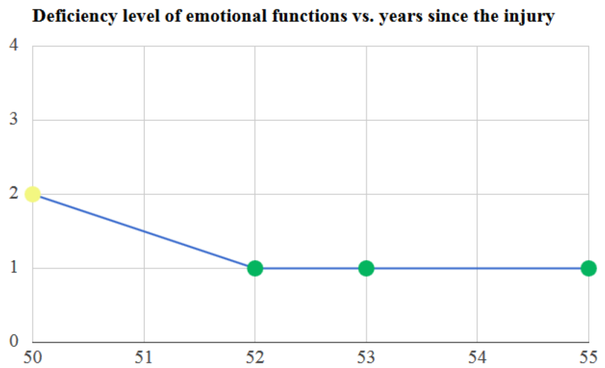
gait patter functions, while his dissimilar attributes are energy and drive functions, defecation functions, remunerative employment, sensation of pain and education.

In the future, expert clinicians will validate the usability of the visualization interface proposed here. In addition the interface will be able to be customized (1) by selecting which attribute is to be predicted, and (2) selecting the number of similar cases that are shown.

Target case
Person with complete paraplegia, 48 years old, male, traumatic injury, ASIA level A, neurological level L3.



Similar cases
Person with complete paraplegia, 68 years old, male, medic injury, ASIA level A, neurological level T4.

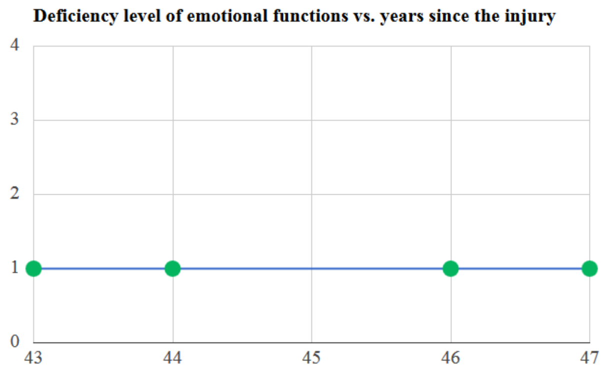


Comparison

Similar attributes
- Energy and drive functions
- Urination functions
- Defecation functions
- Immediate family
- Sports
- Education

Dissimilar attributes
- Gait pattern functions
- Remunerative employment
- Sensation of pain

Person with complete tetraplegia, 61 years old, male, traumatic injury, ASIA level A, neurological level C7.

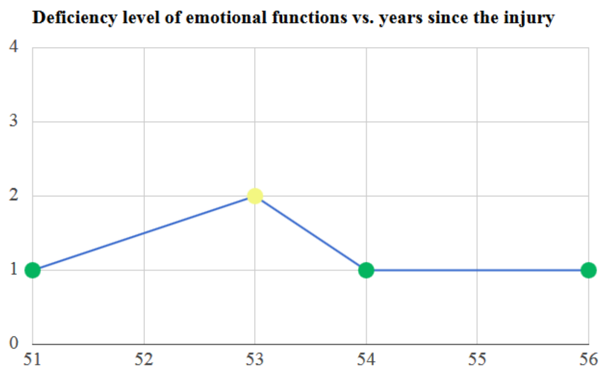


Comparison

Similar attributes
- Energy and drive functions
- Urination functions
- Defecation functions
- Immediate family
- Sports
- Sensation of pain

Dissimilar attributes
- Gait pattern functions
- Remunerative employment
- Education

Person with complete paraplegia, 68 years old, male, traumatic injury, ASIA level A, neurological level T8.



Comparison

Similar attributes
- Urination functions
- Immediate family
- Sports
- Gait pattern functions

Dissimilar attributes
- Energy and drive functions
- Defecation functions
- Remunerative employment
- Sensation of pain
- Education

Figure 5:7 View of the prediction of a target case in the domain of “emotional functions of people with SCI”.

