THESE DE DOCTORAT DE SORBONNE UNIVERSITE

Spécialité

Informatique médicale

Ecole Doctorale Pierre Louis de Santé Publique : Epidémiologie et Sciences de l'Information Biomédicale

Présentée par

Naiara Muro Amuchastegui

Pour obtenir le grade de

DOCTEUR de SORBONNE UNIVERSITE

Sujet de la thèse :

Développement d'un système avancé d'aide à la décision clinique : enrichir la connaissance issue des guides de pratique clinique avec l'expérience

soutenue le 05 de décembre 2019 à Paris
devant le jury composé de (par ordre alphabétique) :

M. Marc CUGGIA	PUPH / Université de Rennes	Rapporteur
M. Grégoire FICHEUR	MCUPH / Université de Lille	Examinateur
M. Joseph GLIGOROV	PUPH / Sorbonne Université	Examinateur
Mme Pascale KUNTZ-COSPEREC	PU / Université de Nantes	Examinateur
Mme Nekane LARBURU	Ingénieur de recherche / Vicomtech	Examinateur
Mme Silvana QUAGLINI	PU / Université de Pavie	Rapporteur
Mme Brigitte SÉROUSSI	PUPH / Sorbonne Université	Directeur
M. Alain VENOT	PUPH / Université Paris 13	Examinateur

Sorbonne Université
Bureau d'accueil, inscription des doctorants et base de données
Esc G, 2ème étage
15 rue de l'école de médecine
75270-PARIS CEDEX 06

Tél. Secrétariat: 01 42 34 68 35
Fax: 01 42 34 68 40
Tél. pour les étudiants de A à EL: 01 42 34 69 54
Tél. pour les étudiants de EM à MON: 01 42 34 68 41
Tél. pour les étudiants de MOO à Z: 01 42 34 68 51
E-mail: scolarite.doctorat@upmc.fr


Except where otherwise noted, this work is licensed under
http://creativecommons.org/licenses/by-nc-nd/3.0/

La vie est courte, l'art est long, l'occasion fugitive, l'expérience trompeuse, le jugement difficile.

Hippocrate.

Remerciements

Au Directeur de cette thèse, Madame le Professeur Brigitte Séroussi. Vous m'avez fait l'honneur d'accepter de diriger cette thèse. Je vous remercie pour les conseils précieux et rigoureux que vous m'avez prodigués. Recevez ici le témoignage de ma gratitude et de mon profond respect.

Aux rapporteurs de cette thèse, Monsieur le Professeur Marc Cuggia et Madame le Professeur Silvana Quaglini. Vous m'avez fait l'honneur d'être rapporteurs de cette thèse. Recevez ici le témoignage de ma gratitude et de mon profond respect. My sincere gratitude and deep respect to the reporters of this thesis, Professor Silvana Quaglini and Professor Marc Cuggia.

Aux examinateurs de cette thèse, Monsieur le Docteur Grégoire Ficheur, Monsieur le Professeur Joseph Gligorov, Madame le Professeur Pascale Kuntz-Cosperec, Madame Nekane Larburu et Monsieur le Professeur Alain Venot. Vous m'avez fait l'honneur d'accepter de juger cette thèse. Veuillez accepter mes sincères remerciements et soyez assurés de mon profond respect.

To my colleagues at VICOMTECH, especially those in the eHealth and Biomedical Applications department where I have been working for the past 5 years, thank you for giving me the opportunity to grow as researcher and for facilitating carrying out this PhD thesis at Sorbonne Université while simultaneously working. A special thanks to Iván Macía, head of the department, who has believed in me throughout this entire process.

A los amigos que he ido haciendo durante este periodo y los que llevan tiempo a mi lado, que me han sufrido y padecido pero sin los cuales estoy segura de que no lo hubiera conseguido. Gracias por vuestro apoyo, confianza y paciencia. Intentaré ser tan buena amiga para vosotros cuando me necesitéis como lo habéis sido para mí.

Nire familiari, zuek izan bai zarete benetan aurrera jarraitzeko indarra eman didazutenak, aita, ama eta Unai. Zuen laguntza eta konfidantza gabe azken hilabetetan ez nukeen lortutakoaz gai izango, ezta azken hilabetetan aurre egindakoari gainditzeko kapaz ere. Beti izango naiz zuekin zorretan, ez ditut eskerrak emateko haina hitz. Eskerrik asko denagatik, espero dut nitaz harro egotea. Asko maite zaituztet.

Table of Contents

Remerciements 2
Table of Contents 3
Glossaire 5
1. Introduction 1
1.1 Background 3
1.2 Problem Analysis Context 3
1.3 Objectives and Research Questions 4
1.4 Research Scope 5
1.5 Approach 6
1.6 Structure 8
2. State-of-the-Art 9
2.1 Clinical Practice Guidelines and their transition to Computer Interpretable Guidelines 9
2.2 Clinical Decision Support Systems 12
2.3 Limits of guideline compliance 14
2.4 Rating the Strength of Recommendation and Quality of Evidence of CPGs 16
2.5 Visual analytics in healthcare 19
2.6 Conclusions 21
3. Research Design Approach 23
3.1 Research Design Concepts 23
3.1.1 Clinical Knowledge Formalization 24
3.1.2 CIG Formalization Module 40
3.1.3 Ontology-based semantic validation 42
3.2 Methodology for an evolutive CDSS 43
3.2.1 Domain-independent CIG formalization 43
3.2.2 Authoring tool 46
3.2.3 Augmenting the clinical knowledge using experience 50
3.3 Visual Analytics 64
3.3.1 Decisional Events visualization 64
3.3.2 Real-World Data visualization 67
4. Use Case: Breast Cancer 70
4.1 Breast Cancer Knowledge Model (BCKM) 71
4.2 Breast Cancer Clinical Practice Guidelines 74
4.3 Non-compliance criteria definition 76
4.4 Breast cancer outcomes 78
4.5 Interaction of the guideline-based CDSS and the experience-based CDSS within DESIREE project 79
5. Validation 81
5.1 Technical Assessment 81
5.1.1 Unit test design and implementation 83
5.1.2 Integration test design and implementation 91
5.2 Clinical assessment 95
5.2.1 Experience generation from retrospective data in simulated BUs 95
5.2.2 Performance and clinical validation with prospective data in real BUs 97
6. Conclusions 101
7. Research contributions 103
7.1 International Conferences 103

7.2 Journals 105
7.3 Awards 105
7.4 Intellectual Property 105
8. Discussion and Future research 106
Bibliography 109
Annex 1 – List of ONK patient IDs corresponding to non-compliant decisions from retrospective cases dataset 118
Annex 2 – Experience-based rule generated for patient 10107 120
Table of figures 122
Table of tables 124

Glossaire

- **AGREE:** Appraisal of Guidelines for Research and Evaluation
- **AHRQ:** Agency for Healthcare Research and Quality
- **APHP:** Assistance Publique - Hôpitaux de Paris
- **BRM:** Business Rules Model
- **BU:** Breast Unit
- **CDSS:** Clinical Decision Support System
- **CIG:** Computer Interpretable Guideline
- **CPG:** Clinical Practice Guideline
- **DE:** Decisional Event
- **DRL:** Drools Rule Language
- **EMB:** Evidence-Based Medicine
- **EORTC-QLQ:** European Organisation for Research and Treatment of Cancer -Quality of Life Questionnaire
- **FHIR:** Fast Healthcare Interoperability Resources
- **GRADE:** Grading of Recommendations Assessment, Development, and Evaluation
- **GUI:** Graphical User Interface
- **HL7:** Health Level Seven
- **NCI:** National Cancer Institute
- **OCEBM:** Oxford Centre for Evidence-Based Medicine
- **ONK:** Onkologikoa Hospital
- **PBC:** Primary Breast Cancer
- **PCP:** Parallel Coordinates Plot
- **PRO:** Patient-Reported Outcomes
- **QoE:** Quality of Evidence
- **RCT:** Randomized Controlled Trial
- **RDF:** Resource Description Framework
- **RWD:** Real-World Data
- **RWE:** Real-World Evidence
- **SORT:** Strength Of Recommendations Taxonomy
- **SWT:** Semantic Web Technologies
- **SaaS:** Software as a Service
- **SoR:** Strength of Recommendation
- **TNM:** Task-Network Model
- **USPSTF:** US Preventive Services Task Force

1. Introduction

Clinical practice relies on the medical evidence reported over time and the latest scientific research results. It is continuously evolving and updating, making it difficult for clinicians to keep track of new evidence (Barth et al. 2016). In order to facilitate this task and engage clinicians into a standardized and evidence-based clinical practice, Clinical Practice Guidelines (CPGs) were developed. These documents summarize the latest evidence demonstrated in a specific medical area (Grimshaw and Russell 1994), translating it into a set of recommendations on how patients should be treated and improving the quality of care in clinical settings. Nevertheless, their effectivity, i.e. clinician's adherence to them (Grol 2001), is not as strong as expected due to several factors, such as lack of awareness, familiarity, agreement, self-efficacy, outcome expectancy and/or inertia of previous practices (Cabana et al. 1999). Other issues in adopting CPGs as the basis of actual clinical care are related to the problem of providing patient-specific recommendations (Seidling et al. 2010) and their implementation since they are paper-based documents, and therefore, their overall maintenance as well as keeping them up-to-date is a tedious and time-consuming task (Dueck et al. 2015). To overcome these issues CPG-based Clinical Decision Support Systems (CDSSs) have been proposed (Sim et al. 2001) to provide personalized evidence-based recommendations. These systems implement a computerized version of the CPGs (i.e. Computer Interpretable Guidelines or CIGs) in order to analyze the patient's available clinical data according to the latest evidence in the least amount of time and in the most reliable way. Nevertheless, implementing the guidelines as CIGs using CDSSs does not inherently ensure clinical compliance with these guidelines. Even if the CDSS facilitates the application of guidelines, several factors may still lead to a clinical decision that does not follow the provided guideline-based recommendation(s). Some examples may include a clinically complex case that is not well defined in the guidelines, and therefore does not have an adequate therapy assigned or the lack of the latest clinical evidence due to the difficulty of updating the guidelines (Hunt et al. 1998). Moreover, healthcare is promoting the shared decision-making between clinicians and patients, seeking better healthcare outcomes in populations as well as individual patients, while also helping to control the costs of care (Légaré and Witteman 2013). This step is very relevant since some of the guideline gaps are caused by not considering patient preferences or taking into account special individual characteristics for each clinical decision. Retrieving observational data and patients' quality of life outcomes related to the provided healthcare during routine clinical

practice results in an accumulation of Real World Data (RWD) beyond the limitations of Randomized Clinical Trials (RCT), which sometimes are not representative of the real population. This RWD information in clinical cases, therapeutic decisions, and observed outcomes may be useful in order to predict the expected outcomes for incoming patients with similar profiles and to promote self-consistency in the management of clinically non-compliant cases. The use of RWD may also potentially lead to the generation of Real World Evidence or Practice-based Medicine, beyond the limitations of existing CPGs.

This thesis proposes an experience-based CDSS that:

(i) models the clinical knowledge and clinical performance in a computerized way,
(ii) addresses the guideline limitations generating new experience-based knowledge and consequently tends to
(iii) improve the clinicians' compliance with the formalized knowledge, and
(iv) explore accumulated RWD to be used in future decisions, giving the possibility to study new clinical hypotheses using visual analytics tools over data.

The proposed experience-based CDSS formalizes the knowledge related to each clinical case and decision in order to evaluate the clinical performance based on different studied clinical outcomes and to report the cases where clinicians do not comply with guidelines, generating new rules that incorporate this new knowledge. Additionally, quality measurements, such as the usability and the strength of the newly generated rules are proposed for evaluating the clinical reliability behind the new formalized clinical knowledge. Finally, some visual analytics tools are developed, allowing the exploration of the gathered RWD, which may help in the detection of trends and patterns in the data that could be translated into new clinical hypotheses to be studied.

This chapter presents the motivation for this research and addresses the main objectives and approach of this work. It is organized as follows: Section 1.1 presents the background of this research. Section 1.2 analyses the existing problems among guideline implementation and compliance in current healthcare practice. Section 1.3 presents the objectives and research questions of this thesis. Section 1.4 addresses the scope of this work and Section 1.5 describes the approach. Finally, Section 1.6 presents the structure of the remaining thesis.

1.1 Background

Clinical Practice Guidelines (CPGs) aim to formalize the latest evidence-based medicine in a standard, documented way in order to facilitate the improvement and standardization of the care provided (Thomas 1999). In order to adhere to best practices with CPG implementation, Clinical Decision Support Systems (CDSS) are proposed as they implement CPGs in a computerized way for faster and more rigorous analysis of ever-growing clinical knowledge and provide the best clinical support, according to CPG formalized evidence, in the least amount of time (Sim et al. 2001). Nevertheless, there are several factors that may lead to non-compliance with the recommendations provided by the guidelines, such as taking into account patient preferences or limitations of the guidelines themselves when reporting the latest evidence, as these documents may be frequently updated as new evidence is emerging on a daily basis (Cabana et al. 1999).

Hence, in cases where clinicians do not comply with guideline-based recommendations, the therapeutic decision to be followed is made based upon the provider´s clinical experience. In current practice, experience or personal criteria are not tracked or stored, and consequently, this information is lost and not made available for the future decision of similar incoming clinical cases. Guidelines, and consequently CPG-based CDSSs, do not evolve or increase their knowledge in an automatic manner in order to manage those limitations. Retrieving and structuring this information in a computerized way would allow the inclusion of this knowledge throughout the CDSS knowledge base, increasing it with the clinician's experience. Moreover, the clinical impact of each clinical decision in the patients' outcomes is not studied to improve clinical practice or extend the available clinical knowledge in an integrative and systematic way.

1.2 Problem Analysis Context

Multidisciplinary Clinical Teams are promoted for their management of complex diseases since they involve all key professionals needed in order to take part in the decision-making process and treatment planning, discussing and agreeing on the best evidence-based therapeutic options for the patient (Taylor et al. 2010). CDSS usage is promoted to support these clinical teams when handling the clinical case study since they can implement different guidelines and evaluate all relevant and available patient data in real-time (i.e. during

multidisciplinary committee meetings), providing the best recommendations in a fast and traceable manner (Patkar et al. 2011). The effectiveness of the CDSS relies on a list of directives pointed out in (Bates et al. 2003) which encompasses:

(i) a short reasoning time to deliver a solution matching clinicians' criteria and providing all the needed information to reach that result,

(ii) providing different recommendations based on reliable, relevant and available clinical data of the patient according to the implemented CPGs,

(iii) fitting the system within the clinicians' reasoning process workflow, providing user-friendly interfaces that help clinicians using the CDSSs and give the desired support by displaying the knowledge in a simple, integrated, and easily interpretable way to final users, and

(iv) tracking the impact of the provided decisions and computing some quality and usability assessment measurements, which may suggest more accurate margins of the evaluated variables within the CPGs in an automatic way, based on the data gathered by the system, since both are closely related with the maintenance and update of the knowledge bases formalized within the CDSS.

Addressing these requirements, which are technical to a large extent, would facilitate the implementation of CPGs among clinical committees. Nevertheless, from the clinical knowledge point of view, CDSS may inherit some weaknesses from CPGs when supporting clinicians, such as (i) incompleteness of recommendations, (ii) lagging behind the latest medical evidence and (iii) non-inclusion of some relevant parameters for the decision-making process, such as patient preferences or clinician preferences (Séroussi et al. 2013).

1.3 Objectives and Research Questions

With the motivation to address the research issues related to the in-completeness, and limitations of CPGs, the generic goal of this research is to study how to enrich the CDSS knowledge base when only represented by CPG contents with the implicit clinical knowledge, mainly reflected as the clinicians' experience, considered when facing a clinical case not adequately addressed by the guidelines. More specifically, the main objective of this research is to implement an experience-based CDSS that:

- Retrieves and stores all the information related to each therapeutic decision, evaluating its clinical performance using different clinical outcomes
- During the decision-making process, provides the needed tools for knowledge gathering and formalization in a computer interpretable way for those non-compliant cases where guidelines are not sufficient to provide best recommendations to the studied patient
- Provides visual analytics tools to analyze the clinical data and healthcare performance results (i.e. clinical decisions and outcomes) in a user-friendly way

The research questions that motivate this PhD Thesis are the following:

- Is it possible to build a tool that facilitates the formalization and update of CPGs in a computerized way?
- Is it possible to extract the (implicit) clinical knowledge from the analysis of guideline non-compliant decisions, and use it to extend a strictly guideline-based CDSS?
- Could retrieving the information of all taken clinical decisions give insights into interesting clinical patterns following patients' outcomes?

1.4 Research Scope

In order to develop an experience-based CDSS, we began investigating the problem by examining CPGs and how they improve the quality of healthcare based on the latest reported Evidence-Based Medicine (EMB). First, causes for low clinicians CPG-adherence rates are studied, such as clinicians' limited knowledge about the guidelines, personal attitudes or trust among CPGs or behavior towards them (Cabana et al. 1999). CDSSs are proposed in order to overcome this type of issue, implementing CPGs as CIGs and facilitating CPG internalization in the decision-making process (Berner and Lande 2016). Even if this implementation increased guideline adherence rates, it still maintains some of the paper-based CPG deficiencies, such as poorly formalized knowledge for some clinical cases, the difficulty in maintaining and updating CPGs, complications handling complex clinical cases or the inclusion of patient preferences during the decision-making process (Bates et al. 2003).

Based on the results of the problem investigation carried out in the design phase of the research and aiming to overcome these issues, this thesis proposes an experience-based CDSS.

1.5 Approach

Figure 1 presents the approach adopted in this research. The rounded rectangles on the top row depict the main phases of the research. The square rectangles depict research activities, the results of which can be used in follow-up research. The direct arrows represent a result/input relation between the research activities. The approach applied in this research is divided into five main phases.

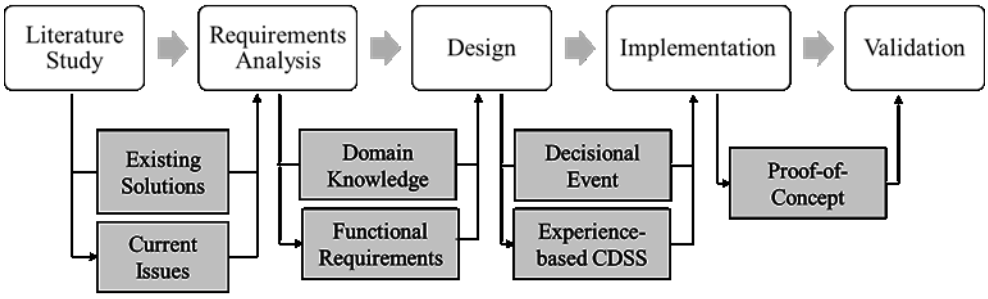


Figure 1: Approach diagram.

The first phase is the *Literature Study* that consists of presenting:

- *Existing Solutions*: to describe the state-of-the-art in CPGs and CDSS usage and address some of the solutions proposed in previous research works.
- *Current Issues*: to present some of the difficulties of CPGs' formalization into CIGs, knowledge gaps and limitations of CPGs, and provide the background information that has motivated this research.

The second phase is the *Requirements Analysis* for new knowledge inclusion during the decision-making process. It includes the following:

- *Domain Knowledge:* It supports the specification of the relation between general-context problem parameters (i.e. formalization of each clinical decision in a computerized way in which computerized guidelines will lean on) and specific-context problem parameters (i.e. formalization of new clinical knowledge based on CPGs non-compliant cases and the evaluation of the formalized contents based on patient outcomes) to augment the knowledge in purely guideline-based CDSS.
- *Functional Requirements*: This identifies the needed requirements when interacting with the guideline-based CDSS in order to include new

knowledge and evaluate the previous evidence formalized in the system based on the outcomes. These functional requirements define the interactions between the guideline-based CDSS, the experience-based CDSS, and the end-user.

The third phase is the design of the Decisional Event structure, which stores all the information related to the decision-making process and the experience-based CDSS, which tracks all of the decisions and generates new knowledge when a non-compliant decision is made.

This research has been conducted in the context of the DESIREE[1] European project, and hence, the design choices are the following:

- Formalize the Decisional Event concept in a computer processable way as the basis to retrieve, model, and exploit all the information related to the decision-making process.
- Develop an experience-based CDSS that will be integrated with the guideline-based CDSS to provide a tool that tracks the decisions and generate new knowledge when the guidelines do not address the studied clinical case.
- Provide some visual analytics tools to explore the generated real-world evidence through the experience-based CDSS.

The fourth phase is the implementation of the experience-based CDSS prototype. The implementation has been conducted in the primary breast cancer use case for supporting Breast Units (BUs) during the meetings for the decision-making process (see Chapter 4).

The validation, also conducted in the use case presented in Chapter 4, is the final phase with the following objectives:

- Verify the design requirements
- Demonstrate the usefulness of the experience-based CDSS for retrieving and formalizing the implicit knowledge when generating new rules that manage guidelines gaps

[1] http://desiree-project.eu/

- Embrace the validation of intermediate results that were essential for the development of the system's architecture. Specifically, the validation of the clinical performance based on the compliance, the outcomes, and the clinicians' interpretation of these results by using visual analytics graphs.

1.6 Structure

The remainder of this thesis is structured as follows:

- Chapter 2 presents the state-of-the-art in CPGs and their adherence and compliance limitations, explores their implementation as CIGs using CDSSs and analyzes measurements of strength and quality of the recommendations to evaluate the guidelines' knowledge. Additionally, provide some visual analytics tools that facilitate the interpretation of this information and make it more accessible to the clinician are analyzed.
- Chapter 3 introduces the main research design approaches for formalizing the clinical knowledge following the Decisional Event structure and the components of the experience-based CDSS, proposing some visual analytics to explore the outcomes of these systems in an intuitive way.
- Chapter 4 describes the implementation of the developments presented in Chapter 3 through a use case in the primary breast cancer domain.
- Chapter 5 evaluates the technical and clinical validation of the proposed system.
- Chapter 6 and Chapter 7 present the results, conclusions and some contributions of this research.
- Finally, Chapter 8 lists various directions for future work.

2. State-of-the-Art

This chapter describes the fundamental aspects of the current state of this area of research. Relevant topics for this thesis are covered, such as the evolution from Clinical Practice Guidelines (CPGs) to Computer Interpretable Guidelines (CIGs) following Evidence-Based Medicine (EBM) principles, formalization of knowledge and experience in the digital domain, evaluation aspects regarding strength of evidence and visual analytics tools for exploratory data analysis of Real World Data registered with such systems.

2.1 Clinical Practice Guidelines and their transition to Computer Interpretable Guidelines

Clinical Practice Guidelines (CPGs) are defined as explicit statements that model and summarize current evidence and clinical judgment, following Evidence-Based Medicine (EBM) principles for a standardized and best practice quality healthcare at the decision-making level (Lobach and Hammond 1997). Implementing CPGs has proved to be a valuable reference when supporting clinicians in their decision-making process as they can provide educational help for practitioners with less experience, improve the clinical care quality by assessing the evidence behind the recommended treatment, ensuring that best clinical practice is followed and help to avoid negligent medical practice or to reduce the biases from reported evidence (Silberstein 2005).

When implementing CPGs, several characteristics from the clinical and development points of view must be considered to ensure good healthcare quality levels and clinicians' satisfaction. Assuring the validity and reliability of their clinical content, along with their clinical applicability in real clinical settings, may help to engage clinicians in their systematic application. Moreover, they must be clear when defining the procedures to be followed and allow some clinical flexibility, being developed in a representative manner to coexist with the current clinical performance procedures within a healthcare system (Sackett et al. 1996; Thomas 1999).

Nevertheless, several clinician adherence barriers cause the dissemination of the guidelines to be tedious and difficult. These barriers are mainly caused by (i) lack of the awareness, (ii) lack of familiarity with the guideline provided recommendations (iii) lack of

agreement due to different clinical interpretations, simplification of the clinical knowledge reported in the guidelines or standardization of clinical cases, (iv) lack of self-efficacy, (v) lack of outcome expectancy, (vi) inertia of previous practice and (vii) other external barriers coming from the patients or environmental factors, out of the clinicians' control (Cabana et al. 1999). In conclusion, barriers related to clinicians' knowledge about the guidelines, attitudes or trust of or behavior towards the guidelines could affect their implementation, compliance, and adherence in real clinical settings (Tunis 1994).

Several methods for guideline integration in clinical settings have been explored, but many barriers still persist. Some studies propose that having timely feedback on the performance and how the clinical behavior changes based on CPG usage could increase the clinicians' likelihood for CPG adherence (Dykes et al. 2005). Including the clinician within the CPG formalization process and encouraging them with clinical performance analysis and study is highly recommended (Hysong, Best, and Pugh 2006).

Actual trends move towards highly interactive computerized systems, trying to intuitively present complex clinical cases, where clinicians may access and check computerized clinical data and take away insights from all of this information in a more natural and intuitive way (Liem et al. 1995). These systems are candidates for more easily accommodating a digital implementation of the CPGs providing evidence-based decision support (Garg et al. 2005). Another objective is to achieve a correct and good quality guideline formalization into computerized languages, following a consistent and adequate methodological development of the clinical processes and the objectives represented in the guidelines. Since CPGs are living documents that report the latest clinical evidence, maintaining and updating them often becomes mandatory. However, CPGs are expressed as textual documents which means their contents lag behind actual knowledge and require new versions based on the reported clinical knowledge being updated (Wang et al. 2002). Furthermore, CPGs are designed to support most common and evidence-backed clinical cases, making them standard and assuring their quality for usual clinical cases but being insufficient for patients that are in gray areas, where lack of evidence exists (e.g. excluded clinical cases in Randomized Controlled Trials or RCTs) or differ from the canon (Bates et al. 2003). In some cases, there are no CPGs formalizing the appropriate scientific evidence to be based on during clinical practice but data corresponding to the opinion of experienced physicians when providing therapeutic decisions is available. The usage of data mining techniques has been proved to help on identifying practice-based decision rules that go

beyond the formalized evidence for helping in the guideline reported evidence completion (e.g. detailing the duration of a treatment administration which is currently not defined in the guidelines but proven to influence the outcomes of the patients) and updating it when needed (Canavero et al. 2017; Toussi et al. 2009).

To facilitate implementation, dissemination, and maintenance-related barriers in the last decade, the representation of the clinical knowledge contained in the CPGs was translated into computerized implementation, known as Computer Interpretable Guidelines (CIGs). CIGs allow for the analysis of computerized clinical data coming from patient electronic medical records and contrasting it with the guidelines in an automatic manner, being able to provide more personalized and reliable advice or treatment recommendations. The following characteristics are key at the core of a CIG´s inherent success and in aiding in their dissemination and implementation throughout healthcare systems:

(i) the use of standardized clinical terminology that facilitates the understanding and univocal interpretation of the clinical data to be analyzed and the clinical knowledge formalized in CIGs,

(ii) the proposition of a model for easy updating the guidelines and facilitating their dissemination over the clinical community, and

(iii) the promotion of quality test tools for assessing the strength of CIG recommendations as a whole and for each of the provided recommendations. This will help in providing optimal personalized guideline-based recommendations at a reasonable cost and implementation effort (Latoszek-Berendsen et al. 2010).

Although several proposals for CIGs representation have been made, there is no leading standardization language that fully satisfies the requirements for the representation of the logic of CPGs (Votruba, Miksch, and Kosara 2004; Kaiser and Miksch 2005; Tu and Musen 1999; Wang et al. 2002). One of these approaches formalizes the clinical knowledge as “Task-Network Models” (TNMs), i.e. models that represent the dependency among actions, structured as hierarchical networks which, when fulfilled in a satisfactory way, provide recommendations (Peleg 2013). Several proposals have been reported following this approach aiming at managing with different clinical modeling challenges, such as GLIF (Boxwala et al. 2004), PROforma (Sutton and Fox 2003), or Asbru (Miksch 1999). Moreover, due to the vast amount of digitalized data coming from the electronic health records to be evaluated and the formalization

of the clinical processes in CIGs, it is highly recommended to apply Semantic Web Technologies (SWTs) (Blomqvist 2014) in order to process the data in a more effective and efficient way, create a proper framework for interoperability between systems and also integrate data from various sources (Argüello et al. 2009; Pruski, Bonacin, and Da Silveira 2011). In addition, along with SWTs, the implementation of standardized terminologies is highly promoted, guaranteeing the interoperability of the implemented knowledge and its univocal interpretation since it allows the representation of the biomedical concepts with stable and unique codes (Ahmadian, Cornet, and de Keizer 2010). Some of the most extended terminologies in cancer domain applications are SNOMED CT[2] and NCI Thesaurus[3] (Bodenreider 2008; Sioutos et al. 2007; Kumar and Smith 2005). Applying these kinds of approaches during the data acquisition and requirements definition process could alleviate the missing data or bad quality data gathering, which results in a poor CIG support and lower guideline compliance (Lanzola et al. 2014).