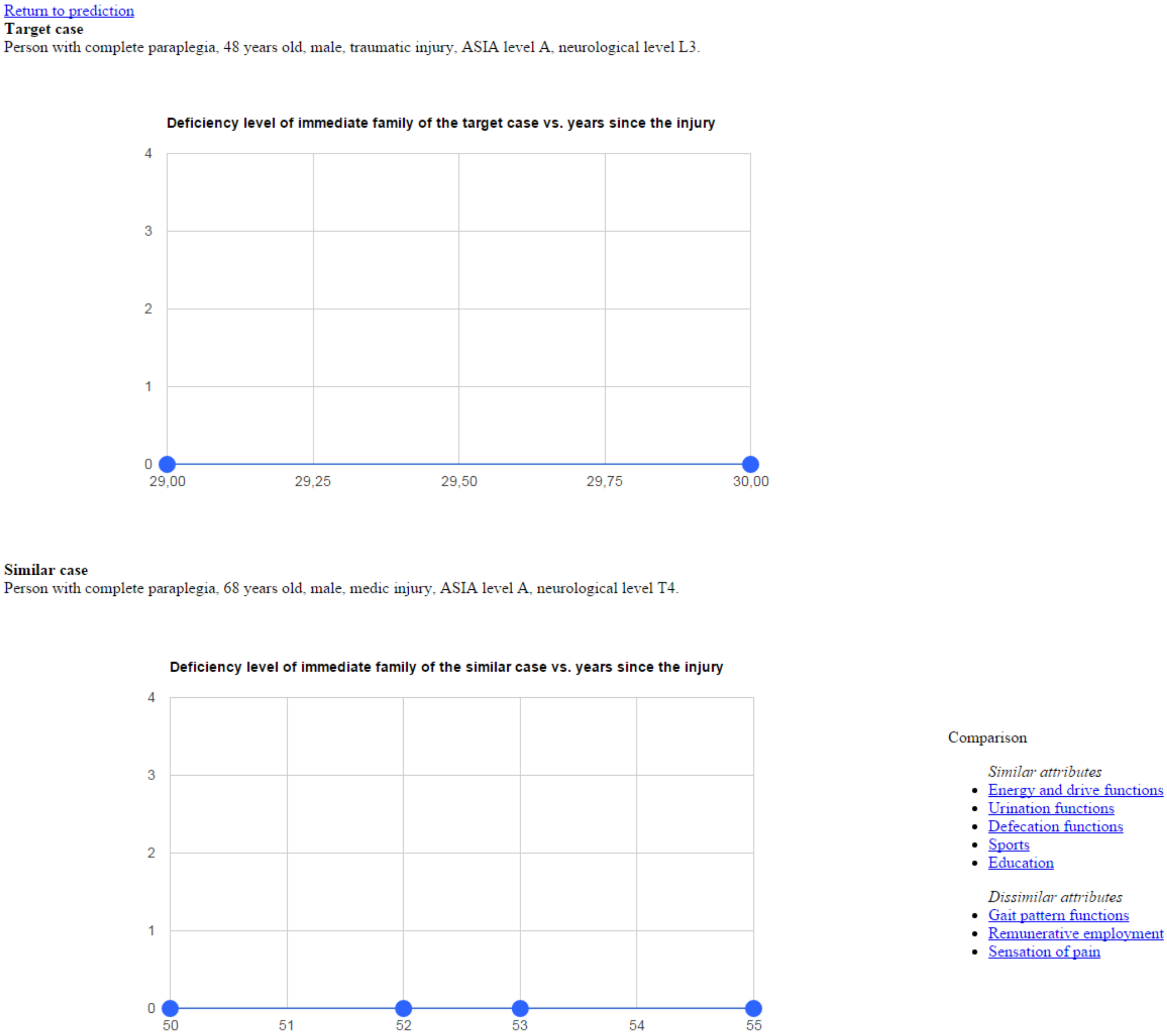


Figure 5:8 View, when selecting the similar attribute *Immediate family,* comparing the target case and the most similar case.

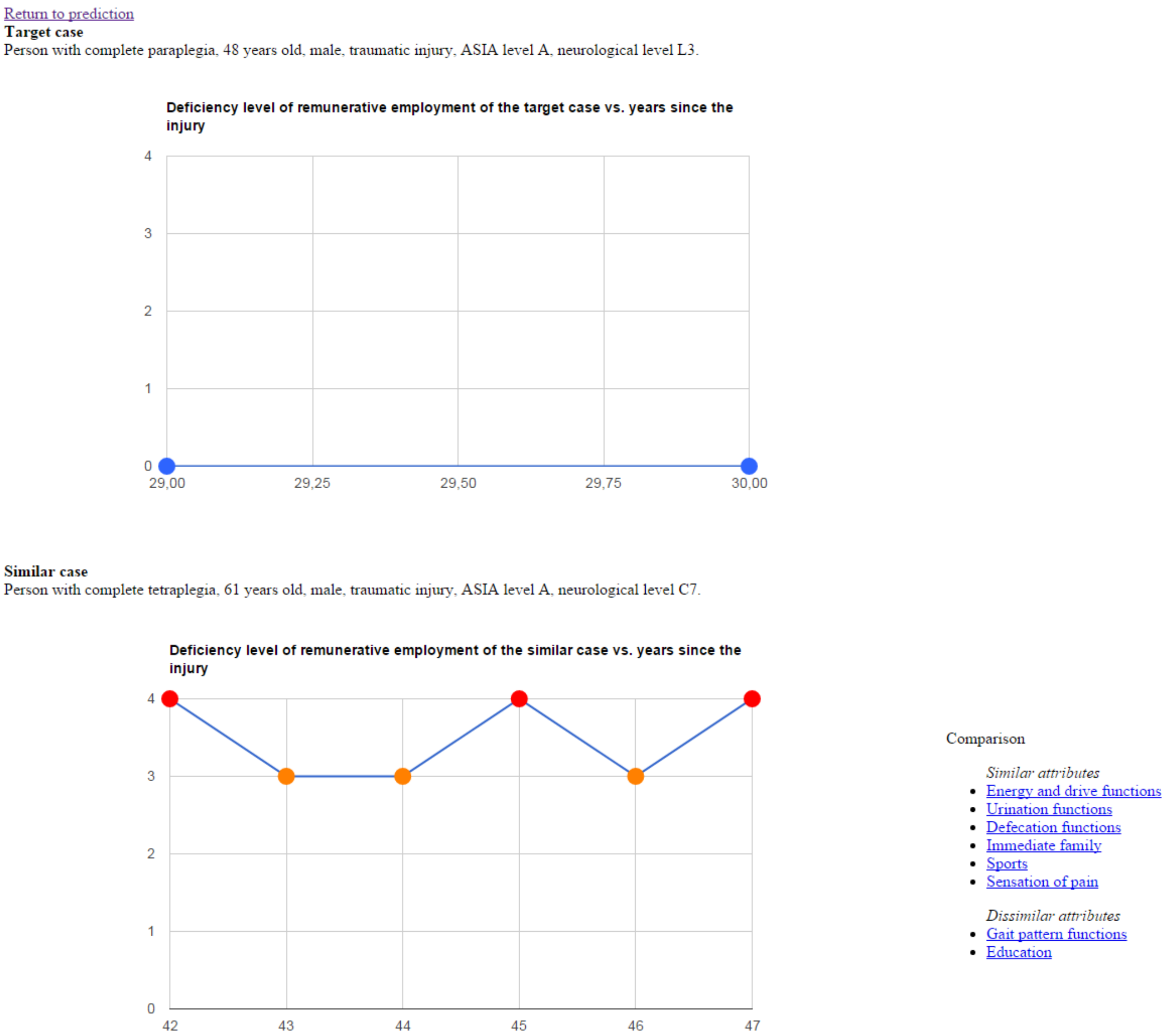


Figure 5:9 View, when selecting the dissimilar attribute *Remunerative employment,* comparing the target case and the most similar case.

5.5 Summary

In this chapter, time series of standardized indicators in the domain of diseases of neurological origin have been successfully used in a case-based reasoning system to predict the next year deficiency levels in four health domains. The use of standard knowledge representations has the additional advantage of providing interoperability among clinical organizations.

Regarding temporal representation, we introduce a temporal abstraction with three levels of granularity and use this in the similarity among cases of a case-based reasoning system. The temporal abstraction comprises four case-representations: *Original*, which comprises of a pair (attribute, time) that constitutes a new attribute, *Previous State*, which consists of the state prior to the one to be predicted; *Long-term change*, which consists of the first state of the series, and a representation of the change from this original state to the state prior to the one to be predicted; *State changes*, which consists of a representation of the change from each state of the time series to the following one.

The resulting system CAPTA is used to predict a prognosis in four clinical domains, and is evaluated using real cases from a hospital database (Institut Guttmann).

The similarity measure of the case-based reasoning system, CAPTA, is composed of two parts, each estimating different relevant aspects of the temporal comparison between two sequences: the time-shifting part and the Allen relations over the rightmost longest common subsequence part. Weights in the similarity measure are computed using a genetic algorithm in two experiments, each one with a different fitness function: one of them minimizing MAE, and another one maximizing accuracy.

This experimental evaluation shows that CAPTA has competitive results in our four prediction domains: emotional functions and urination functions for the spinal cord injury dataset, and emotional functions and executive functions for the acquired brain injury dataset. Since we have evaluated the methods with respect to MAE and accuracy, we can conclude the following:

1. Regarding MAE, CAPTA is better (3 domains) or equal (1 domain); and
2. Regarding accuracy, CAPTA is competitive in three domains (since accuracy differences are not statistically significant) and has better accuracy in one domain.

All in all, CAPTA seems a robust approach for prediction in the four medical domains we have modeled, both considering MAE and accuracy measures. Moreover, we are comparing CAPTA, for each domain, with the best method (SVM, IBk, J84) for MAE and the best method for accuracy. Notice that none of those methods are the best in all domains and evaluation measures. CAPTA is equal or better in all eight domain/measure pairs.

Furthermore, we have not only compared CAPTA with the best methods of Chapter 4, we have augmented decision trees (implemented by J48) with our three temporal representations in order to take into account in our evaluation process whether they could improve the performance of decision trees. This augmentation increases, in principle, the difficulty of CAPTA being competitive in the evaluations we have carried out. However, CAPTA is clearly competitive with respect to these augmented decision trees.

In this chapter, a visualization interface of the prognosis has also been provided. This visualization shows the prediction, the three most similar patients, and their main similar and dissimilar attributes. Therefore, the user can navigate seeing the temporal context of the target case and its comparison with other previous cases.

The next chapter will presents the conclusions of the thesis and discuss the future work.

Chapter 6 Conclusions and Future Work

6.1 Conclusions

The main goals of this work have been:

- to develop an interoperable *prognosis system* with a new case-based reasoning method called CAse-based Prognosis using Temporal Abstractions (CAPTA), evaluated experimentally for prediction about four types of patients with neurological diseases; and

- to develop an interoperable *monitoring system* which contributes to a better understanding of the temporal nature of clinical attributes from a multi-stakeholder perspective.

The knowledge (concepts and relationships) of the prognosis and monitoring systems has been represented using an ontology whose terminology follows the international standards promoted by the WHO. A new automatic translation system from terminology used in patient questionnaires to international WHO standards, based on the interoperable ontology, allowed (see Chapter 3) to address the issue of interoperability in the disability domain. Moreover, a new method has been presented to automatically collect, transform, share and graphically visualize standardized and multidimensional indicators.

In Chapter 4, a monitoring system for the medical domain was developed based on the new ontology and method. The monitoring system does not only include the patient diagnosis but also the main aspects of a person's life (development, participation, and environment) using the International Classification of Functioning, Disability and Health (ICF). The evaluation results indicate that the monitoring system has great potential for two types of users. Professionals found the monitoring system useful to generate new knowledge. People with disabilities found it helpful to enrich personal knowledge with the experiences of other users. As a consequence, the monitoring system has the potential to promote user empowerment and the ability to make decisions with a more informed opinion.

Secondly, in this chapter we have applied several classical methods of machine learning (ML) to prognosis in four clinical domains. We have used the canonical implementation of Weka to experimentally evaluate the performance of these ML methods on these domains without any added knowledge about time, as we pursue in the next chapter. The evaluation results determine the best ML method for each of the four clinical domains, and these specific methods later are used as a basis of comparison with our temporal case-based method in Chapter 5.

Chapter 5 presents our contribution to temporal case-based prognosis. First, we proposed a new way to represent time-varying clinical attributes or indicators; here a time-series of values are represented with sequences of qualitative value changes, and these changes can be described at several levels of abstraction. Second, since CBR focuses on estimating similarity measures between cases, we developed a similarity measure that takes into account time-varying attributes represented as “change sequences”. Since direct comparison of these sequences is not adequate, our similarity takes into account the different Allen relations on time intervals and the possible time-shifting effects of values in an attribute. The similarity measure is a weighted combination of two components, each estimating different relevant aspects of the temporal comparison between two change sequences: the time-shifting aspect and the Allen relations over the rightmost longest common subsequence. The weights in the similarity measure have been estimated using a genetic algorithm. Moreover, we have evaluated prognosis with respect to two different evaluation measures, namely accuracy and MAE; for this reason the genetic algorithm used two fitness functions, one maximizing accuracy and one minimizing MAE, therefore obtaining two sets of weights for the similarity measure when used by the case-based reasoning system for prognosis.

Regarding the prognosis, we developed CAPTA, a case-based system that uses our temporal similarity on time-varying attributes to retrieve the cases most similar of a target patient. CAPTA has been applied to four clinical domains: emotional functions and urination functions for the spinal cord injury dataset, and emotional functions and executive functions for the acquired brain injury data set. The four case bases were created from real patient data from a hospital database (Institut Guttmann). Chapter 5 describes an experimental evaluation of CAPTA for each of these four domains in comparison to the best ML method for this domain selected in Chapter 4. This experimental evaluation shows that CAPTA has competitive results in our four prediction domains. Since we have evaluated the methods with respect to two measures, Mean Absolute Error (MAE) and accuracy, we found that, regarding MAE, CAPTA is clearly better in 3 domains and equal 1 domain; and regarding accuracy, CAPTA is competitive in three domains (since accuracy differences are not statistically significant) and has clearly a better accuracy in one domain.

All in all, CAPTA seems a robust approach for prediction in the four medical domains we have modeled, both considering MAE and accuracy measures. Moreover, we are comparing CAPTA, for each domain, with the best method (SVM, IBk, J84) for MAE and the best method for accuracy. Notice that none of those methods are the best in all domains and evaluation measures. CAPTA is equal or better in all eight domain/measure pairs.

Furthermore, we have not only compared CAPTA with the best methods of Chapter 4, we have augmented decision trees (implemented by J48) with our three temporal representations in order to take into account in our evaluation process whether they could improve the performance of decision trees. This augmentation increases, in principle, the difficulty of CAPTA being competitive in the evaluations we have carried out. However, CAPTA is clearly competitive with respect to these augmented decision trees.