In conclusion, formalizing CPGs into CIGs allows the implementation of decision-support systems that provide patient-specific advice at the point of care. Computerizing guidelines permits the analysis of all patient information, not only focusing on the latest clinical results but also studying all of the relevant medical records in a reliable and efficient manner in the least amount of time, which will help in the inclusion of data mining techniques for identifying relationships between patient specific data, execution paths, process goals and achieved clinical results (de Clercq et al. 2004; Peleg, Soffer, and Ghattas 2008; Ghattas, Soffer, and Peleg 2014). Moreover, it facilitates CPG adherence and the measurement of clinical outcomes and performance related results, such as CPG compliance and the impact of the made decisions on the patients' healthcare, identifying guidelines' grey areas (Sim et al. 2001; Terenziani et al. 2008; Bragaglia et al. 2015; Hommersom and Lucas 2015; Lucas and Orihuela-Espina 2015; Panzarasa et al. 2010; Lanzola et al. 2014).

2.2 Clinical Decision Support Systems

Clinical Decision Support Systems (CDSSs) consist of computerized systems or software developments that aim aiding healthcare professionals in the diagnostic and therapeutic decision-making process (Payne 2000). When these CDSSs are CPGs based, the

[2] http://www.snomed.org/

[3] https://ncit.nci.nih.gov/ncitbrowser/

provided knowledge-driven clinical guidance is based on clinical CIGs knowledge implementation. These systems analyze the relevant clinical characteristics of an individual patient in order to provide patient-specific assessments or recommendations for the best decision-making. In the last decade, CDSSs have proven to be potential tools to improve clinicians' CPG adherence and to support ambulatory patients (Sim et al. 2001; Peleg 2013; Quaglini et al. 2013). Moreover, these systems are able to analyze considerable amounts of structured information coming from patient electronic medical records in a very short period of time, thus achieving an overall improvement in the health care practice, decreasing medical errors and variability while promoting guideline compliance (Sim et al. 2001; Berner and Lande 2016).

CDSSs must verify a list of design requirements in order to successfully support the clinical practitioner during the decision-making process and assure the acceptance as well as the adherence to the CPGs (Isern and Moreno 2008; Bates et al. 2003; Sittig et al. 2008). Some of those requirements are (i) providing a guideline repository that contains the latest available medical evidence for a given clinical domain and keep this knowledge base updated, (ii) have the ability to feed the CDSS directly from electronic medical records and be able to process the relevant information for each case, dealing with the missing data management efficiently (iii) evaluate the clinical data in the least amount of time possible avoiding the inclusion of excessive information which may be overwhelming during the decision-making process and (iv) fit within the clinical reasoning workflow and track its implementation and use impact by analyzing the guideline compliance and the decisions made over time (Lanzola et al. 2014).

From the development point of view, providing tools that would help to maintain the knowledge base updated and following a standardized language when defining the CIGs to be implemented is highly recommended, since this is an important constraint when trying to implement these systems to support medical teams in real clinical settings. For example, in the breast cancer domain, different prototypes of CDSSs that aid in managing care for breast cancer patients have been developed. The success of these prototypes during routine breast unit meetings, however, depends on periodic updates and constant maintenance of the knowledge base in order to upgrade their usage from purely supportive research tools (Séroussi et al. 2017).

Hence, studies have reported that CDSSs do improve care quality and decrease medical errors (Berner and Lande 2016) having a positive impact on the quality of medical practice, but they are quite constraining since they depend on the *a priori* defined domain knowledge and

including new knowledge or updating the implemented guidelines is still not an easy task. Therefore, providing tools that facilitate the implementation, update, and evaluation of computerized guidelines is crucial for the best quality and latest evidence-based clinical support through CDSSs.

2.3 Limits of guideline compliance

Even if CPGs proved to enhance clinical practice, several causes limit their effectiveness and, consequently, the adherence and compliance of clinicians with CPGs. (Grimshaw and Russell 1994; Davis and Taylor-Vaisey 1997).

The complexity of the medical domain makes the formalization of CPGs a difficult task to be achieved successfully. First, formalizing evidence is not a straight forward task and may not reach the correctness and knowledge definition level that clinicians would expect, since opinion and interpretation still have a huge influence on healthcare management. On one hand, CPG development procedures are quite constraining, considering that guidelines are developed for population healthcare management, assuming that the concept of a "standard" patient exits, but might be inaccurate or even wrong for particular patients in real populations (Hurwitz 1999). The opposite can happen as well, when small randomized clinical trials or controlled observational studies are used to report evidence that may need to be generalized, resulting in poorer outcomes when treating bigger populations (Shekelle et al. 1999). The importance of personalization of the guidelines is imperative in order to improve adherence and compliance rates. The work of (Bouaud and Seroussi 2002) states that for breast cancer management guidelines, 66% out of 127 patients fit correctly to be evaluated with standard guidelines, whereas 39% of the cases suffer a bias between the guidelines recommendation and the treatment administered.

A closely related and important issue is the guideline development process or how CPG development working groups are composed. Usually, these teams are comprised of quality auditors or managers who are guided by their opinions, interests, and experience, and who intend to formalize evidence seeking appropriateness of the provided recommendations but ignore the iterative and causal reasoning of clinicians (Woolf et al. 1999). Depending on the clinical context and according to the approaches followed for developing and disseminating, as well as the applied implementation methods, CPGs can be more or less successful when reporting the latest clinical evidence (Grimshaw and Russell 1993). Even if CPGs are audited

to rate their quality of evidence and strength of recommendations, trying to replicate the clinical reasoning process is difficult and translates into simplified, generalized, and in some cases ambiguous vocabulary, which may lack the supporting evidence and will require the clinicians' own opinions for its interpretation. The act of giving way to interpretation and providing purely clinical judgment-based recommendations can be very susceptible to bias and/or directly non-compliant with CPG-based recommendations and following one´s own self-interests (Shekelle et al. 1999). Defining the followed reasoning process as much as possible would help to track and identify the causes of these evidence gaps and to analyze the reasons behind biases from guidelines.

Another point to take into account is that current clinical care is moving towards patient-clinician shared decision-making since patient involvement can provide insights into best health states or outcomes in each case, apart from establishing a partnership that will help clinicians understand their patients' preferences (Say and Thomson 2003). There are particularly complicated cases in which making a clinical decision is a difficult task due to the trade-off between the level of observed symptoms and the impact that those symptoms could have on the patients' life, especially in those cases where the expected medical outcomes are similar for different clinical procedures (i.e. term referred to as "equipoise"), requiring an individualized and personalized healthcare process and the close interaction with the patient for the best decision (Hlatky 1995). Nevertheless, CPGs do not include evidence on patient preferences. It is necessary to overcome many barriers in order to ensure success in this task (Chong et al. 2009):

(i) consider patient preferences as population knowledge that follows some general trends and not only as individual one-off cases, subjective and variable factors, including them as part of the clinical evidence reported in the studies to identify "preference-sensitive" decisions (e.g. those decisions having lifelong implications or an uncertain benefit to the patient, unclear or conflicting evidence, risk of suffering side effects or negatively affecting the patient's quality of life, etc.) of high levels of uncertainty about best clinical procedure to follow (Krahn and Naglie 2008),

(ii) create a clear taxonomy (i.e. systematic categorization) for patients' preferences that will serve as a standardization over all of the involved disciplines (i.e. analysts, economists, clinical psychologists, etc.) that have different point of

views on the measurement of patients' preferences, to label and extract this information in a processable and understandable way (Bastemeijer et al. 2017; Luckmann 2001), and

(iii) build a methodology to synthesize the current evidence on preferences and be able to describe preference-based evidence along with clinical-based evidence, as it is proven to strongly influence the decision-making process (Noble et al. 2015; Froberg and Kane 1989).

In conclusion, even if CPGs aim to improve healthcare outcomes through standardized clinical procedures, their generalized approach lacks significant relevant information, causing low clinical adherence and considerable non-compliance rates in real clinical performance. To overcome these issues, the implementation of more flexible guidelines that facilitate shared decision-making and take into account patients' preferences is being promoted (van der Weijden et al. 2010). Understanding CPGs limitations, identifying "gray" areas (i.e. cases that are complex and which guidelines are not capable of providing satisfactory support) and providing timely feedback to clinicians about compliance rates, guideline biases, and outcomes could significantly improve significantly the clinicians' adherence to CPGs and the quality of the healthcare provided.

2.4 Rating the Strength of Recommendation and Quality of Evidence of CPGs

CPGs rely on the latest EBM to guide clinicians in the decision-making process. Scales such as the Appraisal of Guidelines Research and Evaluation (AGREE) assess the quality of the guideline's development process, not focusing on the clinical content and the quality of the evidence of the provided recommendations (AGREE Collaboration 2003). Hence, to what extent are the recommendations provided in the CPGs based on high-quality evidence? What is considered as high-quality evidence? How can clinicians and CPG developers be confident about those recommendations?

In the last decade, several approaches have been developed in an attempt to answer these questions and formalize evidence-grading systems. The Agency for Health Care Quality and Research (AHRQ)[4] reviewed the ongoing efforts of different medical groups and reported that

[4] https://www.ahrq.gov/

there are currently over 100 proposals for grading the evidence of the guideline recommendations (West et al. 2002). Since many of these approaches were complex and difficult to integrate in daily clinical practice, the AHRQ stated three key elements to be covered by any evidence grading system that would facilitate their dissemination throughout the clinical community (Clair 2005): (i) quality, referring to the validity of the study or the minimal opportunity of bias that it could have, (ii) quantity, when talking about the number of studies taken into account to formalize that evidence and the number of subjects studied within them and, (iii) consistency among other studies on the same topic that could be comparable. Some of the approaches that do accomplish these criteria are the Oxford Centre for Evidence-Based Medicine (OCEBM) Levels of Evidence[5], the Cochrane Collaboration[6], the US Preventive Services Task Force[7] (USPSTF), the Strength of Recommendations Taxonomy (SORT) and the Grading of Recommendations, Assessment, Development and Evaluation[8] (GRADE). The first five are more focused on reporting evidence based on patient-oriented outcomes, which may disagree with disease-oriented outcomes. For example, when analyzing a disease or condition such as providing Doxazosin for treating hypertension or high blood pressure, a disease-oriented outcome would be that it reduces the patient's blood pressure to prevent suffering a stroke whereas a patient-oriented outcome reports that this same treatment increases mortality in people of African ancestry. Patient-oriented outcome approaches are more simplistic in order to facilitate their implementation throughout CPGs and are mainly developed for specific clinical domains or illnesses (Ebell et al. 2004a).

To provide reliable measurements of the quality of the provided recommendations the Quality of Evidence (QoE) and Strength of Recommendations (SoR) are defined. The QoE reflects how confident we are with the provided recommendation(s), and the SoR defines the evidence supporting that recommendation and the benefits/risks tradeoff when following it. Focusing on SORT, this scale provides a uniform rating system, simple and easy to use, for rating the QoE and SoR based on patient-oriented outcomes. The rating system is based on 3 levels of SoR of a body of evidence (*A*, for recommendations based on consistent and good-quality patient-oriented evidence, *B* for recommendations based on inconsistent or limited patient-oriented evidence and *C* for recommendations based on evidence deduced over

[5] https://www.cebm.net/2016/05/ocebm-levels-of-evidence/

[6] https://www.cochrane.org/evidence

[7] https://www.uspreventiveservicestaskforce.org/Page/Name/recommendations

[8] http://www.gradeworkinggroup.org/

consensus, usual practice, opinion or disease-oriented evidence) and 3 levels of QoE (1, meaning good patient-oriented evidence, 2 for limited patient-oriented evidence, and 3 for other kind of evidence, such as consensus, usual practice, opinion or disease-oriented evidence) (Ebell et al. 2004a). Since this approach is based on patient-oriented evidence instead of disease-oriented evidence, it is still insufficient or poor when used on its own. Moreover, this rating system is not able to effectively manage some particular qualitative results, since SORT does not address these type of recommendations (Ebell et al. 2004b).

GRADE, on the other hand, has been adopted by over 65 organizations worldwide trending to be the international benchmark for rating QoE, and SoR in a transparent and explicit way (Guyatt et al. 2013). The primary keypoints of this rating system are (i) the clear separation between QoE and SoR, which means that a particular QoE does not necessarily imply a particular SoR, (ii) the inclusion of patients' outcomes, (iii) identifying explicitly the factors that downgrade (i.e. limitations in the study design, inconsistency or imprecision of the results, indirectness of the evidence, publications bias) or upgrade (i.e. large magnitude of effect, the underestimation of true treatment effect caused by biases results) the QoE of a recommendation, (iv) the transparency of the process of formalizing evidence into recommendations, proposing first the clinical question or recommendation to be studied, then reporting the treatment effects and critical outcomes from available evidence to ultimately assess its confidence when evaluating the followed evidence reporting method and finally analyze the tradeoff between the benefits and risks of following that recommendation, (v) grading the quality of the available evidence on diagnostic strategies, (vi) explicit advice and guidance among values and assumed preferences when making a recommendation even in scarcely available evidence cases, (vii) clear and pragmatic interpretation of SoR levels into "Strong" when the benefits outweighs the risk of following the recommendation, "Strong against" when risks overweigh benefits and "Weak" when risks and benefits are balanced and (viii) simple but methodologically comprehensive approach for rating QoE in 4 grades, "High" when further research won't change the confidence on the expected treatment effect, "Moderate" when further research is likely to have an important impact on the estimated confidence of the treatment, "Low" when further research is very likely to affect the confidence estimation and "Very Low" when the estimation of the effect of the analyzed treatments are unclear (Brożek et al. 2009; Balshem et al. 2011; Maymone, Gan, and Bigby 2014).

In conclusion, GRADE is the most frequently implemented SoR and QoE grading system because of its comprehensive, explicit, and transparent methodology when rating a recommendation to treat a patient. It guides clinicians, aiming to provide the best health care with the most recent evidence and information available in the most objective way. Nevertheless, the assessment of QoE is dependent on subjective opinion, since each step requires clinical judgment and it cannot be completely determined objectively, not assuring the consistency through these assessments.

Hence, measuring the clinical performance, the biases from the latest reported evidence within the CPGs and gathering the outcomes of the patients that followed those treatments to keep evidence as updated as possible is crucial. Nevertheless, accomplishing each of these tasks in a consistent and objective way is still a challenge.

2.5 Visual analytics in healthcare

The clinical data available is increasing exponentially in recent years along with the digitization of healthcare systems. Exploiting these large amounts of heterogeneous data may provide insight for improving healthcare's effectiveness and efficiency, but due to the datasets' magnitude and complexity, these conclusions are difficult to obtain and demonstrate in real clinical settings (Sun and Reddy 2013). Clinicians are overwhelmed by the large amounts of heterogeneous and scattered information they are receiving which in turn requires extensive efforts for their interpretation. Leading to a conclusion about the implicit relationships in the data that could influence patients' health conditions is not a straightforward task. Due to this information overload, some crucial variables and relationships may be ignored, misinterpreted or missed, causing a negative impact on the patient outcomes and clinical performance (Vaitsis, Nilsson, and Zary 2014). To overcome these issues, visual analytics, which is the science of displaying information through easy-to-use interactive interfaces focused on analytical reasoning, is proposed (May et al. 2010). Visual analytics offers timely information in an intuitive and interactive format, facilitating the hypothesis generation, reasoning, and interpretation of the complex data for a given population (Caban and Gotz 2015). Moreover, it permits the discovery of unknown hidden implicit information patterns by highlighting the connections through the analyzed variables within a dataset, customizing the queries to be carried out depending on the formalized hypothesis in each case and allowing the visualization

of complex ideas in a clear and precise way, which is not possible using other approaches (Simpao et al. 2014).

One of the most widespread techniques for visualizing complex multidimensional datasets for discovering patterns among data are the Parallel Coordinates Plots (PCPs) (Inselberg and Dimsdale 1990; Cuzzocrea and Zall 2013) (see Figure 2).

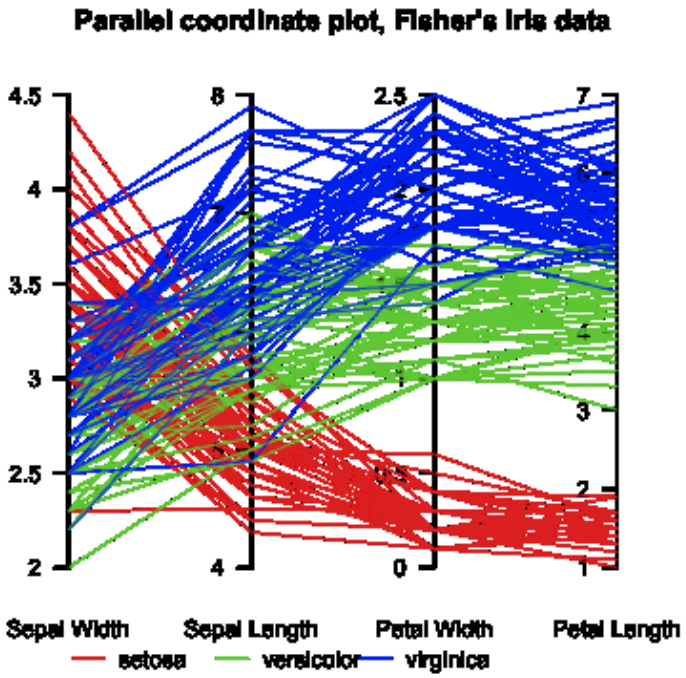


Figure 2: An example[9] of a parallel coordinates plot representation of the Iris Dataset[10].

This technique consists of the visualization of the variables of the dataset as parallel vertical axis while each of the entries or samples draws a horizontal line, linking each matched value of the vertical axis and providing a comparative and continuous view of the data patterns. The visualization of large amounts of data through PCPs allows the integrity of the combination of represented results, the facility to track the path drawn by the data, easing the analysis and causes of it, and provides a way of interaction with the data to explore the highlighted consequences (Cuzzocrea and Zall 2013).

Even if PCP have been proven to be a powerful technique to visualize multidimensional clusters, they can have some scaling problems when visualizing large amounts of data, since

[9] https://upload.wikimedia.org/wikipedia/en/4/4a/ParCorFisherIris.png

[10] https://archive.ics.uci.edu/ml/datasets/iris

there are many intersections and overlapping lines within the visualization that can hide the relevant patterns, creating visual clusters (Holten and Wijk 2010). To minimize this issue, several modifications have been done in PCP representation, such as (i) dimension reordering or reducing the number of variables according to their similarity and relevance for the searched hypothesis, (ii) clustering and filtering the data to show it in an aggregate manner that would simplify its visualization, (iii) apply interactive techniques that help to summarize and visualize line subsets and (iv) using visual enhancement techniques that would help in the identification of the patterns among the data (Zhou et al. 2008).

In conclusion, the potential insights that could be extracted from these large and complex amounts of data may have an important impact on healthcare quality leverage, not only from the research point of view but also in evaluating the performed clinical decisions and their impact on patients' healthcare and outcomes. Nevertheless, the analysis of this vast scale of data is a substantial obstacle to overcome (Shneiderman et al.). The best way to succeed in this task is by analyzing the clinical domain of study and the available data to propose the solution that best fits the actual case needs and research objectives (Gotz, Sun, and Cao 2012). Visual analytics techniques, along with machine learning approaches, are proposed as a promising tool for helping clinicians in current healthcare performance analysis, identifying critical patient groups and validating with real-world evidence regarding the knowledge reported in the clinical guidelines.

2.6 Conclusions

This chapter has analyzed the transition from EBM to the formalization of CPGs and their computer implementation as CIGs to be used by CDSSs. This evolution from paper-based to computer-based guidelines was aimed at improving the adherence to and the promotion of standardized medical practices based on the reported evidence among clinicians for better healthcare quality and patients' outcomes. Previous research has highlighted some adherence barriers and guidelines limitations, such as the difficulty in maintaining the guidelines updated, the identification of gray areas related to preference-based medicine and the missing bias information and tracking of the decisions made over time among others. Regarding the quality of the evidence reported in the guidelines and how confident clinicians can be when following them, the latest scales to measure the SoR and QoE have been explored, since they can be a key factor in promoting adherence among clinicians when consistently and transparently defined

and implemented. Finally, visual analytic techniques have been studied as a tool that helps to understand clinical performance and may allow for the identification of factors to leverage them in a rapid and interactive manner. Therefore, this thesis aims to go beyond this state-of-the-art by proposing a solution that will address the current barriers, as explained in the following chapters.

3. Research Design Approach

This chapter describes the approach followed to design an architecture that:

(i) formalizes the clinical knowledge and new evidence in an interoperable and computer-interpretable way introducing the concept of the Decisional Event,
(ii) provides an experience-based decision support system to augment the knowledge of purely guideline-based CDSS by incorporating guideline non-compliance, and
(iii) research among previous clinical cases and their outcomes seeking new insights leaning on visual analytics techniques.

This chapter is organized as follows: Section 3.1 summarizes the different components needed for the structuration of the Decisional Event concept and CPG formalization in a computer interpretable and semantically interoperable way, serving as the backbone of the rest of the methods presented in this chapter. Section 3.2 presents a domain-independent methodology for CIG formalization, along with an authoring-tool that supports and eases this process, and an experience-based CDSS, which increases the knowledge of purely guideline-based CDSS with new experience-based rules generated from tracking the clinical performance, decisions, and guideline compliance gaps. To conclude, Section 3.3 introduces a visual analytics research tool that enables the evaluation and analysis of the clinical performance of all decisions made over time, which are gathered and reported through the experience-based CDSS in an intuitive and interactive way.

3.1 Research Design Concepts

In this chapter different formalizations are proposed to retrieve, structure, and reuse the clinical knowledge in a standardized and computer interpretable way in order to support the decision-making process and generate new evidence from the different decisions made over time.

A major contribution of this work is the definition and implementation of a virtual data structure, named Decisional Event (DE), which, by following object-oriented programming design principles, allows gathering all the information needed during the decision-making process in a computer processable way. Moreover, using this structure, we provide a way to

augment the actual knowledge with clinical experience that was untracked before. In (Shafiq et al. 2014) the authors presented a knowledge representation of an engineering object, which embodies all associated knowledge and experience within it and leaned on a flexible and standard knowledge representation structure to acquire and store experiential knowledge. This representation was used in (Sanchez et al. 2014) to semi-automatically update the underlying knowledge bases and decision criteria coming from previous experience knowledge (Muro et al. 2016) and to extract experience from patients Electronic Health Records (EHRs) dually represented as an Archetype Model (i.e. definition of the clinical contents) and a Reference Model (i.e. definition of the clinical contents representation structure), by adding a Decisional Model that allows the experience retrieval and usage. Nevertheless, this object did not formalize the CPGs containing Evidence-Based Medicine definition, focusing on the modeling of knowledge through ontologies.

The primary contribution of this thesis is a new formalization of the Decisional Event structure, which encompasses the CPGs formalization as CIGs and later their implementation in CDSSs in order to support clinicians in their decision-making process, track clinical practice performance as well as biases from guidelines' recommendations and outcomes of the treated patients, new real-world evidence that will aim to address the gaps in the guidelines (Larburu et al. 2019). Finally, standard communication protocols and languages have been integrated within the formalization process in order to ensure semantic and technical interoperability among the different modules of the system and facilitate the communication with external platforms.

3.1.1 Clinical Knowledge Formalization

The formalization of the clinical knowledge contained in CPGs into a computer-interpretable structure is a mandatory step in order to move towards a computerized healthcare system, a standardization of the decisions made over time, and ease the assessment of new evidence and patient outcomes'-based results. The transformation of CPGs into CIGs is a legitimate issue since it is a time and resources consuming task. Moreover, providing some tools to ease their maintenance and updating would be helpful. Retrieving clinical performance information would allow the comparison of the provided clinical care among similar clinical cases based on patients' outcomes and identify biases from the guideline-based recommendations.

3.1.1.1 Decisional Event Structure

A decision can be defined as the final conclusion or a resolution reached after analyzing the available information, ideally, based on some validated criteria generated from previous evidence or personal experience. A decision in the clinical domain seeks to provide the best personalized care for the analyzed patient, based on the latest clinical evidence along with the clinicians' intuition and experience (Bate et al. 2012). In this thesis, decisions were formalized into a digital structure named the Decisional Event (DE) to (i) gather all the information related to the decision-making process, reflecting all the rationality for taking a decision and the consequences of such decision for the patient and (ii) allow its later analysis to retrieve information about CPG compliance rates and new clinical evidence, based on the clinical success reported on different patients' outcomes. The DE structure is defined by the following set of components:

- $\boldsymbol{P = \{P_i\}}$: A set of clinical patient parameters that define the clinical case. This structure is flexible enough to handle any kind of clinical data coming from any type of clinical domain, supporting different kinds of data type values (e.g. Boolean, integer).
- $\boldsymbol{R = \{R_j\}}$: A set of clinical statements or rules expressed in a computer-interpretable way as IF-THEN specifications. This formalization originally comes from conditional computer programming, where different statements or Boolean conditions are defined to be accomplished and provide a conclusion or a recommendation. In this case, statements were formalized following the knowledge coming from different clinical sources (e.g. CPGs, local guidelines, experience-based rules generated by the system). Each of the formalized rules consists of the following items:
    - $\boldsymbol{A = \{A_m\}}$: a set of the clinical statements that compose the *conditional part* of the rule (i.e. the IF-part) *a priori* defined in CPGs. These clinical statements must be checked by the clinical parameters $\boldsymbol{\{P_i\}}$ to trigger the consequent part, meaning that they must achieve the mathematical condition imposed on all elements of the set, which in turn, when they are multiple, are connected by logical operators (i.e. and, or). If all conditions are satisfied, the recommendation or THEN-part will be given as a conclusion or recommendation for the studied case.

- ***W***: the recommendation(s) defined as the *consequence part* of the rule (i.e. the THEN-part) describes the conclusion or the recommendation(s) to be followed when the conditional part of the rule was accomplished by the studied clinical case. Each recommendation can be composed of a single order or an ordered list of orders. We identify mainly two kinds of orders: (i) *characterization orders*, whose objective is to provide or define a value for a clinical parameter depending on the clinical variables studied in the conditional part and (ii) *action orders*, which will define the clinical or therapeutic action to perform. A recommendation can be composed of several orders of one type or both types. Hence, giving a recommendation could change the value of some clinical variables while also suggesting an action to perform, expressed as $W = (S_1, S_2, \dots S_l)\ with\ l > 1$ where $\boldsymbol{S_i}$ is each atomic order.

- ***FD:*** the *Final Decision* represents the decision of the clinical team at a time t_0. This final decision is defined as an atomic recommendation from the structural point of view and it can be compliant with one of the recommendations provided by the rules, ***W***, or define a new proposition coming from the clinicians' experience or new clinical evidence not reported in the guidelines, hence being non-compliant with them.
- ***E***: the *Executed Treatment* administered at time t_1 by the clinical team to the patient after the decision was made, is also defined as an atomic recommendation from the structural point of view and it can be compliant with one of the recommendations provided by the rules ***W***, and the final decision ***FD***, taken at the time t_0 or be non-compliant to one or both of them. We measure compliance at these two levels because independently of the compliance of the final decision with the guideline-based recommendations, the final executed treatment may be non-compliant due to new information that was not known at the time of the decision was made. These biases must be recorded since ***E*** is the real treatment given to the patient.
- $\boldsymbol{C = \{C_k\}}$: the set of criteria followed by clinicians to reach an agreement about a final decision ***FD***. These criteria are sorted in different groups that have a closed list of possible Boolean values $\boldsymbol{J_n}$. A single criterion would be defined as a set of justifications $C_1 = \{J_1, J_2, \dots J_n\}\ with\ n \geq 1$. The information stored in

these criteria allows justifying and understanding the reasons behind a non-compliant decision. For example, having a complicated case due to the advanced age of a patient could require informing about the suitability of the provided treatment. Hence, a "*Patient-related restriction"* criterion could be defined that would contain a justification identified as "*Advanced age*" which could be either *true* or *false* (i.e. Boolean value range). These criteria and the justifications within them are defined along with the clinicians for each clinical domain use case, taking into account their experience and intuition when biasing from guideline recommendations. They can also include administrative or patient-related restrictions since they must reflect all possible bias causes (i.e. clinical or external).