Finally in Chapter 5, a visualization interface for the suggested prognosis has also been provided. This visualization shows the prediction of the target patient, the three most similar patients, and the main similar and dissimilar attributes for each retrieved case with respect to the target patient. Therefore, the user can explore the temporal context of the target case in comparison to the most similar cases visualizing the prediction and the main similar and dissimilar attributes.

6.2 Future work

There are three lines of future work we would like to further explore: interoperability, medical domain analysis, and prognosis systems.

Regarding interoperability, incorporating information models, such as EN 13606 or HL7 virtual medical record (vMR) may facilitate information exchange among Medical Health Records (MHRs). Moreover, considering socio-economic impact assessment, we are planning to extend the use of the proposed approach of Appendix I for: (1) automatic socioeconomic impact assessment of other rehabilitation processes, such as functional or cognitive rehabilitation and (2) Quality-adjusted life years (QALYs) evolution in periodic, comprehensive and multidimensional evaluations of people with disabilities of neurological origin. Moreover, the generalization of this methodology to other diseases and rehabilitation processes can be a valuable tool to provide information on and evaluate health policies.

With regards to medical domain analysis, the Institut Guttmann has plans to promote the usage of the monitoring system among patients and clinical staff (especially psychologists and social workers). It is also planned to involve other institutions such as other reference healthcare centers, patient organizations, public health systems, and private enterprise. As 18% of users highlighted, our monitoring system can be extended to other countries. This could be useful for governments, as it would allow get a general picture of disability and of the comparison among different countries. In addition, it would allow the creation of best practices and recommendations at the European level, with direct impact on regional policies of inclusion, support and assistance to people with disabilities. Furthermore, this work can also be generalized to other types of disabilities or chronic conditions. People with visual or hearing disabilities may be users of the monitoring system if questionnaires of these groups are introduced and translated to International Classification of Functioning, Disability and Health (ICF).

With respect to the prognosis system CAPTA, the indexing method of the time-series will be improved in order to provide minimal space consumption and computational complexity. Furthermore, regarding the visualization of the proposed interface, in the future, expert clinicians will validate the usability of the visualization interface proposed here. In addition the interface will be able to be customized (1) by selecting which attribute is to be predicted, and (2) selecting the number of similar cases shown.

6.3 Publications related to this thesis

The work related to this thesis has produced the following publications and patents ordered by date.

Journal articles:

- Subirats et al. (2015) Subirats, L., Lopez-Blazquez, R., Ceccaroni, L., Gifre, M., Miralles, F., García-Rudolph, A., Tormos, J.M. Monitoring and prognosis system based on the ICF for people with traumatic brain injury. International Journal of Environmental Research and Public Health, 12(8), 9832-9847, August 2015. Impact factor: 1.998. Quartile 2 in the category of Environmental Sciences.

- Subirats et al. (2013a) Subirats, L., Ceccaroni, L., Lopez-Blazquez, R., Miralles, F., García-Rudolph, A., Tormos, J. M. Circles of Health: towards an advanced social network about disabilities of neurological origin. Journal of Biomedical Informatics, 46(6), 1006-1029, December 2013. Impact factor: 2.131. Quartile 1 in the category of Computer science, interdisciplinary applications.

- Calvo et al. (2013). Calvo, M., Subirats, L., Ceccaroni, L., Maroto, J. M., de Pablo, C. and Miralles, F. Automatic assessment of socioeconomic impact in cardiac rehabilitation. International Journal of Environmental Research and Public Health, 10(11), 5266-5283, October 2013. Impact factor: 1.998. Quartile 2 in the category of Environmental Sciences.

- Ceccaroni and Subirats (2012). Ceccaroni, L. and Subirats, L. Interoperable Knowledge representation in Clinical Decision Support Systems for Rehabilitation. International Journal of Applied and Computational Mathematics, vol. 11, no. 2, pp. 303-316, June 2012. Impact factor: 0.750. Quartile 2 in the category of Mathematics, applied.

- Subirats et al. (2012). Subirats, L. and Ceccaroni, L., Miralles, F. Knowledge representation for prognosis of health status in rehabilitation. Future Internet journal, MDPI.

Conference papers:

- Subirats et al. (2014). Subirats, L., Ceccaroni, L., Maroto J.M., de Pablo, C., Miralles, F. Medical-treatment recommendation and the integration of process models into knowledge-based systems. Proceedings of the 6th International Conference on Agents and Artificial Intelligence, 6-8th March, 2014, Angers (France). Rank: C. Field of Research: Artificial Intelligence and Image processing.

- Subirats et al. (2013b). Subirats, L., Ceccaroni, L., Gómez-Pérez, C., Caballero, R., Lopez-Blazquez, R. and Miralles, F. On semantic, rule-based reasoning in the management of functional rehabilitation processes. International Symposium on Management Intelligent Systems (IS-MiS 2013). Advances in Intelligent Systems and Computing, 220, 51-58, Salamanca (Spain), 22nd-24th May, 2013.

- Vargiu et al. (2013). Vargiu E., Ceccaroni L., Subirats L., Martin S., and Miralles F. User Profiling of People with Disabilities - A Proposal to Pervasively Assess Quality of Life, Proceedings of the 5th International Conference on Agents and Artificial Intelligence (ICAART 2013), 352-357. February 15-18, 2013, Barcelona (Spain). Rank: C. Field of Research: Artificial Intelligence and Image processing.

- Subirats and Ceccaroni (2011). Subirats, L. and Ceccaroni, L. An ontology for computer-based decision support in rehabilitation. The Tenth Mexican International Conference on Artificial Intelligence (MICAI 2011), 549-559, Lecture Notes in Computer Science, Springer. ISBN 978-3-642-25323-2.

- Subirats et al. (2010). Subirats, L., Torrellas, S., Orte, S., Miralles, F., López, R., García-Rudolph A. and Gil. A. Social Network Enhancing Self-care and Mitigating Isolation in Long-duration Disability with Built-in Knowledge Extraction Tools. In: Medicine 2.0: Social Media and Web 2.0 in Health, Medicine and Biomedical Research (Med 2.0), 2010, Maastricht, The Netherlands. ISSN 1923-2195.

6.4 Patents related to this thesis

Miralles et al. (2013). Miralles, F., Ceccaroni, L., Subirats, L., Tormos Muñoz, J.M, Gil Origuen, A., Lopez Blazquez, R. and Garcia Rudolph, A. System and method for extracting and monitoring multidimensional attributes regarding personal health status and evolution. PCT Patent EP2011/074267, filed December 29, 2011, and issued June 4, 2013.