- ***O(t)***: the set of clinical outcomes of the patient in a time t_n are stored to be able to assess the success or failure of the given treatment and hence, evaluate the quality of the decision made. Different kinds of outcomes can be stored in order to have a global vision of the decision and its impact, not only from the clinical point of view but also by taking into account patient-reported outcomes.

After defining the DE structure, a methodology for its implementation is presented next to retrieve all this information in a computerized way during the decision-making process of a real clinical setting.

3.1.1.2 Decisional Event Object formalization in Java

Considering the Decisional Event concept defined before, a Java[11] programming language-based formalization (see Figure 3) was implemented to be able to apply the proposed theoretical approach through a practical methodology, gathering all the decision-related information in a computer-processable way.

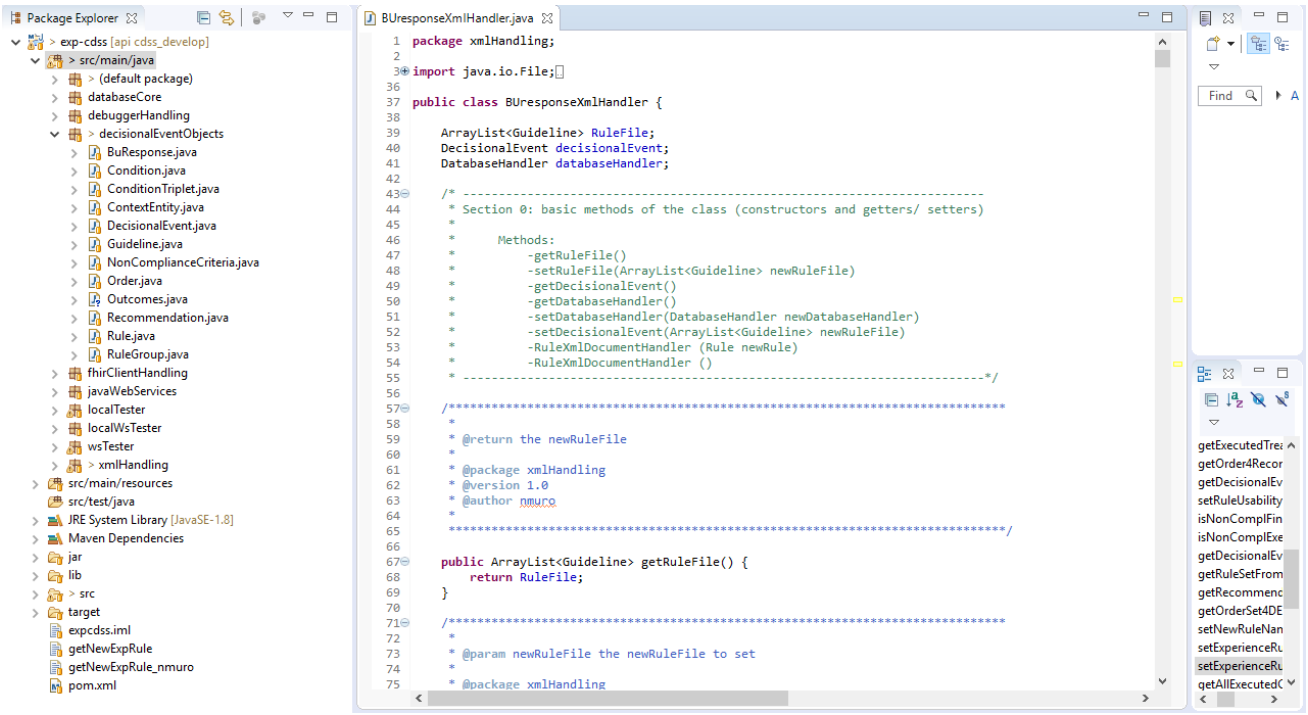


Figure 3: Eclipse editor overview with all the defined packages. The classes under the *decisionalEventObjects* package, and that will be described in this chapter, are deployed on the left part of the application.

Since Java is an object-oriented programming language (Deitel and Deitel 2011), each of the components of the Decisional Event was defined as a compound of 11 main Java classes (i.e. programming archetype objects), each of them with their own attributes, reflecting the different conceptual relationships and dependencies among them. These classes are described below:

- *DecisionalEvent* class: this class comprises the following attributes definition, which matches with the description presented in Chapter 3.1.1.1:

  (i) the ID of the patient on which the decision was made, which identifies the set of clinical parameters, $\boldsymbol{P} = \{P_i\}$, defining that patient,

[11] https://www.java.com/en/

(ii) the hospital or clinical unit where that decision was made,

(iii) the clinical scenario classification depending on the studied clinical domain, defined as the clinical situation in which the patient was at the decision-making time, since each scenario may face different kinds of needs and decisions can be based on the different relevant clinical data for each scenario,

(iv) the set of CPGs used to evaluate the patient and provide evidence-based treatment support,

(v) the set of rules, $\boldsymbol{R = \{R_j\}}$, within those CPGs that were triggered, since their conditional part, $\boldsymbol{A = \{A_m\}}$, was accomplished by the values of the clinical parameters of the patient, $\boldsymbol{P = \{P_i\}}$,

(vi) the set of recommendations, $\boldsymbol{W}$, coming from those triggered rules, $\boldsymbol{R = \{R_j\}}$, that provide the different therapeutic options defined in the queried CPGs for that particular patient,

(vii) the final decision, $\boldsymbol{FD}$, made by the clinical team and

(viii) the real executed treatment, $\boldsymbol{E}$, after a time *t*.

(ix) For compliance tracking, the non-compliance criteria, $\boldsymbol{C = \{C_k\}}$, followed in case of bias from the CPG based recommendations are also stored, along with two Boolean variables that gather if the chosen final decision and the real executed treatment were guideline compliant, compliant among them or none of the options.

(x) Finally, a set of outcomes, $\boldsymbol{O(t)}$, of the patient are also saved after a time *t* to evaluate the clinical impact of this Decisional Event.

The *DecisionalEvent* class refers to 5 different classes (see) within it, which implement each of the formalized concepts, i.e. *Guideline*, *Rule*, *NonComplianceCriteria*, *Recommendation,* and *Outcomes* classes. First, the *NonComplianceCriteria* and *Outcomes* classes will be explained. These classes are required to compute compliance and the usability and strength of the given recommendations. Afterward, the classes used to define the clinical evidence CIG alike will be presented.

```
/**
 * @author nmuro
 *
 */
public class DecisionalEvent {

    String patientID;
    String scenario;
    String unit;
    ArrayList<Guideline> guidelineSet;
    ArrayList<Rule> ruleSet;
    ArrayList<NonComplianceCriteria> nonComplCriteria;
    ArrayList<Recommendation> recommendationsSet;
    Recommendation finalDecision;
    Recommendation executedTreatment;
    boolean compliantWithFD;
    boolean compliantWithET;
    Outcomes outcomes;
```

Figure 4: The *DecisionalEvent* Java class and its attributes declaration.

- *NonComplianceCriteria* class: As stated in the *DecisionalEvent* class definition, the *NonComplianceCriteria* class (see Figure 5) is used in those cases where the final decision is non-compliant with recommendations provided by the CPG-based CDSS. This class gives some hints about this non-compliance and attempts to formalize the clinicians' intuition in a processable manner. To do so, different criteria groups, $\boldsymbol{C = \{C_k\}}$, that have several justifications are defined along with the clinicians when studying a particular clinical domain, which will address those knowledge gaps or will somehow formalize clinicians' experience or intuition.

```
/**
 * @author nmuro
 *
 */
public class NonComplianceCriteria {

    String nonComplianceCriteriaGroup;
    ArrayList<String> Justifications;
```

Figure 5: The *NonComplianceCriteria* Java class and its attributes declaration.

- *Outcomes* class: this class records different treatment and patient-related outcomes to compute the strength of the recommendation that was followed when treating the patient (see Figure 6). This information is retrieved as part of the clinical parameters of the patient after a time *t* of the first administered treatment. In our approach, these outcomes report mainly the toxicity grade and adverse events, if any, based on the criteria defined in the Common Terminology Criteria for Adverse Events (CTCAE) ("Common Terminology Criteria for

Adverse Events (CTCAE) | Protocol Development | CTEP" n.d.) and the quality of the treatment response (i.e. partial, complete, progression of the disease and stable disease as defined in the Response Evaluation Criteria In Solid Tumors (RECIST)[12] criteria (Eisenhauer et al. 2009)). To include the patient's perspective, International Consortium for Health Outcomes Measurement (ICHOM)[13] questionnaires ("ICHOM | ICHOM Standard Sets) were modeled to be part of the scoring of the rule recommendation's strength.

```
/**
 * @author nmuro
 *
 */
public class Outcomes {

    int CTCAE_Grade1_computation;
    int CTCAE_Grade2_computation;
    int CTCAE_Grade3_computation;
    int CTCAE_Grade4_computation;
    int CTCAE_Grade5_computation;
    ArrayList<String> treatmentResponses;
    int ICHOM_computation;
```

Figure 6: The *Outcomes* Java class and its attributes declaration.

During the decision-making process, a patient is studied following the available evidence in the CPGs. Next, the classes built for gathering and formalizing this clinical knowledge in a computer interpretable way are described (see Figure 7).

- *Guideline* class: this class contains guideline related information (e.g. the guideline name, the version number of the implemented CPG, the author or institution that has developed it and its release date) to identify each of the computerized guidelines and their information when asking for some recommendations.

A guideline is composed of several rule groups since it helps to classify the contained knowledge in a more organized way, for example, by scenario or treatment type. *DecisionalEvent* class alike, the *Guideline* class refers to a class named *RuleGroup*.

- *RuleGroup* class: this class defines all rule groups within a guideline by the following attributes: the name of the rule group, the ID used to identify it and

[12] https://recist.eortc.org/

[13] https://www.ichom.org/

the set of rules that composes it (i.e. ruleGroupName, ruleGroupID and rules in Figure 7).

Within the *RuleGroup* class, the *Rule* class is being pointed out. The *Rule* class is one of the most important classes from the development point of view, due to its complexity and relevance when generating new knowledge using the experience-based CDSS.

- *Rule* class: this class has a number of attributes, as follows:

  (i) the name of the rule (i.e. *ruleName*) as it is defined in the CPG,
  (ii) the ID (i.e. *ruleID*) used to identify it, and
  (iii) the author (i.e. *ruleAuthor*) of the rule (e.g. an institution such as the National Comprehensive Cancer Network (NCCN) or a particular hospital),
  (iv) the *ruleContext* array (i.e. container object that holds a fixed number of values of a single type) composed by *ContextEntity* class instances, storing the information used to validate all the clinical parameters defined in the rule towards a semantic model, such as an ontology. The attributes that compose this class define, for example, the ontological class to which the clinical variables are part of, the relationships among them or the class containing the possible values that these variables could take. The objective of this class is to generate a rule context for each rule with all the needed semantic information, which will assure its semantic interoperability and correctness.
  (v) The *ConditionsDescription*, composed by a *ConditionTriplet* array, defines all conditions of the conditional part of the rule, $\boldsymbol{A = \{A_m\}}$, as triplets (i.e. a precedent condition, a binary operator that links it with next conditions (if any) and a consequent condition, to be able to have traceability and maintain dependencies among conditions), whereas the
  (vi) *recommList* stores the list of given recommendations, $\boldsymbol{W}$, if those conditions, $\boldsymbol{A = \{A_m\}}$, are accomplished. Finally,
  (vii) the attribute *tExecuted* measures the times that the rule was triggered, and

(viii) the attribute *tFollowed* gathers the times in which this rule recommendation was chosen as the final decision from all the executed times.

(ix) Computing both variables we could get the *usability* of the rule as tFollowed/tExecuted.

(x) The *strength* of the rule is computed based on the outcomes of the patient saved within the *Outcomes*, ***O(t)***, class and referred by the *DecisionalEvent* class.

```
/**
 * @author nmuro
 */
public class Guideline {

    String id;
    String num;
    String version;
    String author;
    String date;
    ArrayList<RuleGroup> ruleGroup;
```

```
/**
 * @author nmuro
 */
public class Rule {

    String ruleID;
    String ruleName;
    String ruleAuthor;
    ArrayList<ContextEntity> ruleContext;
    ArrayList <ConditionTriplet> conditionsDescription;
    ArrayList <Recommendation> recommList;
    int usability;
    int strength;
    int tFollowed;
    int tExecuted;
```

```
/**
 * @author nmuro
 */
public class RuleGroup {

    String ruleGroupName;
    String ruleGroupId;
    ArrayList <Rule> rules;
```

Figure 7: The *Guideline*, *RuleGroup,* and *Rule* Java classes and their attributes declaration.

After describing how the information is retrieved and processed in the decision-making process, we describe in depth how the clinical knowledge is modeled in computerized format through the different components of the *Rule* class.

- *ContextEntity* class: as stated in the *Rule* class, the *ContextEntity* class will store all the semantic context information linked to the variables that compose the rule, meaning that if there is any semantic restriction defined by the clinical domain, it will be expressed here (see Figure 8). This is accomplished by defining 5 different attributes that describe:
  (i) the kind of class to which the clinical variables of the conditional part of the rule are referring to, in accordance with the clinical knowledge model where the domain knowledge is defined (i.e. an ontology that defines the clinical knowledge of a given domain and reflects all the relationships and restrictions among variables from the clinical point of view),
  (ii) the parent class type, following a classification scheme in the model that groups different variables by their common origin (e.g. the variable can be related to the patient, related to other clinical variables, related to a

lesion or related to a body side and hence classified under these classes, inheriting their definitions) affecting how these variables should be processed,

(iii) the variable name itself and if pointing out an attribute,

(iv) the attribute description and

(v) other referring variables that could be affected by the value or definition of the current variable.

This class is very important for maintaining a semantic rigor and for identifying hierarchical or same-level relationships and restrictions among variables and univocal identification according to the formalized clinical knowledge in the model.

```
/**
 * @author nmuro
 */
public class ContextEntity {

    String kind;
    String type;
    String var;
    String att;
    String ref;
```

Figure 8: The *ContextEntity* Java class and its attributes declaration.

- *ConditionTriplet* class: this class allows for the relation of different conditions in a hierarchical way, following the "Task-Network Models" (TNM) formalization. It consists of two *Condition* instances (i.e. precedent and consequent conditions), and a binary or logical operator, which defines a logic among conditions and the kind of restriction they must check to move forward or stop the evaluation from being performed (i.e. the binary operator defines whether the accomplishment of the condition is mandatory using the operator "AND" or if the conditions are comparable and consequently only one could be accomplished in order to continue with the evaluation process by using the operator "OR") (see Figure 9).

```java
/**
 * @author nmuro
 */
public class ConditionTriplet {

    Condition precedentCondition;
    Condition consequentCondition;
    String binaryOperator;
```

Figure 9: The *ConditionTriplet* Java class and its attributes declaration.

- *Condition* class: this class relies first on the *ContextEntity* class since they must be in agreement in order to assure the correct definition of the conditions along with the clinical domain model and the clinical restrictions coming from the CPGs (see Figure 10). Hence, *modelReferenceAttribute* and *modelReferenceVariable*, that identify the clinical variable to be studied in the condition, must be defined in at least one of the variable definitions created within the context headings of the rule (e.g. for the *modelReferenceAttribute* "Age", the defined *modelReferenceVariable* would be "thePatient", being "thePatient" a particular instance of "Patient" class defined in the ontology, which encompasses "Age" attribute). The *variableType* attribute will identify the typology of the variable that is being studied (e.g. integer, for "Age" variable, since it will be defined by natural numbers) and that must also match with the *type* attribute defined in the rule context when describing that variable. Finally, the *operator* attribute refers to the mathematical operator to be applied when studying the incoming clinical variable, and the *variableValue* will define the threshold or value to accomplish.

```java
/**
 * @author nmuro
 */
public class Condition {

    String operator;
    String modelReferenceAttribute;
    String modelReferenceVariable;
    String variableType;
    ArrayList<String> variableValue;
```

Figure 10: The *Condition* Java class and its attributes declaration.

- *Recommendation* class: This class defines the consequent part of the rule. It is composed of an ID that identifies which rule this consequent part is related to (i.e. *sourceRuleID*), and the order(s) that compose this recommendation (i.e.

orderSet), since a single recommendation can be composed by one or more orders or procedures to be performed (see Figure 11).

```
/**
 * @author nmuro
 */
public class Recommendation {

    String sourceRuleID;
    ArrayList<Order> orderSet;
```

Figure 11: The *Recommendation* Java class and its attributes declaration.

- *Order* class: represents a single procedure or order within a recommendation, which is composed of a number of attributes (see Figure 12). The *stepNumber* attribute represents the numerical order for the current procedure to be performed within a recommendation. Moreover, since some of the rules are not formalized to provide a treatment recommendation but a variable definition, the *assignations* attribute was detailed. This attribute is based on the condition object explained before, since it designates the requirement of a new condition accomplishment for this patient for giving a therapy. For example, if the variables "*Estrogen Receptor"* and "*Progesterone Receptor"* of a patient are negative, an assignation type order which assigns to the variable "*Positive Hormonal Receptors"* the value "*false*" will be defined within the provided recommendation. In this case, the *Order* class contains some domain-dependent variables (e.g. *breastDose*, *boostDose*), but it comprises the most relevant variables for describing any kind of treatment values, hence not limiting the recommendation definition to the selected clinical domain of study within this thesis (i.e. breast cancer).

```java
/**
 * @author nmuro
 */
public class Order {

    String breastdose;
    String boostdose;
    String fractions;
    String fractionsPeriodicity;
    String duration;
    String durationUnit;
    String cycles;
    String conformance;
    String value;
    String stepNumber;
    ArrayList<Condition> assignations;
```

Figure 12: The *Order* java class and its attributes declaration. The first two attribute classes are defined for the specific clinical domain of Breast Cancer.

In conclusion, the Decisional Event object is flexible enough to handle information related to any clinical decision on any clinical domain, since the domain knowledge and the imposed semantic interoperability for the studied clinical setup is defined externally, leaning on guidelines and semantic web technologies. As specified in the classes' definition, the Order class definition is to some extent domain-dependent to provide as much detail as possible regarding the studied clinical case. Nevertheless, the way it was defined allows the description of any therapies in a higher level of specifications definition, not requiring further developments.

Once the decisions are computerized through the Decisional Event structure all information is stored in a database to be able to (i) save all the clinical information in an easily updatable and exploitable platform, (ii) link the decisions taken for a patient with the outcomes and clinical variables over time to evaluate the clinical performance individually and at population level, (iii) calculate and store compliance rates and track non-compliant decisions and their causes, generating new knowledge that will update or increase the already accumulated knowledge and (iv) compute usability and strength of the rules based on the different clinical outcomes tracked over time.

3.1.1.3 Decisional Event relational database design and implementation

The last step to retrieve, save, and generate new knowledge from all the information related to the decision-making process implies the design of a relational database for persistent storage of this information. Taking into account the design requirements of the Decisional Event

object, a relational database composed of 6 principal tables and 9 relational tables was made (see Table 1).

Principal tables names	Relational tables names
decisionalEvent	*desicionalevent_has_rule*
guidelineInformation	*decisionalevent_has_guideline*
rule	*decisionalevent_has_finaldecision*
conditionTriplets	*rule_has_context*
ruleCondition	*decisionalevent_has_executedtreatment*
ruleRecommendation	*decisionalevent_has_recommendation*
	decisionalevent_has_criteria
	guideline_has_rulegroup
	recommendation_has_orders

Table 1: Description of the tables composing the relational database for storing all the information related to the decision-making process.

Figure 13, shows the corresponding tables with the variables they encompass, their matching variables' types, and their different relationships (i.e. 1:n, n:n).

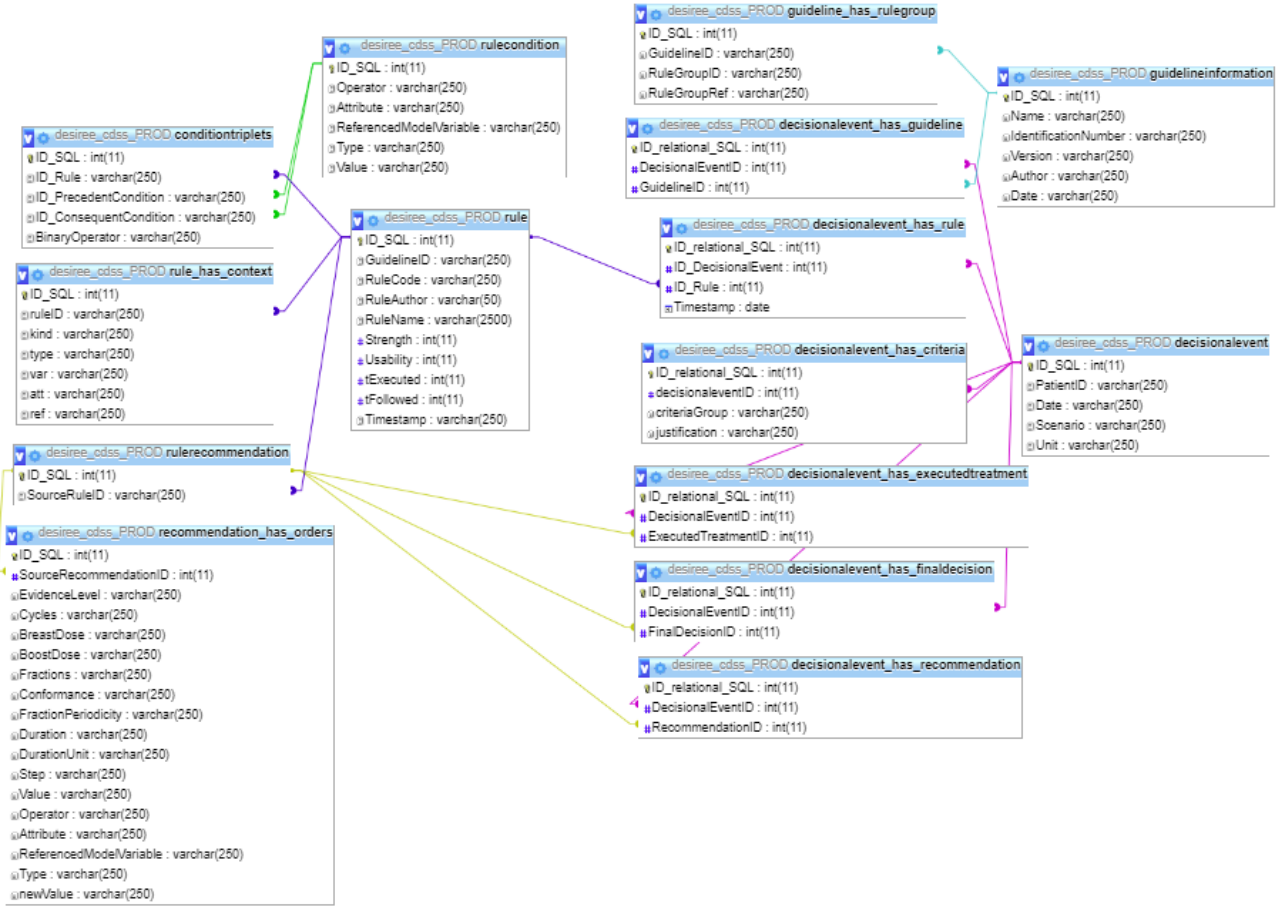


Figure 13. The relational view of the database containing all the information related to the decision-making process.

Even if it is not directly modeled in the Decisional Event database, retrieving the clinical information to be analyzed and processed is mandatory. As stated in the Decisional Event formalization, the clinical parameters are also part of the decision, since they describe the clinical case to be studied. To link patients' clinical data with the related Decisional Event, different relation identifiers were defined. Therefore, even if the information about clinical performance and evidence was saved in one database and the clinical data and outcomes in another, we are able to relate them in a consistent way. This provides additional domain independence.

The methodology to transfer or retrieve the clinical data of a patient was done following the requirements of the Fast Healthcare Interoperability Resources[14] (FHIR) standard, created by the Health Level Seven[15] (HL7) international health-care standards organization. In Figure

[14] https://www.hl7.org/fhir/

[15] http://www.hl7.org/

14, the relational database created to store all clinical data following the FHIR standard is shown.

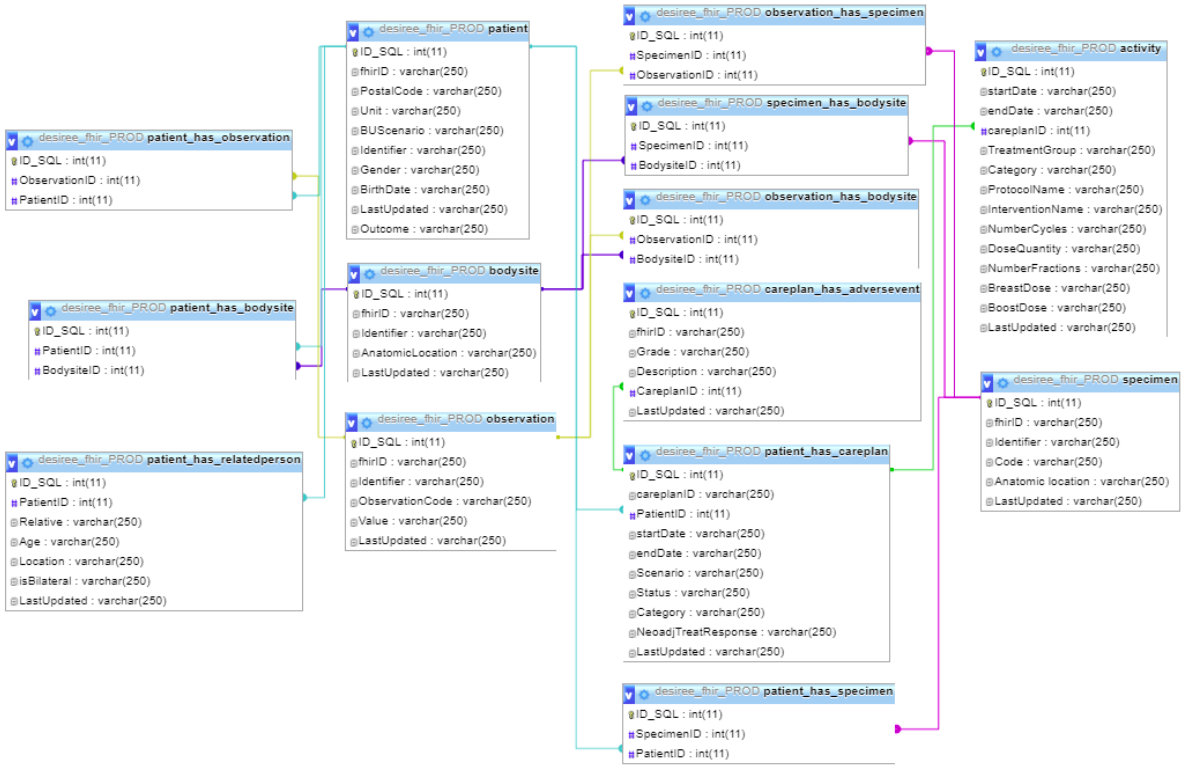


Figure 14: The relational view of the database containing all the clinical data related to the patient following the FHIR messaging definition and structuration.

The FHIR messaging protocol defines some resources and data types with several properties that facilitate the transmission of clinical data in a standardized way. In this case, the FHIR resources *Observation*, *BodySite, Patient*, *Specimen, Careplan, RelatedPerson,* and *Activity* were used. Some extensions were added to those resources to deal with some of the particular needs of the clinical domain, for example, when a specification of treatment attributes is needed, *Careplan* and *Activity* resources may be extended to reach the required level of specificity in the exchanged clinical information.

3.1.2 CIG Formalization Module

As introduced in chapter 2, the formalization of CPGs in a computer interpretable way is required for implementing CPGs within a CDSS. CPGs gather the latest clinical evidence statements or criteria for a given clinical domain, describing the procedures to be followed. In our approach, CPGs are computerized following a rule-based approach that leans on the Decisional Event structure and which replicates the “Task-Network Models” (TNM) (Peleg

2013). Since these models use rules to represent the dependency among actions structured as hierarchical networks, when the conditions are fulfilled in a satisfactory way, the associated recommendation part is provided as a conclusion. To keep the hierarchical relationship among conditions, the Condition Triplet class was used (see Figure 15).