Bibliography

Aamodt, A. and Plaza, E. (1994). Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. AI Communications, 7(i): pp 39-59.

Albrecht, G.L., Devlieger, P.J. (1999). The disability paradox: high quality of life against all odds. Social Science & Medicine. 48(8):977–988.

Allen, J.F. (1983). Maintaining knowledge about temporal intervals. In: Communications of the ACM. pp. 832–843.

Armengol, E. Palaudàries, A., Plaza, E. (2001). Individual prognosis of diabetes long-term risks: a CBR approach. Methods Inf. Med. 40(1):46-51.

Barendregt, J.J.; Bonneux, L.; van der Maas, P.J. (1996). DALYs: The age weights on balance. Bull. World Health Organ. 74, 439–443.

Beale, T. and Heard, S. (2007). An Ontology-based Model of Clinical Information, Stud. Health Technol. Inform., 129: 760-764.

Bickenbach J.E., Cieza A., Rauch A., Stucki G. (2012). ICF core sets. Manual for clinical practice. Hogrefe Publishing. http://www.icf-casestudies.org.

Bottighi, A., Leonardi, G., Montani, S. Portinale, L., Terenziani, P. (2010). Intelligent Data Interpretation and Case Base Exploration through Temporal Abstractions. Case-Based Reasoning. Research and Development. LNCS, vol. 6176, pp. 36-50.

Brown, A. W., Malec, J.F., McClelland, R.L., Diehl, N.N., Englander, J., Cifu, D.X. (2005). Clinical elements that predict outcome after traumatic brain injury: a prospective multicenter recursive partitioning (decision-tree) analysis. Journal of Neurotrauma, 22(10):1040–1051.

Calvo, M., Subirats, L., Ceccaroni, L., Maroto, J. M., de Pablo, C. (2013). Automatic assessment of socioeconomic impact in cardiac rehabilitation. International Journal of Environmental Research and Public Health, 10(11), 5266-5283, October 2013.

Ceccaroni and Subirats (2012). Ceccaroni, L. and Subirats, L. Interoperable Knowledge representation in Clinical Decision Support Systems for Rehabilitation. International Journal of Applied and Computational Mathematics, vol. 11, no. 2, pp. 303-316.

Cieza, A., Brockow, T., Ewert, T., Amman, E., Kollerits, B., Chatterji, S. B., Stucki, G. (2002). Linking health-status measurements to the international classification of functioning, disability and health. Rehabil. Med., 34(5): 205-210.

Cieza, A., Geyh, S, Chatterji, S., Kostanjsek, N., stn, B., Stucki, G. (2005). ICF linking rules: an update based on lessons learned, Rehabil. Med., 37: 212-218.

Cimino, J.J. (1998). Desiderata for Controlled Medical Vocabularies in the Twenty-First Century, Methods Inf. Med., 37(4-5): 394-403.

Chausa, P., Cáceres, C., Sacchi, L., León, A., García, F., Bellazzi, R., and Gómez, E. J. (2009). Temporal data mining of HIV registries: Results from a 25 years followup. In Proceedings of the 12th Conference on Artificial Intelligence in Medicine, AIME '09, pages 56–60. Springer-Verlag.

College of American Pathologists and National Health Service, Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT); 2011. http://purl.bioontology.org/ontology/SNOMEDCT.

Dalal, H.M.; Zawada, A.; Jolly, K.; Moxham, T.; Taylor, R.S. (2010). Home based versus centre based cardiac rehabilitation: Cochrane systematic review and meta-analysis. Br. Med. J. 2010, 340, b5631.

Demšar, J. (2006). Statistical Comparisons of Classifiers over Multiple Data Sets. Journal of Machine Learning Research 7, 1–30.

Donalson, J.J., Ashenfelter, A., Martin, F., Verwoerd, J., Ortega, J.A., Parker, C., Araujo, M. (2013). Method and apparatus for visualizing and interacting with decision trees. US Patent 20130117280 A1, filed November 2, 2012, and issued May 9, 2013.

Ericsson, K., Simon H. (1980). Verbal reports as data. Psychol Rev; 87(3):215–51.

Esling, P. and Agon, C. (2012). Time-Series Data Mining. ACM Computing Surveys, Vol. 45, 1:12.

Ferrucci, D., Levas, A., Bagchi, S., Gondek, D., Mueller, E.T. (2013). Watson: Beyond Jeopardy! Artificial Intelligence 199–200, 93–105.

Frost, J.H, Massagli, M.P. (2008). Social Uses of Personal Health Information Within PatientsLikeMe, an Online Patient Community: What Can Happen When Patients Have Access to One Another's Data. J Med Internet Res. 10(3):e15.

Funk, P. and Xiong, N. (2006). Case-based reasoning and knowledge discovery in medical applications with time series. Computational Intelligence, 22(3-4):238–253.

Geyh, S.; Cieza, A.; Kollerits, B.; Grimby, G.; Stucki, G. (2007). Content comparison of health-related quality of life measures used in stroke based on the international classification of functioning, disability and health (ICF): A systematic review. Qual. Life Res. 16, 833–851.

Gibert, K., García-Rudolph, A., Curcoll, L., Soler, D., Pla, L., and Tormos, J. M. (2009). Knowledge discovery about quality of life changes of spinal cord injury patients: Clustering based on rules by states. In Medical Informatics Europe, pages 578-583.

Hirschberg, D., S. (1977). Algorithms for the longest common subsequence problem J ACM 24, 4, 664-675.

Ibrahim, Z. Al A., Ferrane, I., Joly, P. (2011). A Similarity-Based Approach for Audiovisual Document Classification Using Temporal Relation Analysis," EURASIP Journal on Image and Video Processing.

Izawa, K.; Hirano, Y.; Yamada, S.; Oka, K.; Omiya, K.; Iijima, S. (2004). Improvement in physiological outcomes and health-related quality of life following cardiac rehabilitation in patients with acute myocardial infarction. Circ. J. 68, 315–320.

Jaere, M. D., Aamodt, A., Skalle, P. (2002). Representing temporal knowledge for case-based prediction. ECCBR '02 Proceedings of the 6th European Conference on Advances in Case-Based Reasoning Pages 174-188.