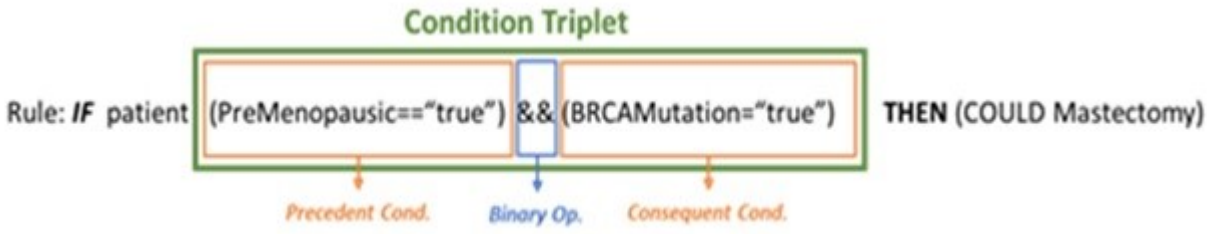


Figure 15. A simple example of a complete Condition Triplet instance, composed of a precedent condition, a binary operator and a consequent condition, following the proposed CIG formalization.

There are three ways of completing a condition triplet in the proposed model:

(i) only the precedent condition is fulfilled, meaning that the rule is composed of only one condition,

(ii) the precedent condition and the binary operator are fulfilled, and the consequent condition remains empty. This means that the conditional part of the rule is composed by n conditions, which will be defined in n-1 condition triplets when $n \geq 3$, where the precedent condition of the n condition triplet would act as the consequent condition of the n-1 condition triplet, linking the first condition triplet to the next one through the logical or binary operator, until the terminal condition triplet, which will be completely filled (see Figure 15), and

(iii) when the three components of the condition triplet are fulfilled (i.e. precedent condition, binary operator, consequent condition) meaning that either it is a rule with two conditions or it is the last condition triplet of the rule, for rules with more than two conditions (see Figure 15).

Once all conditional triplets are completed, the conditional part of the rule is considered as defined. For the consequent part or the recommendation, a rule can have one or more recommendations, each being composed of one or more orders or actions to be executed. All this composes the *Rule* object explained in Chapter 3.1.1.

Once the knowledge is mapped into the *Rule* object, a translation to any computer-executable language could be done. In this work, we created two translation modules, one following the HL7 Standard for the Clinical Decision Support Knowledge Artifact Specification[16] into an XML-based document and another one following the Drools Rule Language[17] (DRL) to generate a DRL extension file that can be triggered by the Drools rule engine. The first format is mainly focused on transferring CIGs in a standardized manner. The second format was used for automatizing the translation of the CPGs into CIGs, which will be plugged into a guideline-based CDSS based on Drools.

3.1.3 Ontology-based semantic validation

The last proposed step regarding the knowledge formalization aims to assure the semantic interoperability of all the clinical concepts and relations defined within the CIG based on a domain ontology and avoiding including corrupted knowledge. To accomplish this objective, all of the clinical variables and their identified values are mapped into an ontology, which semantically reflects all relationships among variables in a (clinical) domain using different object-type and data-type properties (Bouaud, Guézennec, and Séroussi 2018).

As the Decisional Event formalization, and consequently the rules definition, completely relies on the clinical information and relationships defined in the ontology, the proposed formalization is domain-independent and can be used to query any ontology and to search the possible values of any class within it. Moreover, this approach facilitates the usage of standard terminologies and codifications when defining each of the clinical variables within the knowledge model or ontology, such as the National Cancer Institute (NCI) Thesaurus[18], SNOMED CT[19] or Unified Medical Language System (UMLS)[20]. To prove it, a Jena API based tool was developed, able to query any ontology formalized in the Resource Description Framework (RDF) language to be used along with the Java-based CIG formalization module (Artola et al. 2019)[21]. This tool interacts with the methodology presented in Chapter 3.2.

[16] http://www.hl7.org/implement/standards/product_brief.cfm?product_id=337

[17] https://docs.jboss.org/drools/release/5.2.0.Final/drools-expert-docs/html/ch05.html

[18] https://ncit.nci.nih.gov/ncitbrowser/

[19] http://www.snomed.org/snomed-ct/get-snomed

[20] https://uts.nlm.nih.gov/home.html

[21] Paper accesible at : https://www.insticc.org/Primoris/Resources/PaperPdf.ashx?idPaper=80680

3.2 Methodology for an evolutive CDSS

Guideline-based CDSS relies on the knowledge reported within the CPGs to provide the best patient-specific and evidence-based recommendations. These systems are developed in order to promote the compliance of the standardized CPGs aiming to reduce the variability of the clinical practice and improve patients' clinical outcomes. To achieve this objective, CPGs, which are paper-based narrative documents, must be formalized as CIGs.

As stated in Chapter 2.1, a number of proposals have been made for CPG formalization as CIGs. In the proposed approach, the clinical knowledge in CPGs was formally represented as "Task-Network Models" (TNMs) since this formalism allows the representation of the clinical knowledge as Business Rules (i.e. IF-THEN rules), showing the dependencies among the different clinical conditions which, when fulfilled in a satisfactory way, provide some therapeutic recommendations. This CIG approach is implemented in a Java-based formalization of the rules (i.e. *Rule* object presented in Chapter 3.1.1), allowing computer-based reasoning and execution of these formalized guidelines through a rule-execution engine for a given case study.

In the following, the issues faced when formalizing and updating CIGs within a CDSS will be addressed, providing a domain-independent methodology that facilitates CIG maintenance and update in a computer interpretable format. Moreover, an authoring tool that supports and eases the reviewing and validation process of the formalized CIGs is presented. Finally, an experience-based CDSS is proposed which tracks all decisions made over time generating new experience-based rules to deal with the identified guideline limitations, enriching the knowledge base of the CDSS with experience and exploring some outcomes' analysis to assess the strength behind these experience-based recommendations.

3.2.1 Domain-independent CIG formalization

In this Chapter, a tool for formalizing any domain CIG, leaning on the previously formalized Java *Rule* object within the Decisional Event (see Chapter 3.1.1) is proposed. Since the main objective is to generate automatically CIGs with no clinical domain restrictions, a

modular approach based on the Drools[22] business rule management system was done, which encompasses two main components:

(i) an automatic and domain-independent CIG rule file generator in Drools Rule Language (.drl extension) and

(ii) a regular standalone runtime execution environment or rule engine that will communicate with the first module for triggering the dynamically updated CIGs.

The first module presents a generic approach to ease the CPG formalization into a CIG (e.g. DRL file) in a very flexible and domain-independent way, using Drools Rule Language or DRL-based rule templates like the one shown in Figure 16.

```
dialect "java"

template "RuleTemplate"

rule "@{row.rowNumber} - @{name}"
when
    @{object}(@{conditional} )
then

    recom.add("@{action}");
    debug(drools);

end


end template
```

Figure 16: Screenshot of the DRL language-based rule template, used for CIG formalization.

DRL rule templates are Java-based documents that, along with tabular data sources, are able to formalize rules to be triggered by a Java-based rule engine. The main advantage of these templates is that since they are defined following a generic and domain-independent approach, they can support the formalization of any CPGs into CIGs. The DRL rule template is composed of different markers, which identify different sections or requirements needed for rule generation (see Figure 16). The *"dialect"* marker defines the software language used to communicate with the rule engine (i.e. the Drools-based standalone runtime execution environment), meaning that functions written in the specified language (e.g. Java) could interact with the DRL file to include or compute values within it. The *"template"* designates a name to the DRL file that will be used afterward to call and trigger this file within the running environment. The following four markers (i.e. *rule*, *when*, *then*, *end*) define each of the rules.

[22] https://www.drools.org/

First, the *"rule"* marker identifies each rule and gives a description or a rule name, which can establish some information about the formalized knowledge within it. The *"when"* marker formalizes the conditional part of the rule (i.e. the conditions to be accomplished by the evaluated patient), whereas the *"then"* marker formalizes the consequent part (i.e. the recommendations that will be provided when the conditional part was successfully achieved).

Since the rule template can communicate with Java-based classes and methods, we will lean on the previously defined *Rule* object (see Chapter 3.1.1) where all rules are defined in a structured and processable way in Java. First, the knowledge to be formalized is gathered from an external source (e.g. the rules stored in a database) and mapped into a *Rule* object (see Chapter 3.1.1). Afterward, the fulfilled *Rule* is mapped into a Java Map object (i.e. Java-based object that defines key-value pair entries), which simplifies its complexity into four key attributes and their values:

(i) *name*, which defines the name of the rule,
(ii) *object*, describing the variables that must be defined by an input patient to be analyzed,
(iii) *condition* or the conditional part of the rule to be accomplished by the patient clinical data, and
(iv) *action* giving the recommendations defined by the rule.

Once the rule is transferred from the *Rule* object to the Java Map object, it will be outlined to the rule template in the corresponding tags adding an "*end*" marker to identify the end on the rule. When all rules are transferred into the rule template, an *"end template"* marker will be added, referring to the end of the CIG document. This output document will be ready for being called by a standalone runtime execution environment or rule engine that will trigger the formalized rules for any incoming patient information whose data-model matches them.

In conclusion, the proposed approach overcomes the limitations of available tools to formalize business rules, such as Drools spreadsheet tables, since these methods are structurally limited to only allow the generation of rules that contain the same variables definition varying the evaluated variable values. This limits the richness of clinical rules to a rigid domain-dependent structure. In this chapter different Java-mapping tools have been used along with the *Rule* object defined in Chapter 3.1.1 to provide a flexible and domain-independent CIG

formalization module since the clinical knowledge is mapped from external sources instead of being hard-coded within the CIG.

Nevertheless, formalizing the clinical knowledge in any external source such as a database or an ontology requires some programming skills and clinical domain understanding to connect these sources to the presented back-end developments. To ease this step, an authoring tool to define clinical knowledge through a Graphical User Interface (GUI) in a user-friendly way is presented next.

3.2.2 Authoring tool

Formalizing CPGs is not an easy task. First, it requires some clinical knowledge understanding from the technical side, since paper-based documents containing the latest EBM must maintain fidelity to the original CPGs when transcribed into computer interpretable documents. Moreover, we must assure that the richness of the medical narrative is not lost when formalizing this knowledge into CIGs. Once this formalization is successfully accomplished, we may face a third issue: the maintenance of those computer interpretable documents. Translating CPGs into CIGs is mainly promoted not only to facilitate the use of CDSS during the medical practice but also to ease the maintenance, update, and extension of the knowledge that they formalize. Since EBM is periodically updated based on new study results and continuous research, CPGs also need to be regularly revised to include the most accurate and highest quality healthcare procedures. This tedious process implies generating new versions of the CPGs, which must reflect these amendments. Having all guidelines in a computerized manner and providing tools that facilitate their update and maintenance is therefore of utmost importance.

In this section, an authoring tool Graphical User Interface (GUI) that extends the domain-independent CIG formalization module is presented (see Figure 17). This GUI (i) allows the formalization of CPGs into CIGs through a user-friendly interface that does not require any programming skills, (ii) is connected to the ontology-based semantic validation scheme presented in Chapter 3.1.3 and the rule engine described in Chapter 3.2.1, allowing the formalization of semantically validated rules to be included into CIGs that will be available to be queried forthwith and (iii) eases the update and maintenance of the CIGs, supporting the fine-tuning of previously loaded rules or extending the actual knowledge base with new ones.

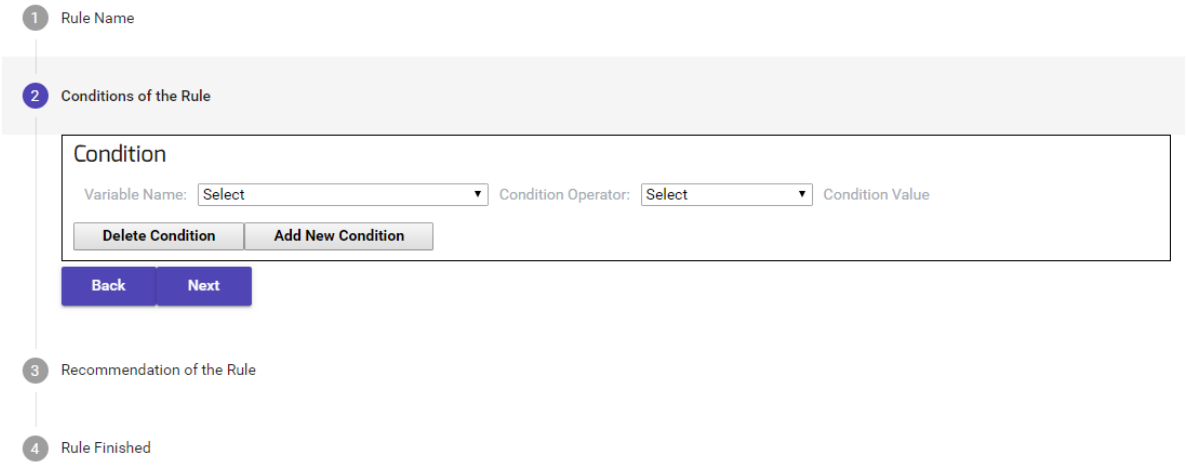


Figure 17: The authoring tool GUI for the definition of new rules.

The presented interface is composed of different combo boxes for individual item selection, which are aligned with the components of the rules defined within Chapter 3.1.1. The application workflow follows the CIG formalization procedure explained in Chapter 3.1.2 to build the Java-based *Rule* objects that will be transferred to the domain-independent CIG formalization module introduced in Chapter 3.2.1, creating the CIG document. Moreover, this interface also integrates the module explained in Chapter 3.1.3 which validates semantically all the formalized clinical concepts and their relationships against an external ontology or knowledge model that structures all relevant variables, the possible values that those variables can take and the relationships among all of them for the studied clinical domain. Integrating this ontology-based semantic validation module assures that the clinical knowledge that is being formalized through the authoring tool GUI complies with the clinical restrictions defined in the knowledge model (e.g. the clinical variable *"Age"* can only accept natural numbers as possible values). Hence, when displaying the *"Variable Name"* combo box in the GUI, an ontology with the clinical domain knowledge is queried, returning all of the variables that may be used to define a condition within the conditional part of a rule (see Figure 18).

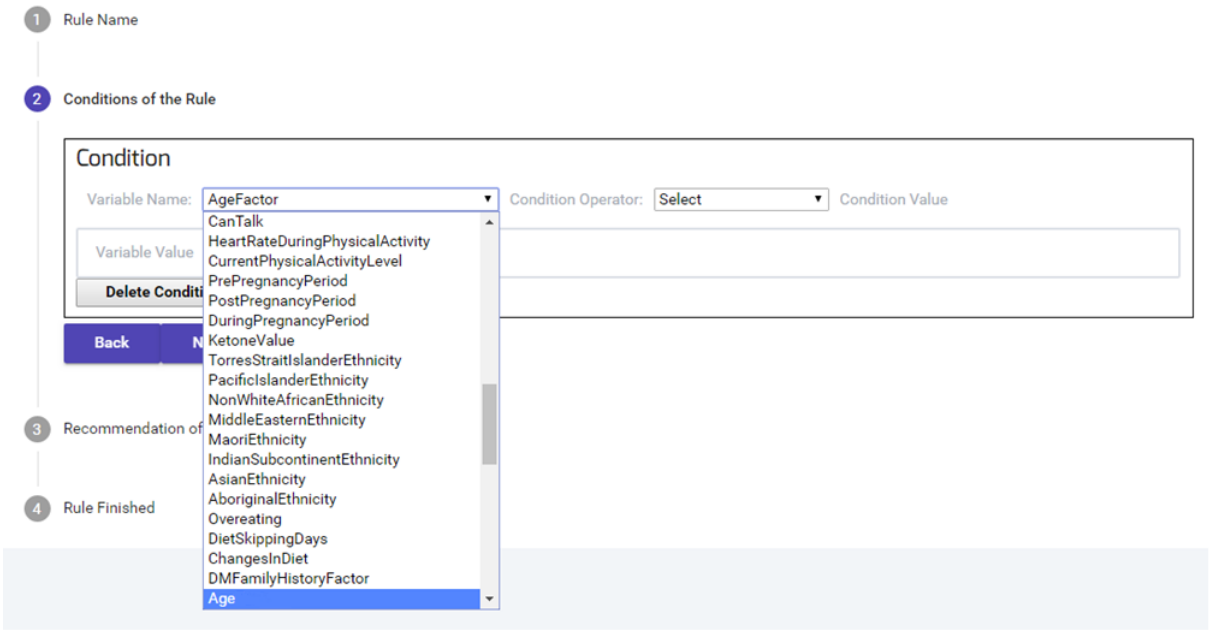


Figure 18: Example of the possible values that a variable within a condition of a rule could take according to the knowledge in gestational diabetes mellitus domain modeled in an ontology.

Once a variable from the provided results set is selected, the rest of the combo boxes will update their content in accordance. This means that if in the first combo box where the variable name should be selected the clinician clicks on *"Age"*, the condition value will be restricted to the possible values defined for that variable in the ontology (i.e. natural numbers) when asking to define the condition threshold (see Figure 18Figure 19).

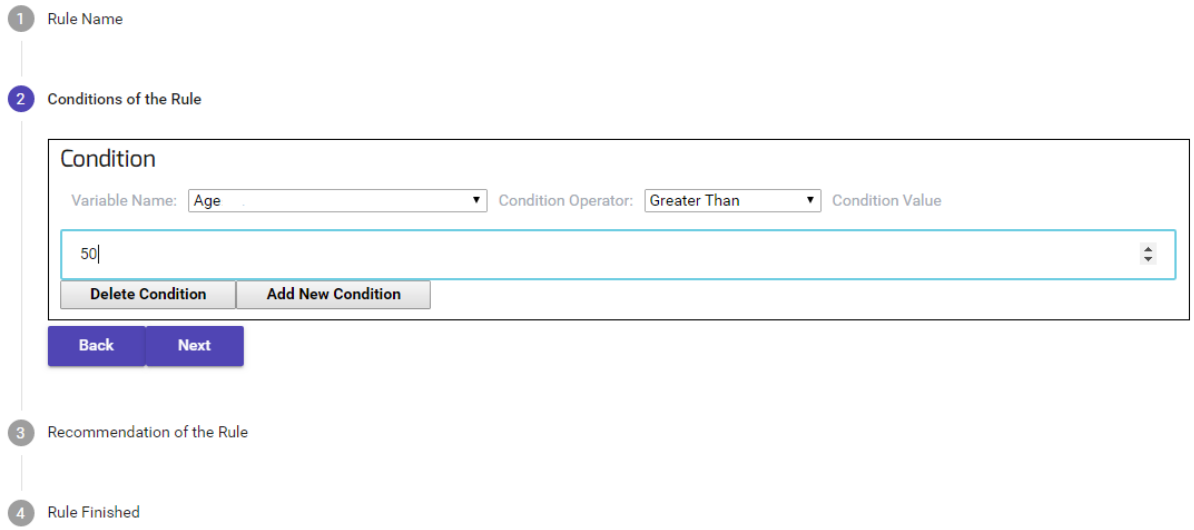


Figure 19: Example of the definition of a condition within a rule, which states that a patient's age must be greater than 50 years old.

Moreover, the condition operator will also be filtered depending on the studied variables, since categorical and numerical values cannot be evaluated in the same way. Once a condition is fulfilled, a binary operator can be selected (i.e. AND, OR) for including more conditions that would compose the conditional part of the rule or go for the definition of the recommendation part of the rule. Nevertheless, when the NONE operator is selected, internally it means that the rule triplet has no consequent condition, hence having only the precedent condition fulfilled (see Figure 20).

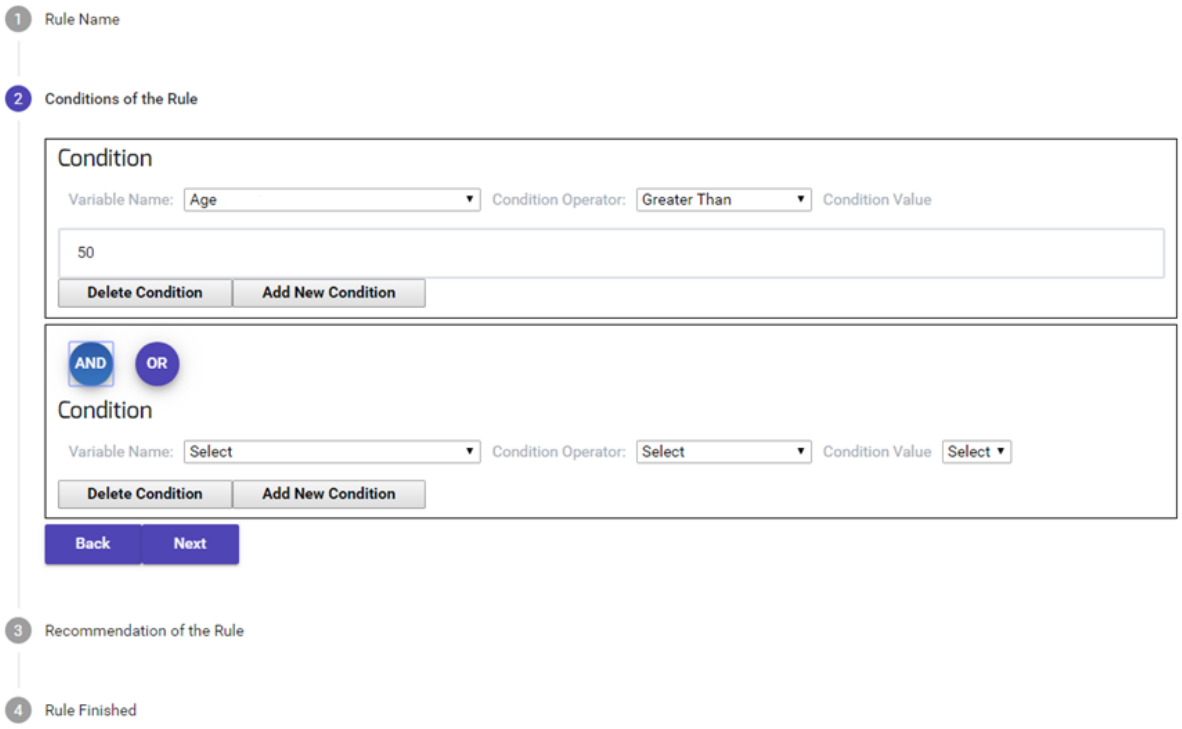


Figure 20: Example of a second condition definition linked by the "AND" binary operator.

This expresses that all conditions defining the conditional part of the rule were already defined (see Chapter 3.1.2). After the conditional part has been defined, the user is guided through the GUI to define the recommendation part of the rule. Following the *recommendation* object definition given in Chapter 3.1.2, a rule may have one or more recommendations and each of them can be defined by one or more orders, understanding order as any atomic action or treatment to be administered to the patient that altogether will compose a recommendation. As the definition of the orders may vary from one clinical domain to another, and taking into account that the domain of the discourse, in this case, is cancer and particularly primary breast cancer, some attributes were defined to deal with this specific illness requirements, such as *"Cycles"* and "*Cycles Duration"* when defining chemotherapy protocols or *"Fractions"* and

"Fractions Periodicity" when describing a radiotherapy treatment. Nevertheless, other attributes have a generic definition, allowing the profiling of any kind of clinical recommendation, such as *"Conformance"* that defines to which extent the recommendation has to be met or followed by the clinician (e.g. may, might should, must, should not, must not), "*Step number*" to enumerate each order following the correct arrangement in which they must be applied or complied, and finally the "*Treatment name*" specifying the exact therapy to be administered.

In conclusion, the presented authoring tool GUI integrates several modules presented in previous chapters in order to facilitate the formalization, maintenance, and update of CIGs for any clinical domain in a user-friendly way and not requiring programming skills.

3.2.3 Augmenting the clinical knowledge using experience

As mentioned before, guideline-based CDSSs are sometimes insufficient for providing the best recommendations for particular clinical cases. This could be the case for several reasons, such as (i) outdated CPGs implementation, to which we proposed a modular architecture that could help in overcoming that issue, (ii) particular clinical cases that do not fit with the guidelines due to their complexity or the lack of supporting evidence formalization in the CPGs or (iii) the inclusion of other factors that are closely related to the decision-making process and can alter the compliance with the guideline proposed recommendations, such as clinical preferences or patient preferences. To deal with those gray areas of CPGs, an experience-based CDSS is proposed in this chapter, whose main objectives are (i) the formalization of the clinical experience or implicit knowledge that guides clinicians when they do not comply with the guidelines, using it to generate new experience-based rules that will cover those guideline gaps and would support future similar clinical cases, and (ii) structuring and gathering all the information related to the decision-making process, allowing the measurement and evaluation of the provided healthcare.

3.2.3.1 Experience-based rules generation

Experience is defined as any practice or action applied by an observational intuition to deal with a situation that has no previous evidence to rely on (Bate et al. 2012). This module is intended to be a tool to acquire all the implicit knowledge used by clinicians when guideline-

based CDSS does not provide the best recommendations for a patient under study, allowing the generation of new experience-based rules that will cover those knowledge gaps.

In Chapter 4, a use case for the Breast Cancer domain is presented. To facilitate the comprehension of the approach proposed in this Chapter, we will support it through two Breast Cancer patients' definition examples (see Table 2). These example patients (i.e. Patient 1 and Patient 2) represent two cases diagnosed as Invasive Ductal Carcinoma, characterized by the following clinical variables:

		PATIENT 1	PATIENT 2
P_1	Age	65	73
P_2	Sex	Woman	Woman
P_3	Number of Pregnancies	0	4
P_4	Number of Lesions	2	2
P_5	Location	Right, Lower outer quadrant	Right, Upper outer quadrant
P_6	Size (mm)	21	21
P_7	BIRADS	5	-
P_8	Ulceration	NO	NO
P_9	Skin metastasis	NO	NO
P_{10}	cT (size)	T2	T2
P_{11}	cN (number)	0	0
P_{12}	cM (metastasis)	0	0
P_{13}	Stage	2a	2a
P_{14}	Histological type	Ductal invasive carcinoma	Ductal invasive carcinoma
P_{15}	Grade	GII	GI
P_{16}	Carcinoma in situ type	Ductal carcinoma	Ductal carcinoma
P_{17}	Estrogen receptor	Positive	Positive
P_{18}	Progesterone receptor	Positive	Positive
P_{19}	HER-2 receptor	Negative	Negative
P_{20}	Ki67 (%)	14	16
P_{21}	Clinical criterion: Tumor Size causes follow-up difficulty	-	True

Table 2: Set of clinical parameters values defining the Patient 1 and Patient 2.

When a cancer patient must be discussed among the involved multidisciplinary clinical team, a meeting is organized to discuss the case and decide which treatment should be administered. During these meetings, the clinicians would first query the guideline-based CDSS to analyze a given clinical case defined by a set of variables to obtain the evidence-based recommendations that best fit the clinical case. The expected output would be a set of recommendations that could be delivered to the patient aiming to improve its clinical condition. Figure 21 shows a rule definition from a local clinical protocol of the Onkologikoa Foundation Hospital[23] for Primary Breast Cancer (PBC) management. This rule would be triggered for both, Patient 1 or 2, which advises providing a neo-adjuvant hormonotherapy, since both patients have a lesion bigger than 20mm (i.e. P_6 in Table 2) with a histological type equal to ductal invasive carcinoma (i.e. P_{14}), with positive estrogen (i.e. P_{17}) and progesterone (i.e. P_{18}) receptors but negative HER2 receptor (i.e. P_{19}) and a Ki-67 percentage lower than 20 (i.e. P_{20}).