Kashfi, H. (2011). The Intersection of Clinical Decision Support and Electronic Health Record: A Literature Review, Proc. of the Federated Conf. on Comp. Science and Inf. Syst., Szczecin, Poland, 18-21, pp. 347353.

Leal, J.; Luengo-Fernández, R.; Gray, A.; Petersen, S.; Rayner, M. (2006). Economic burden of cardiovascular diseases in the enlarged European Union. Eur. Heart, 27, 1610–1619.

Marino, R.J., Barros, T., Biering-Sorensen, F., Burns, S.P., Donovan, W.H., Graves, D.E., Haak, M., Hudson, L.M., Priebe, M.M., ASIA Neurological Standards Committee 2002 (2003). International standards for neurological classification of spinal cord injury. The journal of spinal cord medicine. 26 Suppl 1: S50–6.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychol Rev; 63(2):8197.

Miralles, F., Ceccaroni, L., Subirats, L., Tormos Muñoz, J.M, Gil Origuen, A., Lopez Blazquez, R. and Garcia Rudolph, A. (2013). System and method for extracting and monitoring multidimensional attributes regarding personal health status and evolution. PCT Patent EP2011/074267, filed December 29, 2011, and issued June 4, 2013.

Mont, D. (2007). Measuring health and disability. Lancet, 369, 1658–1663.

Montani, G. Leonardi, A. Bottrighi, L. Portinale, P. Terenziani. (2013). Supporting flexible, efficient and user-interpretable retrieval of similar time series. IEEE Transactions on Knowledge and Data Engineering 25(3): 677-689.

Laxe S., Tschiesner U., Zasler N., López-Blazquez R., Tormos J.M., Bernabeu M. (2012). What domains of the international classification of functioning, disability and health are covered by the most commonly used measurement instruments in traumatic brain injury research? Clin Neurol Neurosurg; 114(6):645–50.

Pang, B.C, Kuralmani, V., Joshi, R., Hongly, Y., Lee, K.K., Ang, B.T., Li, J.; Leong, T.Y., Ng, I. (2007). Hybrid outcome prediction model for severe traumatic brain injury. J. Neurotrauma, 24(1):136–146.

Raj, R., O'Connor M.J. and AK, D. (2009). An ontology-driven method for hierarchical mining of temporal patterns: application to HIV drug resistance research. In Proceedings of the 12th Conference on Artificial Intelligence in Medicine: Artificial Intelligence in Medicine, AIME '09, pages 56–60. Springer-Verlag.

Rector A., Nowlan W., Kay S. Foundations of an electronic medical record. Methods Inf Med 1991; 30:179–86.

Rovlias, A. and Kotsou, S. (2004). Classification and regression tree for prediction of outcome after severe head injury using simple clinical and laboratory variables". Journal of Neurotrauma, 2004, 21(7):886–893.

Sànchez-Marrè, M., Cortes, U., Martinez, M., Comas, J. and Rodriguez-Roda, I. (2005). An approach for temporal case-based reasoning: Episode-based reasoning. In Proceedings of the International Conference on Case-Based Reasoning (ICCBR) 2005, Chicago, August 2005. Edited by H. Munoz-Avila and F. Ricci. Springer, Berlin, pp. 465–476.

Sassi, F. (2006). Calculating QALYs, comparing QALY and DALY calculations. Health Policy Plan. 21, 402–408.

Schmidt, R. and Gierl, L. (2001). Temporal abstractions and casebased reasoning for medical course data: Two prognostic applications. In Machine Learning and Data Mining in Pattern Recognition, volume 2123 of Lecture Notes in Computer Science, pages 23–34. Springer Berlin / Heidelberg.

Schmidt, R., Gierl, L. (2003). Case-Based Reasoning for Time Courses Prognosis. Knowledge-Based Intelligent Information and Engineering Systems Lecture Notes in Computer Science Volume 2773, pp 984-991.

Serrà, J., Arcos, J. L., Garcia-Rudolph, A., García-Molina, A., Roig, T. and Josep Maria Tormos. (2013). Cognitive prognosis of acquired brain injury patients using machine learning techniques. Int. Conf. on Advanced Cognitive Technologies and Applications (COGNITIVE), Valencia, Spain, 108-113.

Shahar, Y. (1997). A framework for knowledge-based temporal abstraction. Artificial Intelligence, 90 (1–2):79 – 133.

Sowa, J.F. (2000). Knowledge Representation: Logical, Philosophical, and Computational Foundations, Brooks Cole Publishing, Pacific Grove.

Subirats, L. and Ceccaroni, L. (2011). An ontology for computer-based decision support in rehabilitation. The Tenth Mexican International Conference on Artificial Intelligence (MICAI 2011), 549-559, Lecture Notes in Computer Science, Springer. ISBN 978-3-642-25323-2.

Subirats, L., Ceccaroni, L., Miralles, F. (2012) Knowledge representation for prognosis of health status in rehabilitation. Future Internet journal, MDPI.

Subirats, L., Ceccaroni, L., Lopez-Blazquez, R., Miralles, F., García-Rudolph, A., Tormos, J. M. (2013) Circles of Health: towards an advanced social network about disabilities of neurological origin. Journal of Biomedical Informatics, 46(6), 1006-1029, December 2013.

Subirats, L., Ceccaroni, L., Gómez-Pérez, C., Caballero, R., Lopez-Blazquez, R. and Miralles, F. (2013). On semantic, rule-based reasoning in the management of functional rehabilitation processes. International Symposium on Management Intelligent Systems (IS-MiS 2013). Advances in Intelligent Systems and Computing, 220, 51-58, Salamanca (Spain), 22nd-24th May, 2013.

Subirats, L., Ceccaroni, L., Maroto J.M., de Pablo, C., Miralles, F. (2014) Medical-treatment recommendation and the integration of process models into knowledge-based systems. Proceedings of the 6th International Conference on Agents and Artificial Intelligence, 6-8th March, 2014, Angers (France).

Subirats, L., Lopez-Blazquez, R., Ceccaroni, L., Gifre, M., Miralles, F., García-Rudolph, A., Tormos, J.M. (2015) Monitoring and prognosis system based on the ICF for people with traumatic brain injury. International Journal of Environmental Research and Public Health, 12(8), 9832-9847, August 2015.

Subirats, L., Torrellas, S., Orte, S., Miralles, F., López, R., García-Rudolph A. and Gil. A. (2010). Social Network Enhancing Self-care and Mitigating Isolation in Long-duration Disability with Built-in Knowledge Extraction Tools. In: Medicine 2.0: Social Media and Web 2.0 in Health, Medicine and Biomedical Research (Med 2.0), 2010, Maastricht, The Netherlands. ISSN 1923-2195.