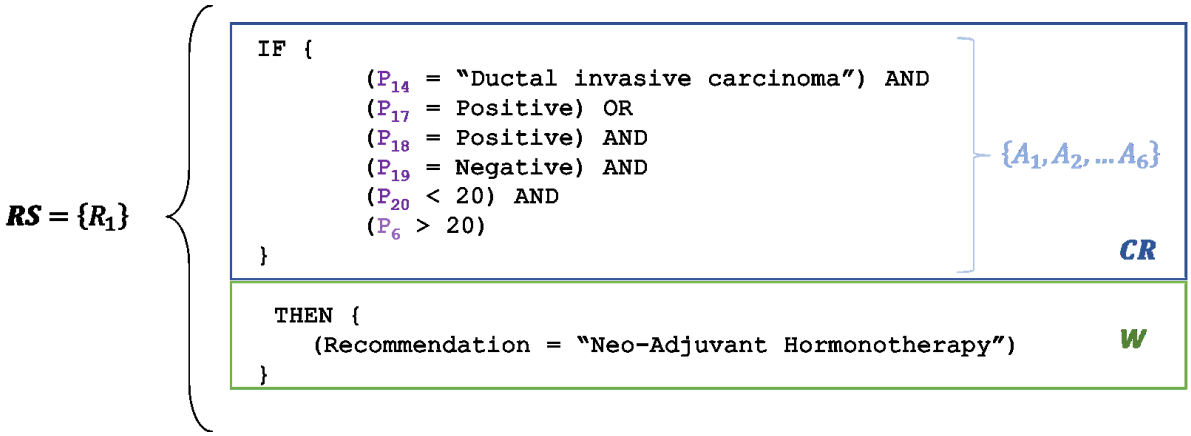


Figure 21: Example of a simplified rule for non-metastatic breast cancer with infiltrating tumor management.

During this decision-making process, the clinicians may or may not be compliant with the guidelines based on (i) the studied clinical case complexity, (ii) the adequacy of the suggested recommendation set from the queried guidelines or (iii) other factors or criteria that may affect the choice of the final decision. If the clinicians stay compliant with the guidelines, selecting one of the guideline-based recommendations as the final decision, a new Decisional Event will be recorded saving all the information related to that decision (see Chapter 3.1.1.1, Decisional Event Structure). On the contrary, if the multidisciplinary clinical team decides to be non-compliant, a justification for this non-compliance will be requested (see Figure 22). The aim of retrieving these justification(s) or non-compliance criteria is to provide some insights or

[23] https://www.onkologikoa.org/en/

information that could specify what the reasoning was behind providing an alternative treatment to the patient that does not rely on CPGs evidence.

Decisional Event (PATIENT 1)

- Clinical Parameters: $\{\boldsymbol{P_i}\}\ with\ 1 \leq i \geq 20$
- Clinical Condition: $\boldsymbol{R}$
  - Set of Antecedents: $\{\boldsymbol{A_m}\}\ with\ 1 \leq m \geq 6$
  - Recommendation: $\boldsymbol{W} =$ "Neo − Adjuvant Hormonotherapy"
- Clinical criterion: $\boldsymbol{C} =$ Tumor Size − clinical variable restriction
  - Justification: $\boldsymbol{J_1}$ (Follow-up difficulty) $= True$
- Final decisión: $\boldsymbol{FD} =$ "Surgery"

Figure 22: Summary of the representation of the data that composes a Decisional Event for Patient 1 in a time t=0. The non-compliance related variables are highlighted in red.

Once these criteria are provided, the system will store the non-compliant decision along with the information retrieved during the decision-making process. Afterward, the experience-based rule using this information will be generated as follows:

- The starting point of the method is to retrieve the set of CPG rules that were queried and triggered during the decision-making process and that will be part of the generated Decisional Event $\boldsymbol{RS} = \{R_1, R_2, \dots R_j\}\ with\ j > 0$. The antecedents (i.e. IF-part) of these rules are defined as $\boldsymbol{CR} = \{A_1, A_2, \dots A_k\}\ with\ k > 0$ and will be included as part of the experience-based rule. The evaluated patient parameters accomplished all of them (see Figure 21Figure 21).
- Thereafter, we identify and retrieve another rule set within the queried CPGs $\boldsymbol{RS'} = \{R'_1, R'_2, \dots R'_u\}\ with\ u > 0$ whose recommendation $\boldsymbol{W}$ match the final decision $\boldsymbol{FD}$ made by clinicians for the analyzed scenario (surgery, for the presented example). The antecedents of this secondary rule set $\boldsymbol{RS'}$ are defined as $\boldsymbol{CR'} = \{A'_1, A'_2, \dots A'_m\}\ with\ m > 0$ and the evaluated patient parameters did not accomplish at least one from each rule $\boldsymbol{R}$' (see Figure 23).

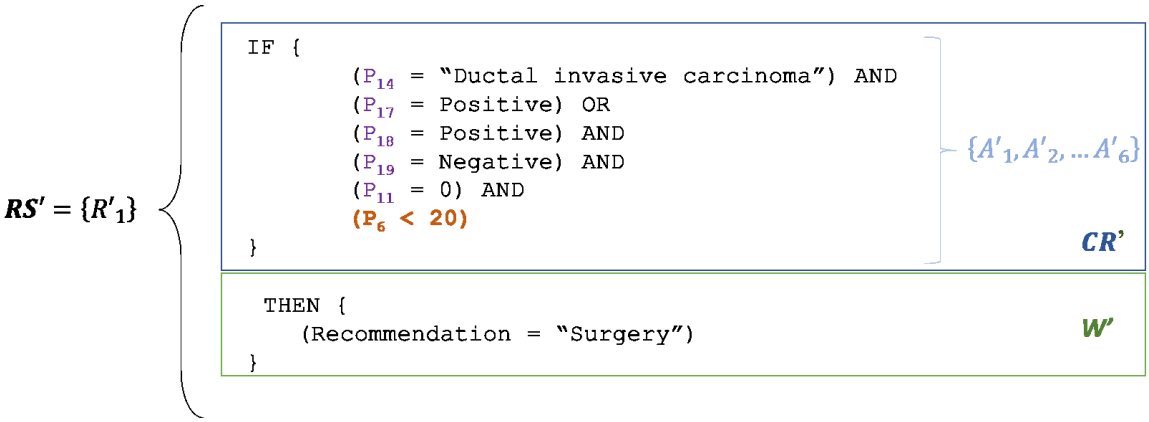


Figure 23: Example of a simplified rule for non-metastatic breast cancer with infiltrating tumor management, matching the Final Decision made by the BU for Patient 1. The conflictive antecedent is highlighted in orange since Tumor Size >20 ϵ CR and Tumor Size ≤20 ϵ CR'.

- From both sets of antecedents, ***CR*** and ***CR'***, *"conflictive antecedents"*, i.e. incompatible antecedents, and antecedents that are complementary, i.e. "*complementary antecedents"* are identified.
    - Complementary antecedents are kept within the conditional part definition of the experience-based rule generated from the non-compliant decision.
        - In some cases, one or more antecedents could be defined in the non-compliant antecedent set ***CR'*** but not in the compliant one ***CR,*** i.e. antecedents defined in the relative complement of ***CR*** formally noted as $\boldsymbol{CR' \setminus CR}$. For this scenario, this new antecedent will be included in the new rule with the patient's clinical parameter as constraint value.
    - When the identified ***CR*** and ***CR'*** sets contain conflictive antecedents, the following steps must be adopted, depending on the relevance of the antecedent (i.e. it was part of the non-compliance criteria) and the reasons for the conflict:
        - If the antecedent is defined in both ***CR*** and ***CR'*** (i.e. when it is defined in $\boldsymbol{CR \cap CR'}$) but with different value constraints, the experience-based rule will include this antecedent with the patient's clinical parameter value as a constraint. In our example, the tumor size in ***CR*** is characterized by *Tumor Size >20* whereas for ***CR'*** is measured as *Tumor Size ≤20*. In the experience-based

rule, it will take the patient value: *Tumor Size* = P_i = 21mm (see blue condition definition in Figure 24).

- Lastly, the set of criteria $\boldsymbol{C_k}$ (e.g. clinical preferences, patient preferences) defined by clinicians in the decision-making process composed by one or more Boolean justifications $\boldsymbol{J_n}$ give us hints about new relevant clinical parameters to include (e.g. because they were not defined in the within the triggered rules of the queried CPGs) or study. They are added as new conditions to accomplish within the experience-based rule (see orange condition definition in Figure 24, where P_{21} represents the variable *Clinical criterion: Tumor Size causes follow-up difficulty*, as stated in Table 2).

Hence, in Figure 24 the new experience-based rule is shown, composed by:

(i) the antecedents checked by both rules, i.e. 'complementary antecedents' $\in \boldsymbol{CR} \cap \boldsymbol{CR'}(same\ P_i\ equal\ value)$ (represented in purple),

(ii) the adjustment of the parameter that was not compliant in one of them characterized by the most restrictive value, i.e. *'conflictive antecedents'* $\in \boldsymbol{CR} \cap \boldsymbol{CR'}\ (same\ P_i\ different\ value)$ (represented in blue),

(iii) the inclusion of a clinical parameter that was only measured in one of the rules, i.e. inclusion of $\boldsymbol{P_i} \in \boldsymbol{CR} \backslash \boldsymbol{CR'} \cup \boldsymbol{CR'} \backslash \boldsymbol{CR}$ (in green) and

(iv) the inclusion of clinical criteria that justifies the non-compliancy from the BU (represented in orange).

```
IF {
        (P14 = "Ductal invasive carcinoma") AND
        (P17 = Positive) OR
        (P18 = Positive) AND
        (P19 = Negative) AND
        (P20 < 20) AND
        (P6 = 21) AND
        (P11 = 0) AND
        (P21= True)
}
THEN {
        (Recommendation = "Surgery")
}
```

Figure 24: The new experience-based rule generated to cope with this CPG non-compliant clinical case

The recommendation(s) of this experience-based rule will be recorded as the final decision chosen by the clinicians during the non-compliant decision definition.

Finally, all generated experience-based rules will be added into the knowledge base, augmenting the CDSS knowledge base with clinical experience that was not identified or formalized before. Since these rules do not have the same evidence levels as the guideline-based ones, they are identified or tagged as experience-based rules when provided to clinicians along with the guideline-based recommendations when studying a new clinical case. In order to ensure the quality of these rules, we propose some usability and strength measurements based on the patients' outcomes.

3.2.3.2 Clinical Evidence Assessment

When augmenting the clinical knowledge base of the CDSS with new knowledge that was identified from particular clinical cases and setups, such as the experience-based rules explained in Chapter 3.2.3, we must ensure that clinical evidence is reliable and that this new knowledge could be applied to any incoming similar case that should match with the experience rules without causing any harm. CPGs measure the quality of evidence and the strength of recommendations based on different studies' results (e.g. randomized clinical trials, observational studies) and different reported outcomes, such as overall survival and disease-free survival rates and treatment responses, which assess their clinical reliability. Since experience-based rules can be considered a way of facilitating new evidence inclusion and updating the actual clinical knowledge, we studied the procedures followed in guideline updating and evidence evaluation based on outcomes and we focused on the measurement of three kinds of outcomes that give a spherical evaluation of the clinical evidence supporting the generated rules: (i) treatment response, (ii) reported toxicities and adverse events after being treated from a clinical point of view and (iii) the measurement of patient-reported outcomes (PROs) to include the patient perspective of the effectivity of the administered treatment. In this chapter, we will describe these outcomes in-depth, and we will quantify the usability of all the rules based on the decisions made over time. Moreover, a proposal of the GRADE scale-based approach for reporting the strength of the experience-based rules recommendations is given.

3.2.3.2.1 Outcomes for usability and strength computation

Outcomes can be used to assess and evaluate the clinical reliability of the knowledge formalized in the CPGs at a time *t*, supporting clinicians during the decision-making process for the most reliable and best healthcare performance assessment when administering a

treatment to the studied patient. In our approach, we propose dynamically storing outcomes in a computerized way to allow for the automatic evaluation and update of the evidence reported by the CPGs. We focused on three kinds of outcomes: (i) first given treatment response, (ii) toxicities or adverse events and (iii) patient-reported outcomes.

3.2.3.2.1.1 Treatment response

Measuring and studying the stage of illness after providing treatment to a patient can give some hints about the effectivity of a therapy. The treatment response refers to some generic outcomes, such as, survival, local relapse, loco-regional relapse, metastasis, exitus related to the disease and exitus not related to the disease, which are used to measure the improvement, stagnation or deterioration of a patient's condition in a clinically objective way after providing an initial treatment. Some of the proposals that specify the criteria to be applied when classifying these outcomes are Response Evaluation Criteria In Solid Tumors (RECIST)[24] and the World Health Organization (WHO) criteria (Choi et al. 2005). Depending on this outcome, administering the therapy or stopping it could be chosen, based on how well this therapy performed for that patient, having a great impact and improving his/ her health condition (Therasse et al. 2000). Nevertheless, this response is not always easy to deduce. When several cumulative therapies are administered to a patient, identifying which of the decisions in the healthcare process was related to which outcome is difficult. In cancer management, where administering a combination of chemotherapy, radiotherapy or hormonotherapy treatments to the patients is common, having these kinds of uncertainties when reporting the treatment response is common. Hence, this response usually relates to neoadjuvant therapies, given as a first step to shrink a tumor before the main treatment is provided, which is usually surgery.

3.2.3.2.1.2 Toxicities

Another approach that can give some hints about the impact that a therapy is having on a patient is the reporting of collateral toxicities or adverse events after its administration. An adverse event is understood as "any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a medical treatment or procedure that may or may not be considered related to the medical treatment or procedure" (Shapiro and Recht 2001). Until now, adverse events were retrospectively studied

[24] https://recist.eortc.org/

to determine their nature, incidence rates, typology, preventability, and impact on patients for a given illness (Zegers et al. 2009; Gagnier, Morgenstern, and Kellam 2017; Vincent, Neale, and Woloshynowych 2001). In previous years, some approaches have been proposed which identify, report, and grade these adverse events or toxicities in a more descriptive way, such as the US National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE)[25] proposal. The CTCAE identifies the different adverse events for a set of disorders, giving a definition that clarifies each adverse event term and some navigations notes to assist the reporter of the adverse event when choosing which is the most accurate effect to report (Trotti et al. 2003). Along with this textual description, a severity grading scale from 1 to 5 is provided, where:

- *Grade 1* refers to asymptomatic or mild symptoms or symptoms exclusively related to clinical or diagnostic observations with no indicated intervention,
- *Grade 2* describes moderate or minimal, local, non-invasive interventions that limit the correct performance of activities of daily living, such as going shopping, using the telephone or managing money,
- *Grade 3* indicates a severe or medically significant adverse event that can cause hospitalization or the prolongation of it, disabling the patient in some cases and requiring external support in order to perform self-care activities of daily living, such as bathing, dressing, and undressing or using the toilet,
- *Grade 4* reports life-threatening consequences and needs from an urgent medical intervention, and
- *Grade 5* means that the patient expired due to the effects of the adverse event.

CTCAE has been proved to be valid and reliable in large heterogeneous cancer patients populations (Dueck et al. 2015).

3.2.3.2.1.3 Patient-Reported Outcomes

Finally, the inclusion of Patient-Reported Outcomes (PROs) along with disease-oriented outcomes are promoted in order to track and evaluate the subjective perspective of the patient regarding the undertaken therapy (Trotti et al. 2003). These PROs give some information about the feelings of the patients, their satisfaction regarding the therapy they had been treated with,

[25] https://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm

and their current quality of life and health condition (i.e. clinical signs or symptoms and functional status of the patient), collected through questionnaires or interviews. The most commonly followed questionnaires promoting an informed and shared decision making between patients and clinicians for healthcare quality improvement are the ones proposed by the International Consortium for Health Outcomes Measurement (ICHOM)[26]. ICHOM has published standard sets of outcomes, measurement tools, and time points and risk adjustment factors for 28 different clinical conditions that are constantly reviewed and updated. In the particular case of breast cancer[27], the outcomes that mostly affect patients have been formalized by taking into account different information sources, such as complications suffered or adverse events during the treatment administration (e.g. CTCAE reported measurements), depression and anxiety episodes reported, overall and cause-specific survival measurements and different quality of life questionnaires. In the specific case of breast cancer, these questionnaires (e.g. the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire – Core Questionnaire EORTC QLQ-C30[28]) aim to report any decrease in the quality of life of the patient regarding (i) breast and body image after the suffered interventions (e.g. EORTC QLQ-BR2, BREAST-Q[29]), (ii) arm, breast and vaginal symptoms that affect sexual and daily life (e.g. EORTC QLQ-BR23), (iii) physical, emotional, cognitive and social functioning, taking into account the ability to work and the overall well-being (e.g. EORTC QLQ-C30), (iv) some collateral affections assessment regarding endocrine symptoms and possible metastasis because of the received therapies (see Figure 25).

[26] https://www.ichom.org/

[27] https://www.ichom.org/portfolio/breast-cancer/

[28] https://qol.eortc.org/

[29] http://qportfolio.org/breast-q/



Figure 25: The representation of the breast cancer standard set of ICHOM[30].

In conclusion, the measurement of these three kinds of outcomes (i.e. treatment response, adverse events or toxicities, and PROs) can report more accurately which are the most relevant factors that matter to breast cancer patients and affect their quality of life from the subjective point of view along with quantitative and objective measurements of the clinical performance and impact that the administered therapies have on patients' health status. These coexistent, different but complementary outcomes, computed as a single overall measurement, could guide clinicians towards a better and more complete clinical performance for the highest quality healthcare.

In the next section, a proposal for computing these outcomes in order to provide a strength for the experience-based rules presented in section 3.2.3.1 is given as well as a usability calculation based on the information gathered on each decisional-making process through the Decisional Event.

[30] https://www.ichom.org/portfolio/breast-cancer/

3.2.3.2.2 Usability and strength computation

In this section, a proposal to compute the trustworthiness and the strength of new knowledge coming from experience is performed. First, a measurement based on the decision-making process performance is carried out in order to compute the usability of a rule. Second, the outcomes explored in the previous chapter are computed altogether in order to provide an overall strength scale measurement.

Usability focuses on measuring the usefulness (i.e. the quality of being used) and usableness (i.e. the capability of being used) of a function in the context of evaluation (Jeng 2005). In the experience-based rules evaluation context, the usability computation is an objective measurement that gives some hints regarding the confidence clinicians have for a given rule recommendation. During the decision-making process, and leaning on the Decisional Event structure, all triggered rules for a given case study are recorded along with the final decision made, which, if compliant, will be one of the suggested recommendations coming from a triggered rule. Hence, two measurements are stored for each rule: the number of times this rule was triggered (i.e. *tExecuted* variable) and the number of events in which this rule recommendation was chosen as the final decision from all the executed times (i.e. *tFollowed* variable). The percentage of the usability of the studied rule is equal to the division of the following two variables: *tFollowed/tExecuted*.

The strength computation of the experience-based rules explores the previously presented outcomes (i.e. treatment response and adverse events following CTCAE) and the GRADE scale guidelines to propose a simplified approach that could compute the strength of the provided recommendations in an automatic way. Taking into account the strength scale proposed by GRADE (i.e. Strong, Strong against, or Weak) we propose including the outcomes in that scale as follows (see Figure 26Figure 26Figure 26):

- When CTCAE reported grades are equal to 1 or 2 and the treatment response has been total, with no relapse, and the rule is followed almost all the executed times, the SoR will be classified as Strong
- When CTCAE reported grades are equal to 3 or 4 and the treatment response has not been totally successful, hence, reporting a relapse, and the rule usability is not so high, the SoR will be classified as Weak

- When CTCAE reported grades are equal to 5 and the treatment has not responded at all, not having any impact on the disease progression and the rule usability is almost null, the SoR will be classified as Strong against.

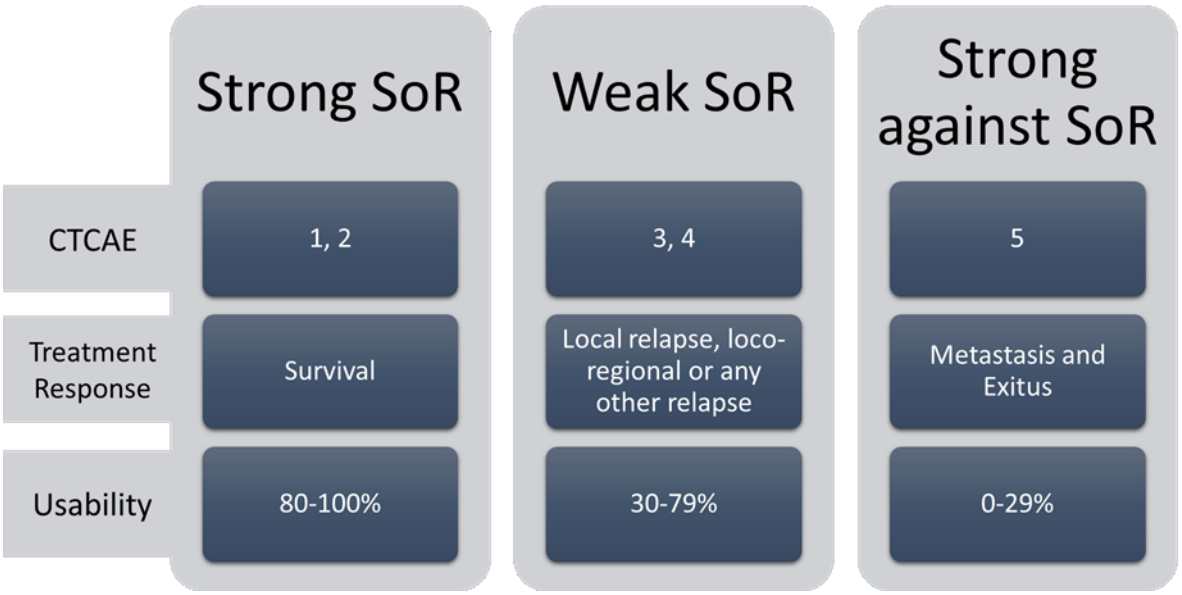


Figure 26: The strength of the recommendation computation approach based on GRADE using the toxicities, the treatment response, and usability values.

To ease understanding, a couple of examples are described next. If a rule in which usability has been computed as 90% (i.e. 90 out of 100 cases followed that recommendation when suggested) has reported that patients treated following the recommendation provided survived and did not suffer any adverse events or, if any, these events were grade 1 or grade 2 toxicities, we could compute the strength of that rule as strong, since the clinical performance is showing good effectivity and positive clinical impact of the rule. Nevertheless, when some discrepancies arise between the measured outcomes and the usability computation, the tendency should be computing the overall strength guided by the worst outcome result. Hence, if the reported CTCAE are from grade 3 or 4, but the treatment response was successful (i.e. the patient survived) and the usability is high (e.g. 80%) the strength will be weak, since the CTCAE detected that this therapy is causing important adverse events to the patient, maybe causing the need for refining the dosage or drug components proposed to be administered by this recommendation.

Finally, ICHOM questionnaire responses computation is proposed as a support to the analyzed strength grading scale. Computing these responses as an overall percentage that defines the improvement or decrease of the healthcare status of the patient from a subjective

point of view can give some insights when other clinical outcomes are not sufficient. Since PROs are a subjective measurement of the health status and quality of life, the results reported on these questionnaires will be processed as when studying the progression or decrease of health-related quality of life test results for cancer patients. The very first answered questionnaire will be used as the baseline for each patient, which will define the initial health status and quality of life level of the patient. Afterward, differential calculations will be made comparing the results of the recently answered questionnaires with the baseline questionnaire, which will estimate if there is an improvement or deterioration in the gathered results. These differential calculations would be made for each question, ultimately providing an overall score that summarizes the quality of life and health status of the patient in an objective way. Moreover, correlations and predictions could also support these measurements when the patient's follow-up is considerably long(Avis, Ip, and Foley 2006).

In conclusion, this chapter has proposed a usability measurement and a strength computation based on some clinical performance and outcomes results in order to back up the evidence behind this formalized experience, opening new research paradigms that would be interesting to explore in a deeper fashion. Usability can give some insights about the confidence or the strength of that rule since it shows how many times we prefer administering that therapy compared to others. Hence, when computed for a considerable number of patients, high usability percentages could be translated into strong strength, medium ranges of percentages into weak strength and low percentages into strong against. To support this measurement, and guided by the GRADE scale, the toxicity grades and the treatment response of the patients treated with the analyzed rule recommended therapy were also included in the strength computation. Lastly, PRO inclusion was explored to support those cases in which the different outcomes and usability measurements differ. Nevertheless, we should bear in mind that since several converging factors and therapy-related outcomes affect clinical performance, the clinicians' opinion will always be the last word when reporting this kind of strength measurements. Hence, even if we can attempt to compute the strengths of treatments in an automatic way, these results must be clinically validated and reviewed by a clinical team to assure their quality and validity.

3.3 Visual Analytics

According to the vast clinical information that we are able to retrieve in a computer processable way through the proposed Decisional Event formalization (see Chapter 3.1.1), it is essential to provide some visualization tools to aid clinicians in exploring the accumulated Real-World Data (RWD), including patient clinical data, decisions, and outcomes. These tools provide insights when evaluating the actual healthcare quality and clinical performance in an aggregated manner, identifying biased therapies or unexpected outcomes in an easy and intuitive way. Moreover, these tools can assist in the validation of evidence and description of new clinical hypotheses when exploring population data that reflects real-world evidence.

In this chapter, we propose two visualizations modules for (i) reporting all Decisional Events, compliance rates and the outcomes that were stored over time and (ii) for providing insights about Real-World Evidence (RWE).

3.3.1 Decisional Events visualization

Providing timely feedback regarding the clinical performance and the outcomes derived from said performance can cause clinicians to be more confident in their daily clinical practice and validate a good healthcare quality provision for hospital managers, clinicians, drug providers, and other stakeholders. In this chapter, a statistical dashboard is presented. It is focused on showing a summary of the information related to the decision-making process and its' impact (i.e. executed treatments, outcomes, and compliance rates) recorded over time through a decisional history at a glance. This dashboard is composed of 7 main charts.

First, a general overview of all decisions made over time is shown to the user (see Figure 27). We provide some filtering options for (i) selecting the X and Y axes to be shown in the graph, based on clinical variables of interest that could give some specific insights when analyzing the clinical performance, (ii) the clinical scenario whose performance the clinician is interested in analyzing or (iii) the hospital where the decision was made (see the graph in the upper left part of the Figure 27). For example, in the specific case of cancer, which is the use case of this study (see Chapter 4), the clinician may be interested in seeing which the most administered treatments for advanced-age patients, regarding their tumor size in the diagnostic phase, are. Hence, "age" and "tumor size" variables would be selected as the X and Y-axis of the graph respectively, in order to easily visualize the relationship between these two variables.

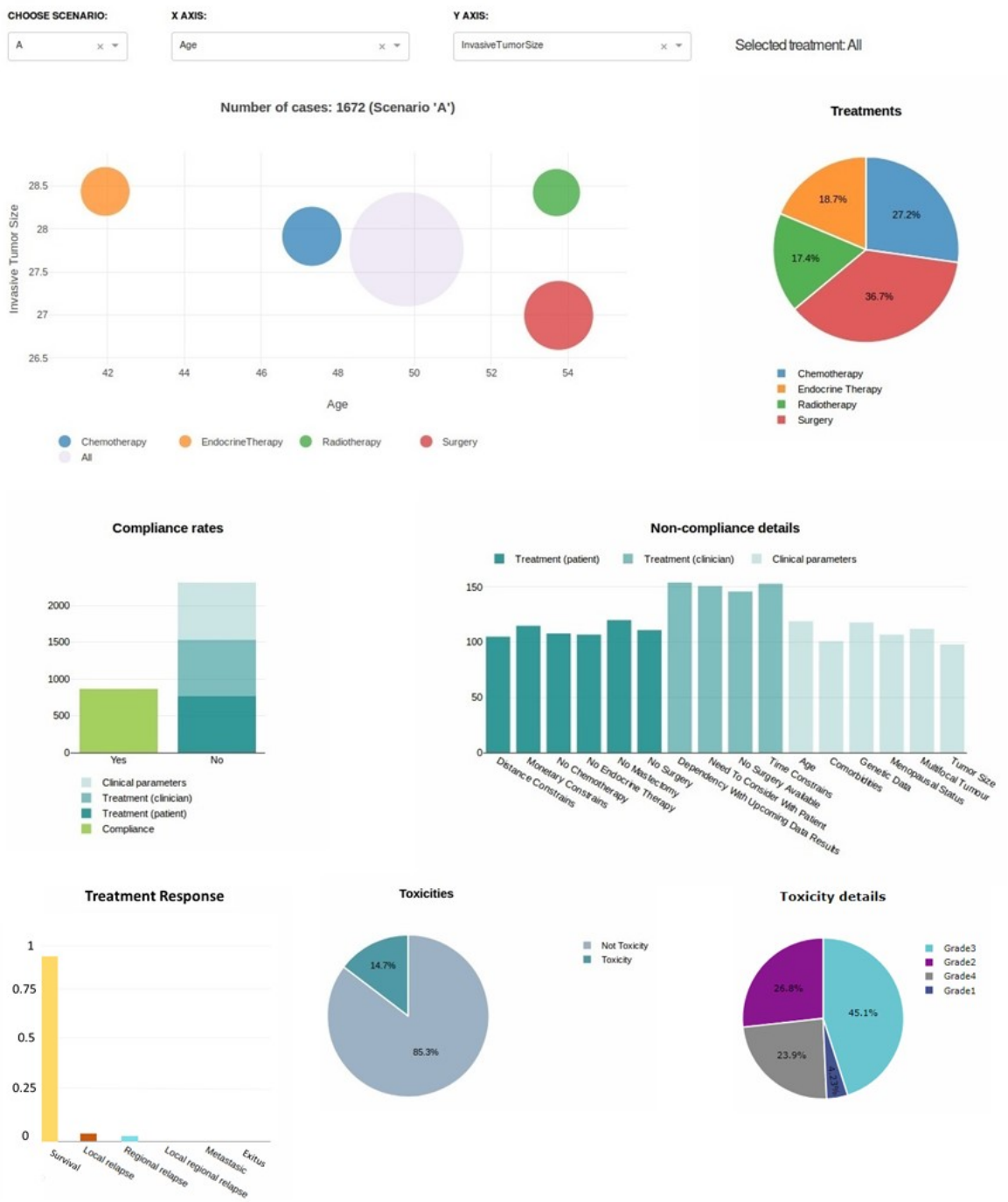


Figure 27: The Decisional Events history visualization dashboard.