Vargiu E., Ceccaroni L., Subirats L., Martin S., and Miralles F. (2013). User Profiling of People with Disabilities - A Proposal to Pervasively Assess Quality of Life, Proceedings of the 5th International Conference on Agents and Artificial Intelligence (ICAART 2013), 352-357. February 15-18, 2013, Barcelona (Spain).

Westra, B.L., Savik, K., Oancea, C., Choromanski, L., Holmes, J.H., Bliss, D. (2011). Predicting Improvement in Urinary and Bowel Incontinence for Home Health Patients Using Electronic Health Record Data. J Wound Ostomy Continence Nurs., 38(1): 77–87.

Whiteneck G., Meade MA, Dijkers M., Tate D, bushnik T, Forchheimer MB. (2004). Environmental factors and their role in participation and life satisfaction after spinal cord injury. Arch Phys Med Rehabil 2004; 85: 1793-803.

Williams, A. (1995). The Measurement and Valuation of Health: A Chronicle; Centre for Health Economics, University of York: Heslington, UK.

World Health Organization (2004). The Global Burden of Disease: 2004 Update. Available online: http://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf (accessed on 16 August 2014).

World Health Organization, International Classification of Functioning, Disability and Health (ICF). http://apps.who.int/classifications/icf.

Subject index

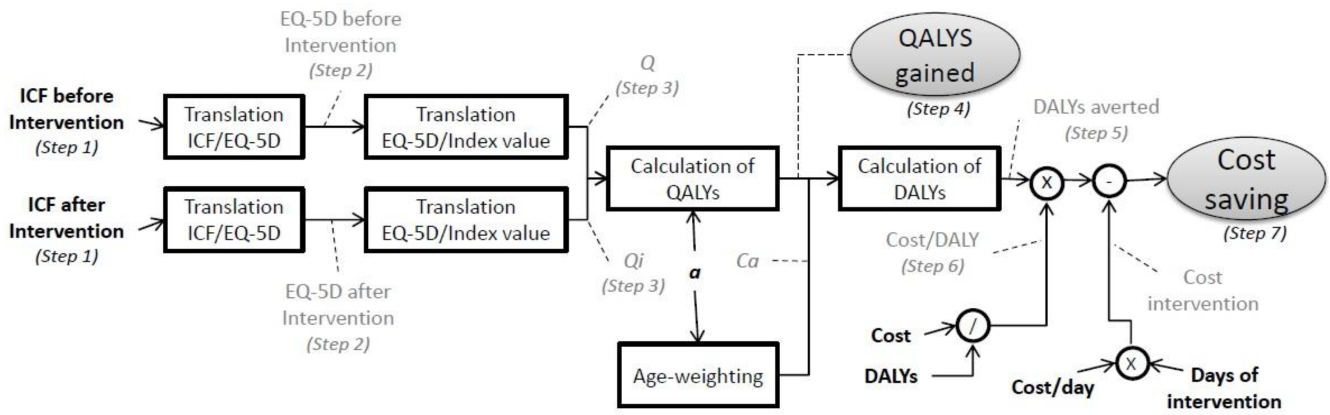
acquired brain injury, 19, 29, 30, 39, 98
Allen, 38, 39, 76, 79
CBR, 37, 38, 71
CDSS, 19, 27, 28, 55, 71
classes, 21, 32, 42, 46, 47, 48, 49
College of American Pathologists and National Health Service, 31
content management system, 57
decision tree, 39
disability paradox, 101
diseases of neurological origin, 39, 94
emotional functions, 61, 62, 68, 69, 70, 81, 82, 83, 85, 86, 88, 92, 93, 94, 95, 98, 108
EN/ISO 13606, 32, 40, 54, 55
evolution, 32, 33, 39, 43, 44, 55, 57, 61, 62, 63, 64, 66, 67, 99, 100, 103
executive functions, 39, 40, 70, 88, 98
granularity, 72, 79, 94
ICD, 19, 31, 32, 46, 48, 49
ICF, 19, 21, 31, 32, 42, 44, 46, 47, 48, 49, 50, 51, 52, 53, 56, 61, 62, 63, 66, 67, 70, 72, 97, 99, 101, 102, 105, 107, 108, 109, 110
individual, 31, 32, 61, 66, 67, 108
instances, 49
interface, 33, 34, 65, 67, 70
LCS, 75, 76, 79
myocardial infarction, 102, 107, 109
PCE, 19, 28, 42
population, 31, 43, 44, 61, 62, 63, 66, 67, 79, 107
prognosis, 28, 39, 72, 98, 100, 101, 104
scenario, 49
SNOMED CT, 19, 31, 32, 42, 46, 47, 48, 49, 50, 51, 53, 102
SPARQL, 19, 52
spinal cord injury, 19, 29, 39, 64, 98
state, 42, 48, 62, 63, 73, 74, 78, 81, 82, 84, 86, 88, 94
temporal abstraction, 38, 72, 79, 94, 104
time-series, 38, 71, 99
urination functions, 69, 72, 84, 95, 98
vMR, 32, 54, 55, 98
World Health Organization, 19, 31

Appendix I: Analysis of the socio-economic impact

Indicators of health outcomes can provide information about the quality of life gained after an intervention such as cardiac rehabilitation. These indicators, together with information provided by governments about the cost of lost quality of life in population and the cost of intervention to revert this, can be used to assess the impact of cardiac rehabilitation in patients who have suffered, e.g., an acute myocardial infarction. In this article Calvo et al. (2013) we present a novel methodology to perform this assessment and validated the benefits of rehabilitation in terms of enhancing quality of life and giving decision support to healthcare policymakers. This methodology, which can be used with data from existing medical health records, takes into account changes in people's functional status, represented by ICF values, resulting from interventions that produce quality-of-life gains but do not change an underlying medical diagnosis. Therefore, this chapter calculates the socio-economic impact in cardiac rehabilitation using the standard used in Chapter 3 to solve interoperability problems.

A 7-step methodology is suggested to assess the impact of performing cardiac rehabilitation (or any other intervention which can be reflected in changes in ICF values). Figure 1 shows a block diagram in order to give a visual representation of the procedure. In dotted lines, the results of each step in the calculation can be found, to better keep track of the method used.