In this Decisional Event history representation chart, each of the drawn bubble ratio is proportional to the number of decisions made matching the selected filters. Nevertheless, to make this information clearer, a second graph (i.e. treatments description graph) representing the number of times a therapy was administered as a percentage, is given (see the graph in the upper right part of the Figure 27). When clicking on any of the bubbles in the first graph, the rest of the representations will be updated according to the selected treatment and the treatment

description graph will be transformed into a graph providing specific sub-treatment information for the selected therapy, as shown in Figure 28 for chemotherapy treatment measurements.



Figure 28: A zoomed overview of all graphs for the specific cases treated with chemotherapy treatments.

Since the visualized data is retrieved from the Decisional Event history stored over time, compliance rates can be also visualized. These graphs will report the clinicians' adherence to

the CPGs and the main causes of non-compliant decisions. We provide this information using two graphs. The first one states an overall compliance rate in several cases, itemizing the non-compliant cases with colors according to their justification for non-compliance (i.e. non-compliance due to clinical parameters, clinicians' restrictions or patients' restrictions). The second graph provides the specific criterion of non-compliance that was selected during the decision-making.

Finally, to evaluate the clinical impact of the decisional history, the results of different patients' outcomes are also visualized, such as treatment response outcomes (i.e. survival, local relapse, regional relapse, loco-regional relapse, metastasis, and exitus) and toxicities or adverse events results based on the CTCAE grades scale computation, by percentage of patients suffering from any toxicities as well as the specifications of the present grades in percentage too (see Figure 27 lower mid-right graphs). These metrics are updated when clicking in any treatment by zooming on the specific information for the selected therapy (see Figure 28 for the specific case of chemotherapy treatment-related measurements).

3.3.2 Real-World Data visualization

The term Real-World Data (RWD) refers to data derived from diverse sources within the healthcare context and from a heterogeneous patient population in Real-World Settings, which can generalize clinical findings to include larger and more inclusive populations of patients, coming to generate Real-World Evidence (Sherman et al. 2016; Juliusson et al. 2009). In this thesis, we focused on providing a tool for exploring Real-World Data to identify factors that influence clinical outcomes, outside of a controlled environment (i.e. clinical trials' reports), providing not only insights about new clinical hypotheses, population trends and outcome expectancy, but also promoting the study of healthcare performance taking into account clinical results for administrative and monetary conclusions.

In our approach, we propose visualizing all the reported clinical information by using a Parallel Coordinates Plot (PCP) graph. The PCP represents the main "n" studied clinical variables as perpendicular parallel axis. To simplify the paradigm of over-plotting information, some data mining techniques have been implemented. These techniques include a pre-processing stage for cleaning the data and to avoid displaying non-representative and null values, a classification of the variables by their categories (i.e. integer, Boolean, string) and a grouping of related variables that provide the same clinical information. Moreover, a correlation

matrix among all studied variables for a given outcome was calculated. This matrix scores of all attributes from the most correlated to the outcome to least correlated, visualizing each of them as parallel vertical axes that respect this order. Once the perpendicular axes are defined, each clinical case will be represented as horizontal lines matching their values on the vertical axis and joining all of the recorded values as a final single line. This step restricts the drawn clinical cases to those with all relevant clinical variables fulfilled. Furthermore, ranges among the vertical axes can be selected to highlight the regions of interests. A color scale or heat map is used to assign colors to these lines, which reflects the possible values for the sought outcome (see Figure 29).

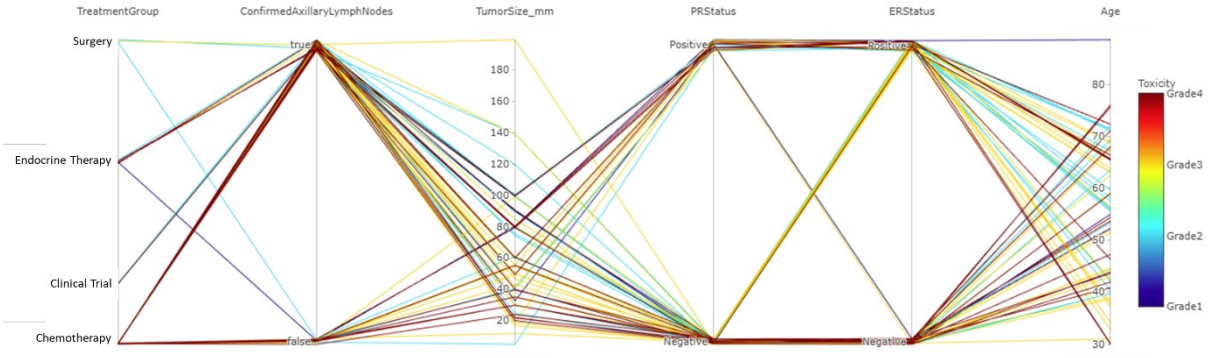


Figure 29: The Parallel Coordinates Plot (PCP) for outcome expectancy showing the results of 67 patients.

Other filters are also provided to the clinicians to be applied over the shown data and to help them identify the patterns of interest, such as grouping the information by a given clinical scenario or by a particular treatment. In order to give more freedom to the clinicians, the outcome to predict can also be changed and the number of represented axes may be modified according to their own criteria for exploring further approaches or different representations of the same information.

In Figure 29 and Figure 30, an example of Primary Breast Cancer (PBC) patients data visualization is shown. In Figure 29, 67 PBC patients were represented among the identified 5 most correlated vertical axes (i.e. the provided treatment, if there were any confirmed positive axillary lymph nodes, the tumor size in millimeters, the progesterone receptor status, the estrogen receptor status, and the age of the patient). The selected outcome was toxicities, painting each patient's clinical data in the color that represents its toxicity grade. In Figure 30, some ranges of restrictions were applied to the visualized axes in order to highlight specific clinical cases. We selected advanced age patients (over 65 years old) treated with chemotherapy

with confirmed positive axillary nodes and positive progesterone and estrogen receptor status. As shown in Figure 30, for at least 3 of the patients with tumors sized from 50, 60 and 80mm had particularly negative outcomes (i.e. Grade 4 toxicities), providing some insight into how chemotherapy affects older patients that have considerably large tumor sizes.

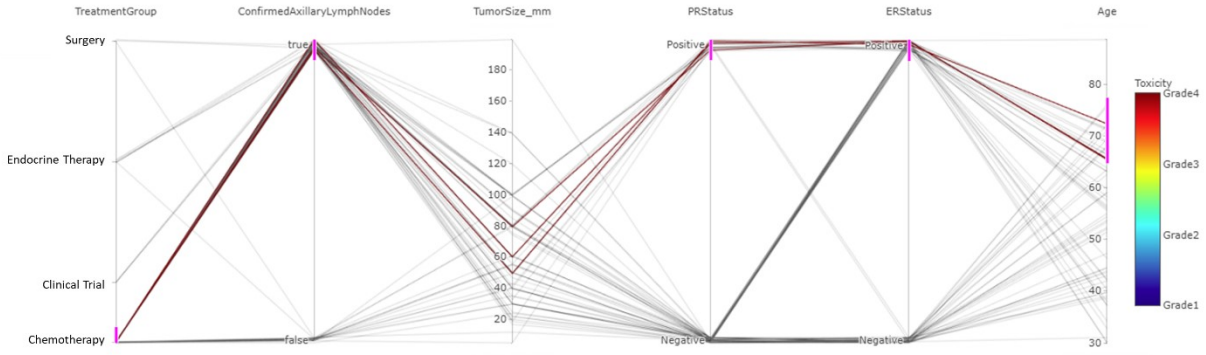


Figure 30: The Parallel Coordinates Plot (PCP) for outcome expectancy when applying some ranges restrictions among variables to highlight patients with higher toxicity grades.

4. Use Case: Breast Cancer

In this chapter, the case study in which the methodology presented in this thesis was implemented is shown. The research described in this thesis was developed within the DESIREE[31] project, funded by the European Union's Horizon 2020 research and innovation program under grant agreement No 690238 ("Decision Support and Information Management System for Breast Cancer | DESIREE Project | H2020" n.d.) and other internal projects of the eHealth and Biomedical Applications department at the Vicomtech[32] Foundation.

The purpose of the DESIREE project was to develop a web-based software ecosystem for the personalized, collaborative and multidisciplinary management of primary breast cancer (PBC) to support and assist specialized multidisciplinary clinical teams, known as tumor boards and more specifically for PBC, as Breast Units (BUs). DESIREE aims to aid the BUs during the case presentation, discussion, and delivery of the treatment, for best-personalized healthcare from diagnosis to follow-up.

Keeping in mind that breast cancer is the most widespread type of cancer in women worldwide, BUs focus on efficiently managing patients those afflicted by the disease, involving the different clinical specialties in charge of PBC management for the discussion of each clinical case setting guideline-based quality procedures and making decisions based on consensus and high quality of care. Nevertheless, daily clinical practice and case presentation during BU meetings is hampered by the complexity of the disease, the ever-growing amount of data available that defines the clinical case, the number of possible therapy options, the short time available to review each case, and the difficulty in deciding on cases that are beyond the guidelines' definition.

DESIREE provides the BUs with the latest available therapy options defined in several local and international guidelines (e.g. NCCN) using a guideline-based CDSS providing evidence-based recommendation for the current case – however complex it may be. In order to overcome the limitations of purely evidence-based support, two additional paradigms

[31] http://desiree-project.eu/;
[32] http://www.vicomtech.org/en/d2/ehealth-and-biomedical-applications

corresponding to experience-based CDSS and patient similarity-based CDSS are proposed within the DESIREE environment.

Within the scope of this project, the major part of this thesis was conducted to develop an experience-based CDSS which co-exists and interacts with a guideline-based CDSS during the decision-making process to overcome the limitations of purely guideline-based CDSS. To achieve this objective, the experience-based CDSS tracks all the decisions made over time, proposing an evaluation that considers several different outcomes. Moreover, this system is able to generate new knowledge from those cases that were non-compliant with the guidelines, which will increase the current knowledge closing some guideline gaps.

In the following chapters, we will present how the proposed methodology was implemented for BU support during primary breast cancer management within the DESIREE project, describing the different components developed for this purpose.

4.1 Breast Cancer Knowledge Model (BCKM)

The Breast Cancer Knowledge Model (BCKM) is an ontology-based formalization comprising the most relevant and meaningful minimum dataset defining a primary breast cancer patient (Bouaud, Guézennec, and Séroussi 2018). The purpose of building this complex model was to provide a common reference and standardized terminology for all modules and components within DESIREE, guaranteeing a clear and univocal definition of all clinical terms. Moreover, technical specifications, such as the structure of the rules in the guidelines, the relationships among the different variables modeled in the CPGs and novel sources of structured and non-structured information (e.g. imaging biomarkers, genetic data), are also defined, providing a holistic view of the patient. The BCKM is able to represent all this information throughout time in a structured and consistent manner, suitable for case presentation and decision support.

Building the BCKM was a multidisciplinary effort among clinicians and knowledge engineers of several institutions within the DESIREE project that, after studying the different available CPGs, protocols, and data sources in PBC along with the clinicians' own experience and relevant cases study, agreed upon the necessary knowledge and how to model it in a structured way leaning on an ontology ("D2.4. Digital Patient Model and Structure" 2019; Bouaud, Guézennec, and Séroussi 2018).

In the DESIREE BCKM, 628 classes, 35 object-type properties, and 11 data-type properties were defined. These classes or entities were structured under three main classes and related or grouped by different data-type and object-type properties. The *DESIREEEntity* class defines all clinical variables that compose a PBC patient, the *ModelEntity* class models the clinical variables classes under *PatientEntity*, *LesionEntity* or *SideEntity*, depending on which of them they are related to (e.g. the clinical variable *Age* is an attribute of the patient, hence, it is related to the *PatientEntity* class). Some object-type properties were built to describe this relationship among variables, such as *hasPatient*, *hasSide* or *hasLesion* object-type properties. Finally, the *DssEntity* class supports the reasoning process by defining some technical requirements for the CIG formalization, such as *RecoEntity* or *OrderEntity* classes (Guézennec G.,. 2018).

Once all entity concepts have been defined, the possible values that they could take on were also identified. For that, attribute concepts were used, under the *DESIREEAttributes* class, related to their entity class by object-type and data-type properties. For example, the *HistologicType* attribute of a *LesionEntity* uses the object-type property *hasRange* to refer to the *BreastCarcinoma* class, within the *DESIREEValue* class (see Figure 31).

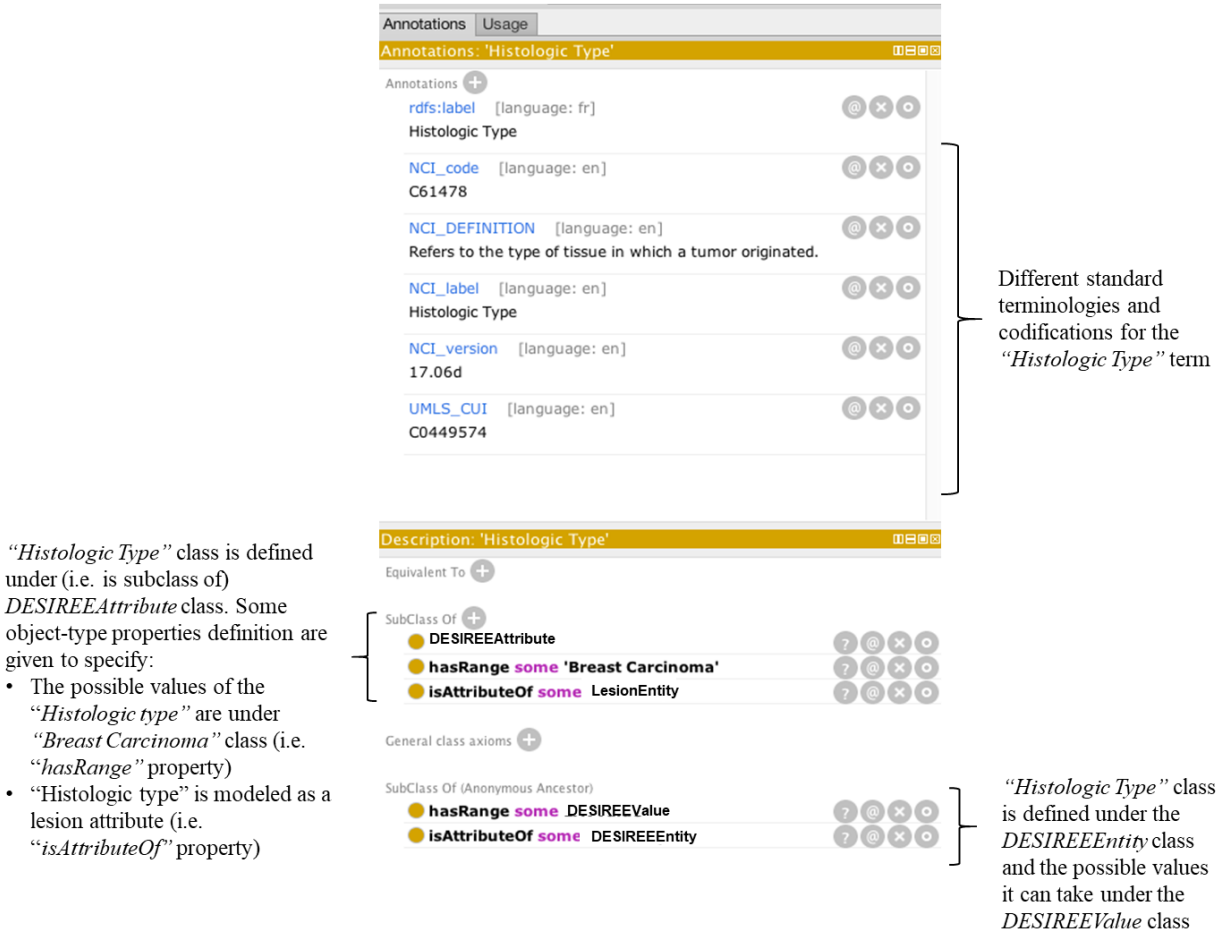


Figure 31: Example of the "Histologic Type" class definition within the DESIREE BCKM.

Finally, in order to assure the standardization, interoperability, and the correctness of the defined terms, some existing authoritative terminologies were implemented, such as National Cancer Institute Thesaurus (NCIt), Logical Observation Identifiers Names and Codes (LOINC) or Systematized Nomenclature of Medicine - Clinical Terms (SNOMED CT). In DESIREE, mainly NCIt was implemented since it is open source and contains a broad number of identified clinical concepts. This was made through four annotation properties (see Figure 31):

(i) *NCI_ DEFINITION*, where the definition of the concept is given,

(ii) *NCI_ALT_DEFINITION All*, where an alternate definition is provided (if any),

(iii) *NCI_code* gives the code identifying the clinical term and

(iv) *UMLS_CUI* provides the UMLS[33] terminology-based code.

[33] https://www.nlm.nih.gov/research/umls/sourcereleasedocs/index.html

Since some of the attributes were not defined in any of the terminologies, an internal code generation was required. The inclusion of the UMLS codification was made to facilitate the interoperability with external systems.

This BCKM ontology was the backbone of all the modules in the project, assuring their interoperability and defining in a standard way all clinical variables and their relationships and restrictions regarding PBC patients' characterization and clinical knowledge defined within the CIGs.

4.2 Breast Cancer Clinical Practice Guidelines

In the DESIREE project, the National Comprehensive Cancer Network (NCCN) and the European Society for Medical Oncology (ESMO) guidelines were studied. Moreover, as explained in Chapter 2, to promote clinicians' adherence to CPGs, local protocols followed by 2 clinical partners on the project were also formalized: Onkologikoa foundation[34] protocol and Assistance Publique- Hôpitaux de Paris[35] (APHP) protocol.

Five scenarios have been defined within the DESIREE consortium to characterize the PBC patient management stage along the breast cancer clinical pathway:

- Scenario A: After Diagnosis – Treatment has not started
- Scenario B: After Neo-adjuvant Therapy – No surgery
- Scenario C: After surgery – After Neo-adjuvant Therapy
- Scenario D: After surgery – No Neo-adjuvant Therapy
- Scenario E: After surgery – Incomplete Adjuvant Therapy

For each scenario, the studied clinical parameters, the applied CPG rules, and the recommended treatments or procedures may differ. Moreover, the CPGs themselves divide the contained knowledge into the following groups:

(i) rules that characterize patients based on their clinical parameters' values,
(ii) rules that provide procedure, tests and examination information in the breast cancer domain and

[34] https://www.onkologikoa.org/en/
[35] https://www.aphp.fr/

(iii) rules that diagnose the patient, based on the knowledge gathered from (i) and (ii) rule groups.

When formalizing paper-based CPGs or local protocols into CIGs, a two-step approach was followed within DESIREE. First, CPGs were formalized in a human-readable language named Natural Rule Language (NRL). NRL tagging allows matching the defined content with an underlying knowledge model in an easy way. In Figure 32 an NRL representation of a simple rule is provided as an example. This rule is defining that a lesion (i.e. a *LesionEntity* type class named in this particular case *theLesion*) with null estrogen and progesterone receptors, will set the clinical variable *HormonalReceptorPositive* related to the same lesion (i.e. *theLesion*) to false since no analyzed lesion receptor had a positive value. Hence, *HormonalReceptorPositive* is a Boolean (i.e. its value range is defined within the BCKM with a *hasRange* data-type property pointing at *BooleanValue* range) entity whose possible values are true or false. On the contrary, *ERResult* and *PRResult* are two classes defining the *LesionEntity* whose possible values are defined as an integer or natural number (i.e. *IntegerValue* range in the BCKM).

```
Context: LesionEntity
-- ==============================
Action Rule "NCCN-GR02-RL11 | NCCN: IF ER = 0 AND PR = 0 THEN HR-"
-- ==============================
        "theLesion" is a LesionEntity,
If
                theLesion.ERResult = 0
        AND     theLesion.PRResult = 0
Then
        set theLesion.HormonalReceptorPositive to false
;
```

Figure 32: NRL representation of a simple rule.

The second step formalizes these model and CPG dependencies and deductive reasoning, as an XML-based computer processable and executable language, named Desiree Guideline Document Language (DGDL). The DGDL was built as a model-independent language, allowing the reasoning over one or more models at a time (i.e. the BCKM in this case) but decoupled from the models. This kind of approach facilitates the implementation, update, and evolution of the formalized knowledge within CIGs for new incoming relevant data and evidence. From the structure point of view, the DGDL is composed of a number of declarations for describing the rules, inspired in HL7 Standard-Clinical Decision Support Knowledge Artifact Specification[36], and detailed in (Guézennec G., 2018). In Figure 33, an

[36] https://www.hl7.org/implement/standards/product_brief.cfm?product_id=337

example of the rule expressed in NRL in Figure 32 is given, following the XML schema that contains the references and the logic expressions to the BCKM.

```
<Rule id="NCCN-GR02-RL11" name="NCCN: IF ER = 0 AND PR = 0 THEN HR-" context="LesionEntity">
    <Context kind="global" var="theLesion" type="LesionEntity"></Context>
        <Conditional>
            <If>
                <BinaryOperator op="AND">
                    <Condition op="EQUAL">
                        <ModelRef var="theLesion" att="ERResult"></ModelRef>
                        <Literal type="integer">0</Literal>
                    </Condition>
                    <Condition op="EQUAL">
                        <ModelRef var="theLesion" att="PRResult"></ModelRef>
                        <Literal type="integer">0</Literal>
                    </Condition>
                </BinaryOperator>
                <Then>
                    <Seq>
                        <Assign>
                            <ModelRef var="theLesion" att="HormonalReceptorPositive"></ModelRef>
                            <Literal type="boolean">false</Literal>
                        </Assign>
                    </Seq>
                </Then>
            </If>
    </Conditional>
</Rule>
```

Figure 33: DGDL representation of a simple rule.

This has been the standard format used within DESIREE for CIG definition and rule transmission among different modules, such as the guideline-based CDSS and the experience-based CDSS. An alternative approach to CIG formalization was proposed in Chapter 3.1 and 3.2 in order to mitigate the risk of erroneous knowledge formalization during the two-step methodology for CPG formalization and provide formalization tools easing this process in an integrative way among the architecture presented in this thesis.

4.3 Non-compliance criteria definition

In order to identify the main causes of non-compliant decision in PBC management, a list of the most common non-compliance criteria was formalized within the BCKM, with the help of the clinicians involved in the DESIREE project. These criteria were classified into three main groups as follows:

(i) treatment-related patient restrictions, or patient preferences related criteria,

(ii) treatment-related clinician restrictions, due to clinicians' habits or available medical facilities for diagnosis, treatment, and follow-up, and

(iii) clinical parameters related to clinical restriction, meaning complex clinical case restrictions.

Several justification options were defined within each of these groups to be selected when a CPG non-compliant decision is made for a given patient during a BU tumor committee, in order to define these reasons for guideline non-compliance and later use this information when generating a new experience-based rule (see Figure 34). For example, when an advanced-age patient must be treated, some doubts may emerge since guidelines do not reflect all of the potential complications that may arise when providing the suggested therapy to this particular patient. Hence, the clinicians may evaluate the trade-off between the benefits and harms that the CPG-recommended treatment could cause to the patient and, in case other options are considered to be more suitable, choose the "*Age*" justification within the clinical parameter related to the clinician restriction group. This non-compliance justification, along with the final decision taken by the BU to treat this patient, will be transferred to the experience-based CDSS, where a new experience-based rule will be generated following the procedure explained in Chapter 3.2.3.

This new experience-based rule will be triggered for any clinically similar incoming patient matching and accomplishing the conditions defined within its conditional part when queried by the guideline-based CDSS, providing the experience-based recommendation along with the guideline-based recommendations.

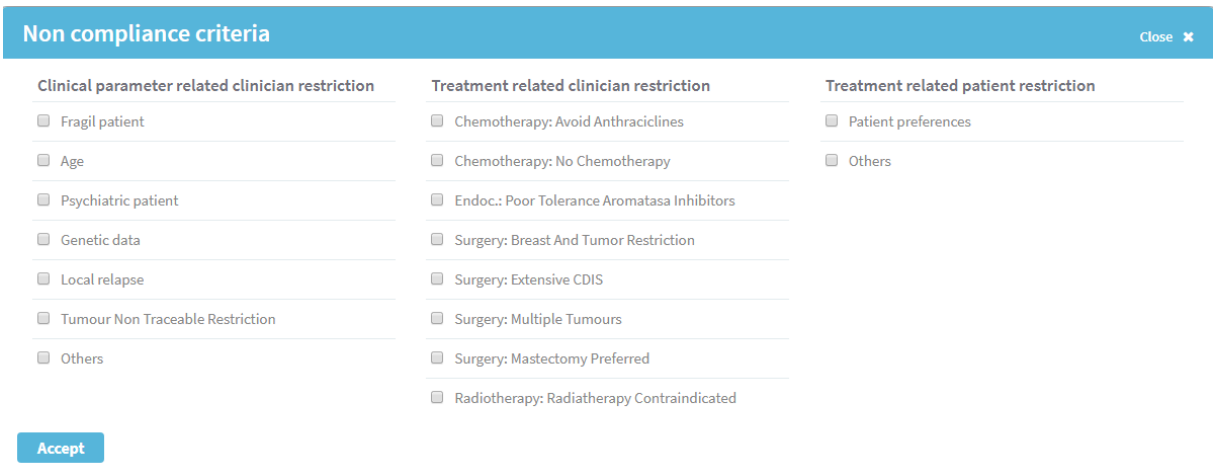


Figure 34: Non-compliance justifications options provided to BUs during the decision-making process.

4.4 Breast cancer outcomes

As stated in Chapter 3, several different outcomes were studied for measuring the strength of the provided recommendations. The specific outcomes used for PBC management were:

(i) the first treatment response measurement, identifying if there was any decrease in the lesion size after neo-adjuvant therapy administration and before the main treatment, which is usually a surgical procedure. This response can be reported as a survival, when the treatment was successful and shrank the lesion size or did not cause any worse status, different kind of relapses (i.e. local, regional or loco-regional) when the return of the PBC is reported after already being treated, a metastatic situation or the exitus of the patient.

(ii) the toxicities or adverse events reported following the Common Terminology Criteria for Adverse Events (CTCAE) v5[37], and

(iii) the patient-reported outcomes through the European Organization for Research and Treatment (EORTC) QLQ-BR23 and ICOHM-BREAST questionnaires, which focuses on the quality of life evaluation during the post-diagnosis and treatment phases.

These outcomes were gathered and related to each patients' Decisional Event to provide some insights into the healthcare quality and performance achieved in each case. Moreover, these results were shown in the parallel coordinates visual analytics graphs to study the relationships among the diverse clinical variables and the reported toxicities and treatment responses. Finally, an approach was also proposed within this thesis (see Chapter 3.2.3.2.) in order to provide a recommendation's strength based on the computation of these three kinds of outcomes.