Figure 1 Block diagram of the methodology used.



Step (1) The ICF values before and after rehabilitation for the ICF categories are identified. These categories are the ones relevant to perform the mapping into the EQ-5D-5L questionnaire. To compute EQ-5D-5L dimensions, if there are data of more than one ICF category, the average of their ICF categories will be performed following the methodology described in 3.1. Equation 5 – QALYs gained shows the mapping of the questionnaire into ICF categories according to Geyh et al. (2007). The visual analogue rating scale, which refers to how good or bad the user feels his/her health is "today", is also mapped into "emotional functions". We can omit Step 1 if we already have the EQ-5D-5L values.

Step (2) The EQ-5D-5L values related to the ICF values found in Step 1 are identified. The 0–4 range of ICF values is mapped into the 1–5 range of the EQ-5D-5L questionnaire to obtain the weighted health status of each patient before and after cardiac rehabilitation.

Step (3) The values of the five top items in the EQ-5D-5L can be calculated using a utility-weighted algorithm (Williams (1995)), which has been recommended for use in economic evaluation. The index value calculator, downloadable from the EuroQol Group website, automatically calculates the health-status weights before and after rehabilitation from the EQ-5D-5L profiles.

Step (4) These health-status weights are used in QALYs as quality-weighted years of life. When used in order to calculate the gains in terms of life expectancy, quality-adjusted after rehabilitation, the number of QALYs gained is obtained as in Equation 5 – QALYs gained.

$$QALYs\ gained = \sum_{t=a}^{a+L^i} \frac{Q_t^i}{(1+r)^{t-a}} - \sum_{t=a}^{a+L} \frac{Q_t}{(1+r)^{t-a}}$$

Equation 1 – QALYs gained.

where *L* is the life expectancy for the subject at age *a*; *i* is the condition of a person after the rehabilitation process; *t* represents the individual years within the range of life expectancy; *r* is the discounting rate along years ($r = 0.03$); *Q* is the weight of the quality of life related to a year of life, calculated from the EQ-5D-5L values. Therefore, L^i and Q^i are the life expectancy and the weight of the quality of life after the rehabilitation process. We assume $L^i = L$, which means the same years of life expectancy before and after rehabilitation, standardized to 80 years in males and 82.5 in females (Sassi (2006)).

Step (5) In order to make a global economic assessment, we convert QALYs gained to DALYs averted, as in Equation 2. QALYs and DALYs are related by a conversion factor Ca, defined in Equation 3, where C = 0.1658 and β = 0.04, which takes into account the age-weighting implicit in DALYs, giving more weight to ages around 25 years (Barendregt and Bonneux, 1996).

$$DALYs\ averted = QALYs\ gained \times C_a$$

Equation 2 – DALYs adverted.

$$C_a = C \times a \times e^{-\beta \times a}$$

Equation 3 – Conversion factor.

Step (6) The cost/DALY is identified, in order to give a global estimate of cost savings after having calculated the DALYs averted. The cost/DALY can be directly obtained from official documents or inferred from the cost associated to a disability in terms of DALYs lost, as in Equation 4.

$$Cost / DALY = \frac{Cost\ associated\ with\ disability}{DALYs\ lost\ due\ to\ disability}$$

Equation 4 – Cost/DALY.

Step (7) The overall cost savings due to a health intervention is calculated from the DALYs averted and the cost/DALY, taking into account the cost of the intervention, as in Equation 5.

$$Cost\ savings = (DALYs\ averted \times \frac{Cost}{DALY}) - Cost\ intervention$$

Equation 5 – Cost/DALY.

The following table shows the results obtained for mean QALYs gained and mean cost savings, when calculated for the seven populations described in Calvo et al. (2013). The proposed socioeconomic assessment, which was validated with data of 200 patients taking part in cardiac rehabilitation, showed that this health intervention offers, on average, a gain in QALYs of 0.6 and cost savings of 8,000 €.

QALYs gained and cost savings under different sorts of cardiac rehabilitation.

	QALYs gained	Cost saving
All patients (200 cases)	0.6	8,000 €
Only motor rehabilitation (92 cases)	0.5	9,000 €
At least motor and psychological rehabilitation (24 cases)	0.5	7,000 €
Only motor and psychological rehabilitation (14 cases)	0.2	−200 €
At least motor and sexual rehabilitation (94 cases)	0.6	9,000 €
Only motor and sexual rehabilitation (86 cases)	0.6	9,000 €
Motor, psychological and sexual rehabilitation (10 cases)	1.0	16,000 €

Based on the results obtained, performing motor, psychological and sexual rehabilitation at the same time results in the highest gain in quality of life and the most saving in costs. On the other end of the spectrum, performing motor and *psychological* rehabilitation at the same time results in little improvement in quality of life, and implies a positive comprehensive socioeconomic cost. As far as the other groups are concerned, they do not significantly diverge from the results obtained taking into account all patients for whom, on average, cardiac rehabilitation results in a gain in QALYs of 0.6 and cost savings of 8,000 €.

Due to changes of little magnitude observed, on average, in the quality of life of patients who undergo motor and psychological rehabilitation only and at the same time, it is suggested to allocate special attention when deciding about the eligibility for this combination of cardiac-rehabilitation processes. In these cases, according to the results obtained, adding sexual rehabilitation should be considered, if it makes sense for the specific patient under treatment. Nonetheless, it has to be taken into account that these results are based on changes in values of categories of the ICF, a classification which describes the functioning status using only five values. A patient could experience a slight recovery, to which the ICF is not sensitive. This representation in five levels particularly affects results when accounting for psychological changes using as data source the Beck Depression Inventory (BDI) questionnaire, which has 64 values.

These results show the importance of cardiac rehabilitation after acute myocardial infarction. Although other studies had already proved its value by calculating gains in quality of life and cost-utility ratios, those measures were only able to compare interventions that modify QALYs in the same domain. From cost-utility ratios, it can be concluded the most effective intervention based on the minimum cost needed to gain a QALY. Although this outcome gives an objective measure when comparing interventions, it must be applied to the same disease or dysfunction, since costs of treating different pathologies are in most cases not comparable.

This study provides a measure to more globally assess the cost savings of performing cardiac rehabilitation so it can be compared to any other intervention in the overall framework of healthcare. Moreover, the methodology proposed is also applicable to measure any kind of recovery that can be expressed as changes in ICF values before and after an intervention.