37 https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_5x7.pdf

4.5 Interaction of the guideline-based CDSS and the experience-based CDSS within DESIREE project

Within the DESIREE project, the experience-based CDSS was implemented and used with the following objectives:

(i) the formalization, retrieval, storage and analysis of all decisions made over time by project participant hospitals' BUs,
(ii) the evaluation of these decisions quality and impact based on different reported clinical outcomes and CPG compliance measurements,
(iii) the generation of new experience-based rules from the identification of non-compliance causes to manage guideline gaps or reported biases, and
(iv) the provision of different visualization dashboards, which by allowing the analysis of the obtained data and conclusions, help understanding and reporting new evidence or validating the actual clinical care in a visual and user-friendly way.

With these objectives in mind, the methodology explained in Chapter 3.2 was implemented as a Software as a Service (SaaS) within a secure cloud-based software infrastructure for the coordination and management of clinical cases in BUs. This kind of implementation also facilitates the accessibility and communications among the different modules in the project.

In DESIREE, the experience-based CDSS and the guideline-based CDSS interact in order to augment the knowledge base of the CDSS with experience, being able to provide support when complex cases out of the guidelines are evaluated. To assure the correct communications between these two modules, the experience-based CDSS was linked to the BCKM in order to formalize the clinical contents using the established terminologies. First, all rules composing the knowledge that the guideline-based CDSS leans on and which are formalized as a DGDL XML-format file (Guézennec G., 2018) were retrieved, stored and saved into a database following the *Rule* structure presented in Chapter 3.1.1. Afterward, when a non-compliant case is identified during a BU meeting, the following information is transferred to the experience-based CDSS:

- the non-compliant patient identifier,

- the unit or hospital where the decision was made,
- the queried CPGs by the BU during the decision-making process,
- the triggered rules and the provided recommendations,
- the non-compliance criteria, and
- the newly defined final decision.

This information is used to generate the new-experience-based rule which will be formalized following the DGDL XML-format and transferred to the guideline-based CDSS. A more detailed technical description of how this was done can be found in (Muro N., 2018).

For the clinical data management, a FHIR standard-based server was used to transfer the anonymized clinical data (each patient was assigned with a DESIREE identifier) to all modules and allow the reasoning over it. The experience-based CDSS relies on two modules to perform this task (i) a FHIR decoder that retrieves the FHIR bundle with all of the patient´s clinical information to be studied and (ii) a FHIR transcoder, that transforms this information into processable parameters for the experience-based CDSS and stores this information in the database presented in the Chapter 3.1.1. This clinical data, along with the recorded decisions, is statistically analyzed and presented lately through the visual analytics module.

5. Validation

The objective of this chapter is to provide the validation protocol which was followed to test the successive versions of the experience-based CDSS on both technical and clinical sides. First, a technical validation was performed to assure a correct unitary and integrative functioning of the system from the performance point of view. Following this, a clinical validation was made for studying the clinical impact and validity of the experience-based CDSS in real clinical settings. This validation methodology was performed within the DESIREE project (Gautier P., Pelayo S., Blancafort C. 2017).

5.1 Technical Assessment

In this chapter, a validation framework design is presented on the requirements and procedure to evaluate and technically test each of the functionalities of the experience-based CDSS and its integration with other modules concerning its quality, efficiency and assuring an error-free software implementation.

The technical validation mainly focuses on defining the quality of the analyzed software at two levels, (i) regarding the level at which the system meets the specified requirements and (ii) the degree to which the system meets the needs or expectations of the end-user ("IEEE Recommended Practice for Software Design Descriptions" 1998). To achieve this objective, several information about the quality of the software and risk of its failure to end-users (i.e. if it meets the design requirements, responds correctly to the expected kind of inputs, performs its functions within an acceptable time obtaining the expected results, is usable and meets the requirements to be running in its intended environments) is measured and evaluated through a technical validation. The purpose of the technical validation is ensuring that the software delivers the functionalities or results that were identified when defining the software design and detecting and solving failures that may affect the software's performance.

The methodology followed to technically validate the work developed in this thesis was inspired by the ISO/IEC 25010:2011 International standard protocol or SQuaRE (Software engineering-Software product Quality Requirements and Evaluation) regarding system and software quality[38], which defines a model for measuring quality in use or outcome of interaction

[38] https://www.iso.org/standard/35733.html

when a product is used in a particular context of use. It defines a set of characteristics and relationships between them, which provide the basis for specifying quality requirements and evaluation of the whole system (Miguel, Mauricio, and Rodriguez 2014).

For the technical validation assessment of the developments made within the DESIREE project, a simplified version of the ISO/IEC 25010:2011 standard was applied, measuring 5 out of 8 defined characteristics (see Table 3).

Characteristic	Definition of the ISO/IEC 25010:2011
Functional suitability	Degree to which a system provides functions that meet the stated needs when used under specified conditions, measuring its appropriateness, accuracy, and compliance.
Performance efficiency	Performance relative to the amount of resources used under specified conditions, hence, a time behavior and resource-utilization balance.
Compatibility	Measuring to what extent a system can be replaceable, interoperable and co-exist (i.e. exchange information with other systems or components, and/or perform its required functions) with other systems.
Usability	Degree to which a system can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use
Reliability	Degree to which a system or component performs specified functions under specified conditions for a specified period of time in an appropriate and accurate way.

Table 3: The selected ISO/ IEC CD 25010: 2007 model characteristics to be tested for the software quality evaluation.

Usually, software testing techniques are classified into three levels: (i) unitary tests applied to small pieces of code to verify their correct functioning and the achievement of the searched objective. Once these unitary tests are validated, the next step is performing (ii) integrations tests that verify that communications among different modules of software are

correctly performed and that the generated input/output in each of them is the expected input/output. Finally (iii) the entire system validation is performed, testing the platform as a whole and validating that all requirements are achieved as a whole. In the following chapters, the unitary and integration validations performed on the experience-based CDSS will be presented, as some system validation results within the DESIREE project platform.

5.1.1 Unit test design and implementation

During the technical validation of the experience-based CDSS, different information regarding each of the system's components was collected to have a short technical description of the requirements and objectives of each functionality within this software system. The retrieved characteristics are described below:

- A *component ID*, to identify each of the components or functionalities in a unique and unequivocal manner and a *component name* that describes the functionality a bit further as a caption.
- A specification of the *function* to be tested and a *function description* with some clarifications about the expected performance and results.
- An *input configuration*, defining the input requirements needed for the correct functioning of the system and the interactions needed from the users' point of view (e.g. click on the options provided through an interface in order to proceed in one way or another in the functionality's pathway).
- The description of the *expected output* for that functionality of the system after performing over the received input.
- The SQuaRE (Software Quality Requirements and Evaluation) characteristics of the analyzed functionality
- The *test* function performed over the software component
- The obtained *result*

A classification of all of the subsystems or modules composing the experience-based CDSS is detailed in Table 4.

Subsystem ID	Subsystem name	Description
SUBSYS_1	Rule Generation module	Receive the response from the BU and analyzes the Final Decision taken by Clinicians. In case the FD is not compliant with the CPGs' provided recommendations a new rule is generated to address this clinical case.
SUBSYS_2	Data exploitation and Knowledge Discovery	Exploits and visualizes the clinical data retrieved from FHIR messaging using data mining and visual analytics techniques.
SUBSYS_3	(i) FHIR Decoder and (ii) FHIR Transcoder	Retrieves and parses each new patient case from the FHIR server and translates the FHIR bundle into the clinical information to be stored as clinical parameters that define a patient within a Decisional Event into the decisional history database.
SUBSYS_4	Decisional History	Not a functional module but structural, representing all needed information in a computer interpretable way for the experience-based CDSS. It contains all the information related to a decision. The other modules rely on this structure for storing, computing and retrieving all of the necessary information.

Table 4: Classification of all of the subsystems or modules of the experience-based CDSS.

Once all of the subsystems have been defined, Table 5 shows all modules within each subsystem. These modules have been tested with 4 (i.e. Functional suitability, Performance efficiency, Usability, Reliability) of the 5 selected characteristics of the ISO/ IEC CD 25010: 2007 standard since Compatibility characteristics are measured during the integration test.

Component ID	Component name	Function	Function description	Input configuration	Expected output	Description of the performed tests
SUBSYS_1_01	Rule Generation module	Get BU response	Retrieves all of the information used during the decision-making process by the BU, including the final decision made (compliant or non-compliant)	XML format response file containing: (i) the patient ID, (ii) the triggered guideline(s) IDs, (iii) the triggered rule(s) IDs from the identified guideline(s), (iv) the provided recommendation(s) from the guideline(s), (v) the final decision taken by the BU, (vi) the scenario in which this patient is and to which this decision affects, (vii) if the made final decision was	This information is stored as a Decisional Event in a Java-based structure and in a database (see SUBSYS_4). It will automatically trigger SUBSYS_1_03to update some triggered rule parameters. Only if the case was CPG non-compliant, the SUBSYS_1_02 will be triggered as well.	Check communications and correct data storage and triggering of SUBSYS_1_02 and SUBSYS_1_03.

				CPG non-compliant, the followed non-compliance criteria.		
SUBSYS_1_02	Rule Generation module	New experience-based rule generation	Generates a new rule from a CPG non-compliant decision tracked by SUBSYS_1_01	Same as SUBSYS_1_01	XML format rule that is stored in a database and added into the guideline-based CDSS rule-base to be triggered for incoming similar patients	Check communications for posting the new rule into the guideline-based CDSS, the correct data storage in the experience-based CDSS database and the correct structure (and clinical content supervision in 70% of the technical tests and 100% of clinical tests) of the generated XML files

SUBSYS_1_03	Rule Generation module	Parameters Provider	Returns the usability and strength (if available from standard guidelines) of the triggered rule(s) during the decision-making process for a given patient	XML format response file containing: (i) the triggered guideline(s) ID(s), (ii) the triggered rule(s) ID(s) from the identified guideline(s) (Since this function is triggered by SUBSYS_1_01 the same message is used)	The usability and strength metrics for the triggered rule(s) added into the BU response XML (input message for SUBSYS_1_01)	Check the update of the usability metrics in the experience-based CDSS database and the communications and modification of the BU response XML file accordingly
SUBSYS_2_01	Data exploitation and Knowledge Discovery	Correlation matrix generation	Analyzes the clinical data stored in the experience-based CDSS database (see SUBSYS_3_01 and SUBSYS_3_02) and generates a	Experience-based CDSS containing clinical data processed from FHIR server (see SUBSYS_3_01 and SUBSYS_3_02)	Correlation matrix among all clinical variables	Check the calculated correlations and most correlated variables, communications between the generated matrix, the database and

			correlation matrix between all clinical variables to be shown using the parallel coordinates graph given objective outcome (SUBSYS_2_013)			the visualization interfaces (SUBSYS_2_013)
SUBSYS_2_02	Data exploitation and Knowledge Discovery	Decisional Events and outcomes visualization	Visualization of the information retrieved by the SUBSYS_4 using a dashboard interface	Experience-based CDSS database containing all decisional history (SUBSYS_4)	Visualization of the information related to all decisions made over time and the outcomes of those decisions stored in the experience-based CDSS database in a Graphical User Interface (GUI) dashboard	Check communications between the database and the visualization interfaces

SUBSYS_2_03	Data exploitation and Knowledge Discovery	Parallel coordinates plot	Visualization of the information in the SUBSYS_3 using a parallel coordinates plot	Experience-based CDSS containing clinical data processed from FHIR server (see SUBSYS_3_01 and SUBSYS_3_02)	Visualization of the results from SUBSYS_2_01 plotting them in a parallel coordinates graph Graphical User Interface (GUI)	Check communications between the database and the visualization interfaces
SUBSYS_3_01	FHIR Decoder and Transcoder	FHIR Decoder	Parser to read the clinical information coming in a JSON message that follows the FHIR standard	JSON format message following the FHIR standard, which contains all clinical information of the analyzed patient, including the outcomes	Java-based object containing all the clinical information in a processable way	FHIR message generation and correct structuration, communication with FHIR server and correct gathering of the information in the Java-based object
SUBSYS_3_02	FHIR Decoder and Transcoder	FHIR Transcoder	Parser to read a Java-based object and make the needed operations over the clinical	Java-based object containing all of the clinical information in a processable way	Storage of all of the clinical data in the experience-based CDSS database to be used	Correct communication between the Java-based object and the database and

			data and save it in a database		in SUBSYS_1 and SUBSYS_2	correct transmission of all the clinical information
SUBSYS_4_01	Decisional History	Decisional Event generation and storage	Gathers all information in a computer interpretable way for the experience-based CDSS. It contains all the information related to a decision (SUBSYS_1_01).	The Decisional Event object from the input of SUBSYS_1_01	Storage of the decision in the experience-based CDSS database	Decisional Event generation and database communication and correct information storage

Table 5: Table describing the measured features of each of the functionalities of the experience-based CDSS for each performed unit test.

5.1.2 Integration test design and implementation

After testing one-by-one each of the units in the four subsystems that make up the experience-based CDSS and ensuring their correct functioning, integration tests between the different modules and the capability of communication with other external modules were studied and performed.

Following the SQuaRE characteristics defined within the ISO/IEC 25010:2011 standard, compatibility tests were performed throughout all of the communication pathways identified in the system architecture design and implementation, detailed in Figure 35, to ensure correct data reception, exchanges among modules and expected results output.

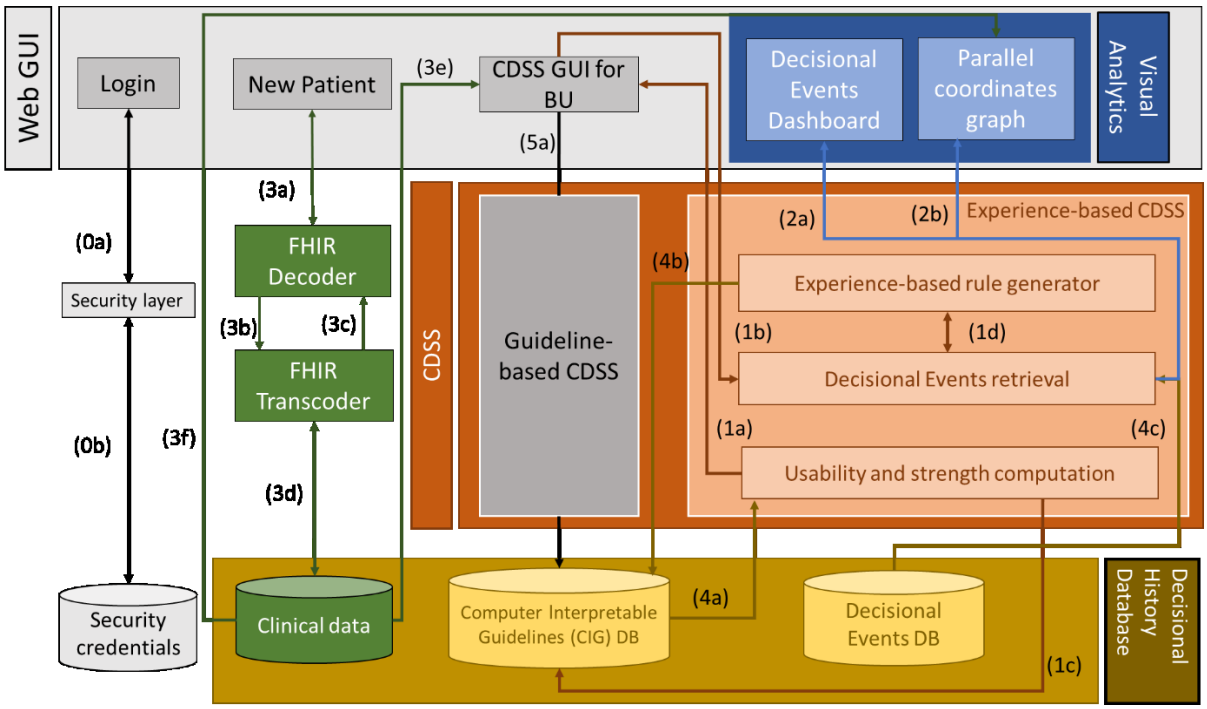


Figure 35: System architecture with all modules and communications.

In the following table (Table 6), a description of all the information exchanges and test results are provided.

Communication ID	Communication description	Transmitter module	Message	Receptor module
3a	FHIR bundle retrieval/sending	FHIR Decoder/ GUI	JSON with patient's clinical information structured following the FHIR standard	GUI/ FHIR Decoder
3b	FHIR bundle parsing into Java-based object with all clinical information	FHIR Decoder	JSON with patient's clinical information structured following the FHIR standard	FHIR Transcoder
3c	From Java-based object with all clinical information to FHIR standard-based bundle translation (e.g. to send care plans)	FHIR Transcoder	Java-based object with all of the clinical information	FHIR Decoder
3d	Clinical information storage into DB	FHIR Transcoder	Java-based object with all of the clinical information	Clinical Data
3e	Clinical information retrieval to be sent to the BU committee's interface to be evaluated by guidelines	Clinical Data	Gathered database record with the necessary clinical information	CDSS GUI for BU (Web GUI)
3f	Clinical information retrieval to be shown in parallel coordinates (data mining analysis in the Decisional Events retrieval module). Correlation matrix generation or update using all clinical attributes.	Clinical Data	Gathered database record with the necessary clinical information	Parallel Coordinates Graph (Visual Analytics)
4a	Gather the computation and strength measurements from the database for the triggered rules for the analyzed clinical case	Computer Interpretable Guidelines (CIG) DB	Triggered rules IDs	Usability and Strength computation

4b	Supplement the guidelines with experience rules from CPG non-compliant cases	Experience-based rule generator	The Java-based object containing the new experience-based rule	Computer-Interpretable Guidelines (CIG) DB
4c	Decisional Events retrieval in a Java-based format for clinical performance evaluation	Decisional Events DB	Decisional Event record in the database	Decisional Events retrieval
1a	Visualization of the usability and strength computation metrics for the triggered rules to be visualized by the BUs	Usability and Strength computation	XML-based document containing Usability and Strength computation metrics for the triggered rules	CDSS GUI for BU (Web GUI)
1b	Decisional Event generation from the BU for a given patient	CDSS GUI for BU (Web GUI)	XML-based document containing the information of the generated Decisional Event	Decisional Events retrieval
1c	Update the usability and strength computation metrics after a new case study into the database	Usability and Strength computation	XML-based document containing Usability and Strength computation metrics for the triggered rules	Computer Interpretable Guidelines (CIG) DB
1d	Record a CPG non-compliant case and generate a new experience-based rule from it	Decisional Events retrieval/ Experience-based rule generator	The Java-based Decisional Event information	Experience-based rule generator/ Decisional Events retrieval
2a	Visualization of all Decisional Events over time and their different associated outcomes to be visualized in	Decisional Events retrieval	All stored Decisional Events information	Decisional Events Dashboards (Visual Analytics)

	the Decisional Events Dashboards (Visual Analytics) GUI			
2b	Visualization of all clinical cases with the outcomes generated after a Decisional Event was performed	Decisional Events retrieval	All clinical information linked to the generated outcomes from decisions made over time	Parallel Coordinates Graph (Visual Analytics)

Table 6: Definition of all integrations and communications exchanges within the system.

In addition to the defined communications, the architecture comprises other necessary functional interactions detailed in Table 7.

Communication ID	Communication description	Transmitter module	Message	Receptor module
0a	Security validation	Login interface/ Security layer	Login credentials transmission	Security layer/ Login interface
0b	Credentials verification	Security layer/ Security credentials	Login credentials validation	Security credentials/ Security layer
5a	Clinical data analysis by guidelines	CDSS GUI for BU (Web GUI)	Clinical information of the patient to be analyzed by the CPGs	Guideline-based CDSS

Table 7: Other interactions defined for the correct functioning of the system

5.2 Clinical assessment

This chapter aims to validate the proposed system design and determine whether experience-based CDSS could handle the limitations of purely guideline-based CDSS and impact on the prescribing behavior of clinicians when implemented in practice.

This clinical validation was performed within the DESIREE project, in which the inclusion criteria applied to the participating patients were (i) having Primary Breast Cancer, (ii) being older than 18 years old and (iii) being informed about the study, signing a patient consent that allowed for the use of their clinical data for research purposes.

During this clinical validation, two different scenarios were set up. First, a validation with retrospective patients (i.e. patients to which some decisions were applied following past BUs meetings) in simulated BU meetings (i.e. performed under pilot conditions as close as possible to real conditions) was carried out to track the compliance rates and generate new experience-based rules to be applied in future cases. In the second phase, prospective patients were studied during simulated BUs to validate the previously generated rules and check the experience-based CDSS performance in clinical practice settings.

5.2.1 Experience generation from retrospective data in simulated BUs

In this first step of the clinical validation, retrospective patient cases were analyzed within simulated BUs. Retrospective patients concerning previously discussed cases, which were provided with a clinical decision. Simulated BUs tried to replicate a real BU with a high fidelity. To achieve it, these simulated BUs were organized as if they were real, involving at least the three mandatory specialties for Breast Cancer care (i.e. radiotherapist, medical oncologist, and surgeon) for making decisions regarding real clinical cases. Once this decision was made, a comparison with a previously recorded decision was studied to compare (i) the DESIREE platform performance when supporting the decision-making process and the quality of the provided decisions and (ii) the identification of non-compliance cases together with the information about the reasons that did support them. This section will focus on the second point since the first one is highly explained in (Gautier P., 2019).

For the identification of the non-compliant cases, two procedures were followed. First, all retrospective cases recorded in the platform from the 4 different clinical sites (i.e. Assistance

Publique-Hôpitaux de Paris[39] (APHP) Hôpital Européen Georges-Pompidou[40], APHP- Hôpital Saint-Louis[41], Eresa Grupo Médico[42], Onkologikoa[43]) were studied, comparing the recorded decisions with the guideline-based recommendations provided using the system. As seen in Table 8, a total of 821 decisions made in the four different clinical sites were analyzed to report non-compliance cases from retrospective data. Of all of the 821 decisions replicated using the system, 249 cases were not provided with guideline CDSS support, mainly due to missing data.

Hospital	Number of retrospective decisions	Missing recommendations
APHP- HEGP	237	74
APHP-SLH	233	95
ERE	86	20
ONK	265	60
Total	821	249

Table 8: Number of recorded retrospective decisions taking into account all clinical scenarios (i.e. A, B, C, D, defined in Chapter 4)

To compare the decisions and rate their similarity, three different levels were defined: (i) actual similarity, when two decisions were strictly identical, (ii) potential similarity, when a decision can be defined as an acceptable abstraction of the other and (iii) different decisions, when neither of the previous options were applicable.

During this validation, we focused on the third group, which was the target group of the experience-based CDSS. Every clinical site was asked to review these results to report the non-compliance causes that generated a different final decision from the CPG recommendations. We focused on the clinical cases from Onkologikoa Hospital (ONK from now on) since their own local protocol was formalized in a computer-interpretable way using the methodology presented in Chapter 3 to be used within the guideline-based CDSS and due to the proximity with the clinical center, which allowed for a deeper knowledge and understanding of the studied clinical cases. Overall, we analyzed 159 retrospective patients originating from this clinical site. From the recorded 265 clinical decisions made for the 5 different defined scenarios, only 11 of

[39] https://www.aphp.fr/

[40] https://www.aphp.fr/contenu/hopital-europeen-georges-pompidou-1

[41] https://www.aphp.fr/contenu/hopital-saint-louis

[42] https://www.eresa.com/

[43] https://www.onkologikoa.org/en/

those decisions were identified as non-compliant with the implemented Onkologikoa protocol (i.e. 4,15% non-compliant decisions).

In Table 9, the number of non-compliant decisions by scenario is specified, where 8 out of 11 non-compliant decisions were made in scenario A, 1 in scenario B and 2 in scenario D. Moreover, the followed non-compliance criteria for these cases are defined. Since a non-compliant case could have one or more non-compliance justifications, there are more criteria than non-compliant cases, as stated by the times they were reported over all non-compliant cases. In Annex 1, the list of ONK patient IDs corresponding to non-compliant decisions from retrospective cases dataset is shown.

Scenario	Number of generated experience-rules	Followed non-compliance criteria (appearing times)
A	8	Tumour non-traceable restriction (1) Patient preferences (2) Surgery: Breast and Tumor restriction (1) Genetic data (1) Age (2) Others (e.g. comorbidity scale value, clinical trial) (3)
B	1	Age (1)
D	2	Age (1) Surgery: Mastectomy Preferred (1) Patient Preferences (1)

Table 9: Number of non-compliant decisions identified for ONK retrospective patients' cases dataset depending on the scenarios

Once these cases were identified and studied, experience-based rules were generated with this knowledge. Afterward, these rules were included as part of ONK protocol rule set to be validated during the second phase with prospective patients.

5.2.2 Performance and clinical validation with prospective data in real BUs

Once some knowledge has been formalized analyzing previous clinical performance, during the second validation phase prospective patients were analyzed in simulated BUs using the DESIREE platform.

During these sessions, 12 non-compliant decisions were recorded, 4 coming from ONK and 8 from APHP-GPHE. The reported reasons were that, even if the provided recommendations were adequate, they did not fit the studied patient. Following this, some of the APHP-GPHE decisions were updated as compliant and considered as potentially similar

decisions, since the non-compliance was due to guideline inaccuracy. In Table 10, a description of the final identified prospective non-compliant cases is given.

Patient ID	Scenario	Provided recommendation	Final decision	Non-compliance criteria	CPG
10107	A	Oncology	Simple lumpectomy Axillary sentinel lymph node biopsy	Other (Uncertainty of the extension of invasion)	ONK
10154	A	Oncology	Simple lumpectomy, Axillary sentinel lymph node biopsy	Tumour Non-Traceable Restriction	ONK
10004	A	Simple lumpectomy Axillary sentinel lymph node biopsy	Letrozole	Age	ONK
10220	A	Simple lumpectomy Axillary sentinel lymph node biopsy	Letrozole	Age	ONK
40146	A	Chemotherapy	Quadrantectomy, Mastopexy, Clips on tumor bed, Axillary sentinel lymph node biopsy	Surgery: Breast And Tumor Restriction	APHP
40153	D	Radiotherapy or Endocrine therapy or Surgery (Breast re-excision for positive margins + mastectomy)	Breast re-excision for positive margins	Others (Treatment-related patient restriction)	APHP
40162	D	Endocrine Therapy	Breast RT standard fractionation Breast: 50 (Gy) / 50 fr, Boost radiation of the tumor bed Boost: 16 (Gy) / 8 fr, Anastrozole	Others (Treatment-related patient restriction)	APHP

40164	C	Endocrine Therapy	Breast RT standard fractionation Breast: 50 (Gy) / 50 fr, Boost radiation of the tumor bed Boost: 16 (Gy) / 8 fr, Infra clavicular node irradiation Breast: 45 (Gy) / 20 fr, Internal mammary node irradiation Breast: 45 (Gy) / 20 fr, Supra clavicular node irradiation Breast: 45 (Gy) / 20 fr, Anastrozol	Others	APHP
40170	D	Endocrine Therapy	Breast RT standard fractionation Breast: 50 (Gy) / 50 fr, Boost radiation of the tumor bed Boost: 16 (Gy) / 8 fr, Anastrozole	Others	APHP

Table 10: The 13 non-compliant prospective cases reported from the second clinical validation phase.

Once a non-compliant case was reported, the experience-based CDSS was queried, introducing both, the non-compliant criteria and the final decision. Using that information and the studied patient clinical data, a new rule that addressed the identified clinical need was created. An example of the rule generated for the patient with ID 10107 is given in Annex 2, following the HL7 Standard for the Clinical Decision Support Knowledge Artifact Specification[44] XML-based document explained in Chapter 3.

The final conclusion of this validation is that the different breast units find the experience-based CDSS useful since it is able to generate new knowledge to deal with those cases that are complex or restricted by external factors, such as patient preferences. Nevertheless, assessing the impact of the generated experience-based rules in a short time and with limited CPG non-compliant cases is complicated. Using the experience-based CDSS along with the guideline-based CDSS in real clinical settings for larger periods of time will allow recording a considerably higher number of compliant and non-compliant decisions, which would trigger the generation of new experience-based rules that would cover identified CPG gaps and would be validated by their applications for several patients with similar clinical

[44] http://www.hl7.org/implement/standards/product_brief.cfm?product_id=337

descriptions not covered by the CPGs, reporting the experience-based CDSS impact on daily medicine in real-time.

6. Conclusions

EBM is continuously growing, evolving and updating as more evidence and research studies' results arise. Clinicians and healthcare systems must maintain their awareness and keep track of this new knowledge in order to provide top-tier service while employing the most cutting-edge technology in clinical settings. To facilitate this task, paper-based CPGs were built, which aim to report the latest EBM, standardizing clinical practice and making it available and accessible for supporting clinicians during clinical decision-making processes by their implementation as CIGs in CDSSs. Nevertheless, some barriers that move clinicians towards a weaker CPG adherence than expected had been identified. From one side, CPG implementation, maintenance, and update drawbacks require high clinical domain knowledge and programming skills, aside from being very difficult and time-consuming tasks. On the other hand, clinicians remain hesitant about CPG´s contents, procedures definitions, and recommendations compared to traditional practices. Moreover, there are still some limitations regarding the evidence reported in the CPGs, since some clinically complex cases are outside of the performed clinical study-design, being uncovered by the CPGs and not providing any suitable recommendation and patient-preferences are still not fully integrated within the CPGs formalization process to be taken into account during the decision-making process.

The objectives and research scope of this thesis were to study how to enrich the clinical knowledge represented in the CPGs with implicit knowledge or experience that support those cases where there is no adequate recommendation to be provided. To achieve it, a methodology to retrieve and study the clinical performance, and gather the implicit knowledge from all decisions made over time for generating new experience-based rules has been presented. First, an object called the Decisional Event (DE) has been formalized, which encompasses all the information related to a clinical decision in a computerized way. This object was the backbone of all the developments presented in this thesis, providing a structure and a formalization methodology for domain-independent CIG generation, tracking clinical performance based on all the decisions made over time to generate RWD to be explored by visual analytic techniques and gathering CPG non-compliant decisions to generate new experience-based rules from them.

The domain-independent CIG formalization methodology was proposed to cope with the difficulties of CPG formalization, update and maintenance in a computerized way and to ease its implementation in any CDSS as an executable file since it relies on the DE structure

which is not constrained by any specific technology. This architecture is domain-independent since all the domain-knowledge comes from external ontologies and eases the formalization process due to an intuitive and easy-to-use interface, which directly formalizes the introduced knowledge following the DE structure methods and being translated into an executable CIG in a dynamic way.

As studied in the state-of-the-art, there are other barriers that may affect clinician's CPGs compliance, such as complex cases or shared decision-making along with the patients. In this thesis, an experience-based CDSS was presented that dealt with the limitations of a purely CPG based CDSS by identifying non-compliant cases and the reasons for biasing from guidelines' recommendations. Using this information, new rules that would augment the knowledge of the guidelines were generated. Moreover, since the DE is able to gather all the information related to the clinical decisions, including CPG compliance rates and different patient outcomes, a usability measurement computation and a proposal for measuring the strength of the experience-based rules based on the different patient outcomes have been proposed.

Moreover, to explore all the RWD gathered in the DE structure in an intuitive and interactive way, some machine learning algorithms were applied along with visual analytic techniques for new knowledge and clinical hypotheses discovery over the data using statistical reports regarding the clinical performance and PCPs for clinical information visualization given a sought-after outcome.

The proposed methodology was implemented in the Primary Breast Cancer domain and tested within the DESIREE project, in which the experience-based CDSS was technically and clinically validated.

7. Research contributions

In this chapter, all the research contributions made by the author during her research career are listed.

7.1 International Conferences

N. Muro, E. Sanchez, M. Graña, E. Carrasco, F. Manzano, J. M. Susperregi, A. Agirre, and J. Gomez. Hygehos Ontology for Electronic Health Records. 4th International Conference on Innovation in Medicine and Healthcare (InMed-16) 15-17 June 2016, Tenerife, Spain. *Best Paper Award.*

N. Larburu, **N. Muro**, I. Macía, E. Sánchez, H. Wang, J. Winder, J. Bouaud and B. Séroussi. Augmenting Guideline-based CDSS with Experts' Knowledge. 10th International Joint Conference on Biomedical Engineering Systems and Technologies, (HEALTHINF, BIOSTEC-17) 21-23 February 2017, Porto, Portugal.

N. Muro, N. Larburu, J. Bouaud, J. Belloso, G. Cajaraville, A. Urruticoechea, and B. Séroussi. Augmenting Guideline Knowledge with Non-Compliant Clinical Decisions: Experience-Based Decision Support. 5th International Conference on Innovation in Medicine and Healthcare (InMed-17) 21-23 June 2017, Algarve, Portugal.

B. Seroussi, G. Guézennec, J. B. Lamy, **N. Muro**, N. Larburu, B. Sekar and J. Bouaud. Reconciliation of multiple guidelines for decision support: a case study on the multidisciplinary management of breast cancer within the DESIREE Project. American Medical Informatics Association 2017 Annual Symposium (2017 AMIA) 4-8 November 2017, Washington D.C., USA.

N. Muro. Development of an Advanced Clinical Decision Support System: Enriching the Guideline-Based Knowledge with Experience. Doctoral Consortium of the 16th Conference on Artificial Intelligence in Medicine (AIME 2017) 21-24 June 2017, Vienna, Austria.

N. Muro, N. Larburu, J. Bouaud and B. Séroussi. Weighting Experience-Based Decision Support on the Basis of Clinical Outcomes' Assessment. EFMI STC 2017: "The

practice of patient-centered care: Empowering and engaging patients in the digital era". 22-23 October 2017, Tel Aviv, Israel.

B. Sekar, J. B. Lamy, **N. Muro**, A. Ugarriza Pinedo, B. Seroussi, N. Larburu, G. Guézennec, J. Bouaud, F. Guijarro Masero, M. Arrúe, H. Wang. Intelligent Clinical Decision Support Systems for Patient-Centered Healthcare in Breast Cancer Oncology. IEEE International Conference on e-Health Networking, Application & Services, 17-20 September 2018, Ostrava, Czech Republic.

N. Larburu, M. Arrúe, **N. Muro**, R. Álvarez, J. Kerexeta. Exploring Breast Cancer Patterns for Different Outcomes using Artificial Intelligence. IEEE International Conference on e-Health Networking, Application & Services, 17-20 September 2018, Ostrava, Czech Republic.

B. Seroussi, J. B. Lamy, **N. Muro**, N. Larburu, B. Sekar, G. Guézennec, J. Bouaud. Implementing Guideline-Based, Experience-Based, and Case-Based Approaches to Enrich Decision Support for the Management of Breast Cancer Patients in the DESIREE Project. EFMI STC 2018: "Decision Support Systems and Education – Help and Support in Healthcare", 14-26 October 2018, Zagreb, Croatia.

N. Muro, N. Larburu, J. Bouaud and B. Séroussi. Gathering real world evidence through the evaluation of decision history. 17th International Conference on Informatics, Management and Technology in Healthcare (ICIMTH), "Health Informatics Vision: From Data via Information to Knowledge" 5-7 July 2019, Athens, Greece. *Best Student Paper Award.*

N. Muro, N. Larburu, J. Torres, J. Kerexeta, G. Artola, M. Arrúe, and B. Séroussi. Architecture for a Multimodal and Domain-Independent Clinical Decision Support System Software Development Kit. 41st Engineering in Medicine and Biology Conference (IEEE EMBC) 23-27 July 2019, Berlin, Germany.

G. Artola, J. Torres, N. Larburu, R. Álvarez, and **N. Muro**. Development of a Gestational Diabetes Computer Interpretable Guideline using Semantic Web Technologies. 11th International Conference on Knowledge Engineering and Ontology Development (KEOD) 17-19 September 2019, Vienna, Austria. *Candidate to KEOD 2019 best paper award.*

7.2 Journals

N. Muro, E. Sanchez, C. Toro, E. Carrasco, S.A. Rios, F. Guijarro, and M. Graña. Experience-Based Electronic Health Records. Cybernetics and Systems, Taylor & Francis, 2016, 47, 126-139

B. Sekar, J.B. Lamy, N. Larburu, B. Séroussi, G. Guézennec, J. Bouaud, **N. Muro**, H. Wang, and J. Liu. Case-Based Decision Support System for Breast Cancer Management. International Journal of Computational Intelligence Systems, 2018, 12, 28-38, ISSN 1875-6883

7.3 Awards

The following papers were awarded as Best Research Paper and Best Student Paper:

N. Muro, E. Sanchez, M. Graña, E. Carrasco, F. Manzano, J. M. Susperregi, A. Agirre, and J. Gomez. Hygehos Ontology for Electronic Health Records. 4th International Conference on Innovation in Medicine and Healthcare (InMed-16) 15-17 June 2016, Tenerife, Spain. *Best Paper Award.*

N. Muro, N. Larburu, J. Bouaud and B. Séroussi. Gathering real world evidence through the evaluation of decision history. 17th International Conference on Informatics, Management and Technology in Healthcare (ICIMTH), "Health Informatics Vision: From Data via Information to Knowledge" 5-7 July 2019, Athens, Greece. *Best Student Paper Award.*

7.4 Intellectual Property

The software architecture presented within this thesis has been registered in the Intellectual Property registry of the Basque Government by the author under the SS-502-19 reference.

8. Discussion and Future research

The research described in this thesis can be further exploited in different directions. This section discusses some of the research challenges and lists some of the future work opportunities.

There are some interesting points discussed over the literature which are worth highlighting here since they remain still unsolved. Addressing the different compliance levels defined not only when comparing different guidelines but also for different recommendations within the same guideline is one of them. Since the defined recommendations do not achieve the lowest level of detail in their definition (i.e. defining the doses, fractions, and durations of the therapies to be provided or the specific name of the drugs) sometimes decisions that should be considered as non-compliant because of the recommendations' definition in the guidelines are not. For example, when more than one therapy is recommended to be provided but when executing the treatment the order of the therapies is permuted it can be considered as a non-compliant decision depending on the impact of this change. In this thesis all formalized guidelines were defined not only at the same level of detail, being comparable among each other, but also giving low-level information of each recommendation which has allowed to clearly identify non-compliant decisions and the reasons of non-compliance. Nevertheless, this is an ideal situation that is not applied to all CIG formalizations because of its high costs in the knowledge modeling effort. In Leonardi et al. (Leonardi et al. 2018) process mining techniques are used to compare processes on a different level of abstraction and identify branches among given recommendations, which could help to identify non-compliances and making guidelines as specific as possible. Another arisen point is related to the generation of multiple rules, coming from the experience of the clinicians or exploiting the data using data mining techniques, and how to handle and evaluate them. From one side, machine learning and data mining techniques can help in predicting the generation, fine-tuning or extension of some rules, suggesting these changes from the conclusions obtained from the data. Moreover, these techniques can be applied in some scenarios where no guideline has been formalized yet but strong knowledge and data corresponding to the opinion of leading experts is available. Another application could be clustering similar knowledge to minimize the number of generated rules and avoid including multiple redundant knowledge (Canavero et al. 2017).

Regarding some future lines that could be further developed, first, the proposed domain-independent CIG formalization can be improved to include dependencies among rules, since the proposed approach could be simplistic in some of the cases. Extending the formalization to manage rules dependencies has been studied achieving a balance between computation and representation complexity and efficiency. One of the studied methodologies was adding brackets for grouping and separating operations among conditional triplets following a given order.

Another point to ease the CIG formalization would be moving towards the support of Natural Language processing techniques for automatic structuration of CPG contents in the proposed Rule structure. Nowadays, due to the complexity of the medical vocabulary and expression variability at an individual level, this step is not possible yet, requiring human supervision from a domain expert.

Other challenges raised were the computation of different outcomes for providing a computerized SoR and QoE value automatically and updating it over time when those outcomes were updated or more information was retrieved. Even if the actual computation of the SoR remains quite subjective and some of the analyzed formalizations, such as GRADE, give some insights about how to rate this value, there is still a significant influence on the part of the clinician. Moreover, the inclusion of Patient-Reported Outcomes in an automated way along with clinical outcomes is not an easy task, due to the nature of those outcomes differing. In addition, the computation of disease-oriented and patient-oriented outcomes as a whole would be a very interesting point of view and source of information, computing these outcomes automatically still requires further research.

Limiting or reducing the data to be typified by the clinicians and the patients could be very helpful for encouraging patient outcome reporting. Many clinical variables measured by ICHOM and EORTC-QLQ questionnaires can be deduced or computed from wearables. Exploring other information sources such as Ecological Momentary Interventions (EMIs) to deliver short questionnaires to frequently report patients' psychological status and provide timely advice and feedback would help to involve the patient and make the experience more participant-friendly during the healthcare delivery evaluation.

Finally, the inclusion of the patient in the decision-making process is reaching the current clinical practice, necessitating more strategies for shared decision-making and clinical

case evaluation since the patient can contribute with their own physical status knowledge and the can clinician study this information while relying upon the most recent reported clinical evidence. In this thesis, we include some restrictions coming from the patient (i.e. non-compliance criteria due to patients' preferences) that do force biasing from CPG-based recommendations during the decision-making process. This fact is also a very interesting trend that is gaining strength within bioinformatics and the healthcare fields.

Bibliography

AGREE Collaboration. 2003. “Development and Validation of an International Appraisal Instrument for Assessing the Quality of Clinical Practice Guidelines: The AGREE Project.” *Quality & Safety in Health Care* 12 (1): 18–23. https://doi.org/10.1136/qhc.12.1.18.

Ahmadian, Leila, Ronald Cornet, and Nicolette F. de Keizer. 2010. “Facilitating Pre-Operative Assessment Guidelines Representation Using SNOMED CT.” *Journal of Biomedical Informatics* 43 (6): 883–90. https://doi.org/10.1016/j.jbi.2010.07.009.

Argüello, M., J. Des, M. J. Fernandez-Prieto, R. Perez, and H. Paniagua. 2009. “Executing Medical Guidelines on the Web: Towards Next Generation Healthcare.” In *Applications and Innovations in Intelligent Systems XVI*, edited by Tony Allen, Richard Ellis, and Miltos Petridis, 197–210. Springer London.

Artola, Garazi, Jordi Torres, Nekane Larburu, Roberto Álvarez, and Naiara Muro. 2019. “Development of a Gestational Diabetes Computer Interpretable Guideline Using Semantic Web Technologies.” In , 104–14. https://www.scitepress.org/PublicationsDetail.aspx?ID=06lGUPYo334=&t=1.

Avis, Nancy E, Edward Ip, and Kristie Long Foley. 2006. “Evaluation of the Quality of Life in Adult Cancer Survivors (QLACS) Scale for Long-Term Cancer Survivors in a Sample of Breast Cancer Survivors.” *Health and Quality of Life Outcomes* 4 (December): 92. https://doi.org/10.1186/1477-7525-4-92.

Balshem, Howard, Mark Helfand, Holger J. Schünemann, Andrew D. Oxman, Regina Kunz, Jan Brozek, Gunn E. Vist, et al. 2011. “GRADE Guidelines: 3. Rating the Quality of Evidence.” *Journal of Clinical Epidemiology* 64 (4): 401–6. https://doi.org/10.1016/j.jclinepi.2010.07.015.

Barth, Julian H., Shivani Misra, Kristin Moberg Aakre, Michel R. Langlois, Joseph Watine, Patrick J. Twomey, and Wytze P. Oosterhuis. 2016. “Why Are Clinical Practice Guidelines Not Followed?” *Clinical Chemistry and Laboratory Medicine* 54 (7): 1133–39. https://doi.org/10.1515/cclm-2015-0871.

Bastemeijer, Carla M., Lennard Voogt, Johannes P. van Ewijk, and Jan A. Hazelzet. 2017. “What Do Patient Values and Preferences Mean? A Taxonomy Based on a Systematic Review of Qualitative Papers.” *Patient Education and Counseling* 100 (5): 871–81. https://doi.org/10.1016/j.pec.2016.12.019.

Bate, Louise, Andrew Hutchinson, Jonathan Underhill, and Neal Maskrey. 2012. “How Clinical Decisions Are Made.” *British Journal of Clinical Pharmacology* 74 (4): 614–20. https://doi.org/10.1111/j.1365-2125.2012.04366.x.

Bates, David W., Gilad J. Kuperman, Samuel Wang, Tejal Gandhi, Anne Kittler, Lynn Volk, Cynthia Spurr, Ramin Khorasani, Milenko Tanasijevic, and Blackford Middleton. 2003. “Ten Commandments for Effective Clinical Decision Support: Making the Practice of Evidence-Based Medicine a Reality.” *Journal of the American Medical Informatics Association* 10 (6): 523–30. https://doi.org/10.1197/jamia.M1370.

Berner, Eta S., and Tonya J. La Lande. 2016. “Overview of Clinical Decision Support Systems.” In *Clinical Decision Support Systems*, edited by Eta S. Berner, 1–17. Health Informatics. Springer International Publishing. https://doi.org/10.1007/978-3-319-31913-1_1.

Blomqvist, Eva. 2014. “The Use of Semantic Web Technologies for Decision Support – a Survey.” *Semantic Web* 5 (3): 177–201. https://doi.org/10.3233/SW-2012-0084.

Bodenreider, Olivier. 2008. “Comparing SNOMED CT and the NCI Thesaurus through Semantic Web Technologies.” In *KR-MED*.

Bouaud, Jacques, Gilles Guézennec, and Brigitte Séroussi. 2018. “Combining the Generic Entity-Attribute-Value Model and Terminological Models into a Common Ontology to Enable Data Integration and Decision Support.” *Studies in Health Technology and Informatics* 247: 541–45.

Bouaud, Jacques, and Brigitte Seroussi. 2002. “Impact of Site-Specific Customizations on Physician Compliance with Guidelines.” *Studies in Health Technology and Informatics* 90: 543–47.

Boxwala, Aziz A, Mor Peleg, Samson Tu, Omolola Ogunyemi, Qing T Zeng, Dongwen Wang, Vimla L Patel, Robert A Greenes, and Edward H Shortliffe. 2004. “GLIF3: A Representation Format for Sharable Computer-Interpretable Clinical Practice Guidelines.” *Journal of Biomedical Informatics* 37 (3): 147–61. https://doi.org/10.1016/j.jbi.2004.04.002.

Bragaglia, Stefano, Federico Chesani, Paola Mello, and Marco Montali. 2015. “Conformance Verification of Clinical Guidelines in Presence of Computerized and Human-Enhanced Processes.” In *Foundations of Biomedical Knowledge Representation*, edited by Arjen Hommersom and Peter J. F. Lucas, 81–106. Lecture Notes in Computer Science 9521. Springer International Publishing. https://doi.org/10.1007/978-3-319-28007-3_6.

Brożek, J. L., E. A. Akl, P. Alonso-Coello, D. Lang, R. Jaeschke, J. W. Williams, B. Phillips, et al. 2009. “Grading Quality of Evidence and Strength of Recommendations in Clinical Practice Guidelines.” *Allergy* 64 (5): 669–77. https://doi.org/10.1111/j.1398-9995.2009.01973.x.

Caban, Jesus J., and David Gotz. 2015. “Visual Analytics in Healthcare – Opportunities and Research Challenges.” *Journal of the American Medical Informatics Association* 22 (2): 260–62. https://doi.org/10.1093/jamia/ocv006.

Cabana, Michael D., Cynthia S. Rand, Neil R. Powe, Albert W. Wu, Modena H. Wilson, Paul-André C. Abboud, and Haya R. Rubin. 1999. “Why Don’t Physicians Follow Clinical Practice Guidelines?: A Framework for Improvement.” *JAMA* 282 (15): 1458–65. https://doi.org/10.1001/jama.282.15.1458.

Canavero, Isabella, Anna Cavallini, Lucia Sacchi, Silvana Quaglini, Natale Arnò, Patrizia Perrone, Maria Luisa DeLodovici, Simona Marcheselli, and Giuseppe Micieli. 2017. “Safely Addressing Patients with Atrial Fibrillation to Early Anticoagulation after Acute Stroke.” *Journal of Stroke and Cerebrovascular Diseases: The Official Journal of National Stroke Association* 26 (1): 7–18. https://doi.org/10.1016/j.jstrokecerebrovasdis.2016.08.022.

Choi, Jung-Hye, Myung-Ju Ahn, Hyan-Chul Rhim, Jin-Woo Kim, Gang-Hong Lee, Young-Yeul Lee, and In-Soon Kim. 2005. “Comparison of WHO and RECIST Criteria for Response in Metastatic Colorectal Carcinoma.” *Cancer Research and Treatment : Official Journal of Korean Cancer Association* 37 (5): 290–93. https://doi.org/10.4143/crt.2005.37.5.290.

Chong, Christopher AKY, Ing-je Chen, Gary Naglie, and Murray D. Krahn. 2009. “How Well Do Guidelines Incorporate Evidence on Patient Preferences?” *Journal of General Internal Medicine* 24 (8): 977–82. https://doi.org/10.1007/s11606-009-0987-8.

Clair, Joel St. 2005. “Table 1, AHRQ Scale of Research Grades and Levels.” Text. February 2005. https://www.ncbi.nlm.nih.gov/books/NBK20542/table/A5857/.

Clercq, Paul A de, Johannes A Blom, Hendrikus H.M Korsten, and Arie Hasman. 2004. “Approaches for Creating Computer-Interpretable Guidelines That Facilitate Decision Support.” *Artificial Intelligence in Medicine* 31 (1): 1–27. https://doi.org/10.1016/j.artmed.2004.02.003.

"Common Terminology Criteria for Adverse Events (CTCAE) | Protocol Development | CTEP." n.d. Accessed January 10, 2019. https://ctep.cancer.gov/protocoldevelopment/electronic_applications/ctc.htm.
Cuzzocrea, A., and D. Zall. 2013. "Parallel Coordinates Technique in Visual Data Mining: Advantages, Disadvantages and Combinations." In *2013 17th International Conference on Information Visualisation*, 278–84. https://doi.org/10.1109/IV.2013.96.
"D2.4. Digital Patient Model and Structure." 2019. Tech. Rep.
Davis, D. A., and A. Taylor-Vaisey. 1997. "Translating Guidelines into Practice: A Systematic Review of Theoretic Concepts, Practical Experience and Research Evidence in the Adoption of Clinical Practice Guidelines." *CMAJ* 157 (4): 408–16.
"Decision Support and Information Management System for Breast Cancer | DESIREE Project | H2020." n.d. CORDIS | European Commission. Accessed August 1, 2019. https://cordis.europa.eu/project/rcn/199908/factsheet/en.
Deitel, Paul, and Harvey Deitel. 2011. *Java How to Program*. 9th ed. Upper Saddle River, NJ, USA: Prentice Hall Press.
Dueck, Amylou C., Tito R. Mendoza, Sandra A. Mitchell, Bryce B. Reeve, Kathleen M. Castro, Lauren J. Rogak, Thomas M. Atkinson, et al. 2015. "Validity and Reliability of the US National Cancer Institute's Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE)." *JAMA Oncology* 1 (8): 1051–59. https://doi.org/10.1001/jamaoncol.2015.2639.
Dykes, Patricia C., Kim Acevedo, Jodie Boldrighini, Carole Boucher, Katherine Frumento, Peggy Gray, Danielle Hall, et al. 2005. "Clinical Practice Guideline Adherence Before and After Implementation of the HEARTFELT (HEART Failure Effectiveness & Leadership Team) Intervention." *The Journal of Cardiovascular Nursing* 20 (5): 306–14.
Ebell, Mark H., Jay Siwek, Barry D. Weiss, Steven H. Woolf, Jeffrey Susman, Bernard Ewigman, and Marjorie Bowman. 2004a. "Strength of Recommendation Taxonomy (SORT): A Patient-Centered Approach to Grading Evidence in the Medical Literature." *The Journal of the American Board of Family Practice* 17 (1): 59–67. https://doi.org/10.3122/jabfm.17.1.59.
———. 2004b. "Simplifying the Language of Evidence to Improve Patient Care: Strength of Recommendation Taxonomy (SORT): A Patient-Centered Approach to Grading Evidence in the Medical Literature." Journal of Family Practice. February 1, 2004. https://link.galegroup.com/apps/doc/A113299028/AONE?sid=lms.
Eisenhauer, E.A., P. Therasse, J. Bogaerts, L.H. Schwartz, D. Sargent, R. Ford, J. Dancey, et al. 2009. "New Response Evaluation Criteria in Solid Tumours: Revised RECIST Guideline (Version 1.1)." *European Journal of Cancer* 45 (2): 228–47. https://doi.org/10.1016/j.ejca.2008.10.026.
Froberg, Debra G., and Robert L. Kane. 1989. "Methodology for Measuring Health-State Preferences—III: Population and Context Effects." *Journal of Clinical Epidemiology* 42 (6): 585–92. https://doi.org/10.1016/0895-4356(89)90155-8.
Gagnier, Joel J., Hal Morgenstern, and Patrick Kellam. 2017. "A Retrospective Cohort Study of Adverse Events in Patients Undergoing Orthopaedic Surgery." *Patient Safety in Surgery* 11 (May). https://doi.org/10.1186/s13037-017-0129-x.
Garg, Amit X., Neill K. J. Adhikari, Heather McDonald, M. Patricia Rosas-Arellano, P. J. Devereaux, Joseph Beyene, Justina Sam, and R. Brian Haynes. 2005. "Effects of Computerized Clinical Decision Support Systems on Practitioner Performance and Patient Outcomes: A Systematic Review." *JAMA* 293 (10): 1223–38. https://doi.org/10.1001/jama.293.10.1223.

Gautier P., Pelayo S., Blancafort C. 2017. “D7.1 - Design of the Validation Protocol.” Tech. Rep.
Gautier P., PelayoS., Blancafort C., Bouaud J., Belloso J., Muro N. 2019. “D7.4 - Report on DESIREE Platform Assessment.” Tech. Rep.
Ghattas, Johny, Pnina Soffer, and Mor Peleg. 2014. “Improving Business Process Decision Making Based on Past Experience.” *Decision Support Systems* 59 (March): 93–107. https://doi.org/10.1016/j.dss.2013.10.009.
Gotz, D. H., J. Sun, and N. Cao. 2012. “Multifaceted Visual Analytics for Healthcare Applications.” *IBM Journal of Research and Development* 56 (5): 6:1-6:12. https://doi.org/10.1147/JRD.2012.2199170.
Grimshaw, J. M., and I. T. Russell. 1993. “Effect of Clinical Guidelines on Medical Practice: A Systematic Review of Rigorous Evaluations.” *Lancet (London, England)* 342 (8883): 1317–22.
Grimshaw, J M, and I T Russell. 1994. “Achieving Health Gain through Clinical Guidelines II: Ensuring Guidelines Change Medical Practice.” *Quality in Health Care* 3 (1): 45–52.
Grol, Richard. 2001. “Successes and Failures in the Implementation of Evidence-Based Guidelines for Clinical Practice.” *Medical Care* 39 (8): II46–54.
Guézennec G., Bouaud J., Séroussi B., Prébet C., Muro N. 2018. “D5.1 - Design and Development of the Knowledge Model and the Clinical Guideline Model.” Tech. Rep.
Guyatt, Gordon, Andrew D. Oxman, Shahnaz Sultan, Jan Brozek, Paul Glasziou, Pablo Alonso-Coello, David Atkins, et al. 2013. “GRADE Guidelines: 11. Making an Overall Rating of Confidence in Effect Estimates for a Single Outcome and for All Outcomes.” *Journal of Clinical Epidemiology* 66 (2): 151–57. https://doi.org/10.1016/j.jclinepi.2012.01.006.
Hlatky, Mark A. 1995. “Patient Preferences and Clinical Guidelines.” *JAMA* 273 (15): 1219–20. https://doi.org/10.1001/jama.1995.03520390079037.
Holten, Danny, and Jarke J. Van Wijk. 2010. “Evaluation of Cluster Identification Performance for Different PCP Variants.” Computer Graphics Forum. June 1, 2010. https://doi.org/10.1111/j.1467-8659.2009.01666.x.
Hommersom, Arjen, and Peter J. F. Lucas. 2015. “An Introduction to Knowledge Representation and Reasoning in Healthcare.” In *Foundations of Biomedical Knowledge Representation*, edited by Arjen Hommersom and Peter J. F. Lucas, 9–32. Lecture Notes in Computer Science 9521. Springer International Publishing. https://doi.org/10.1007/978-3-319-28007-3_2.
Hunt, Dereck L., R. Brian Haynes, Steven E. Hanna, and Kristina Smith. 1998. “Effects of Computer-Based Clinical Decision Support Systems on Physician Performance and Patient Outcomes: A Systematic Review.” *JAMA* 280 (15): 1339–46. https://doi.org/10.1001/jama.280.15.1339.
Hurwitz, B. 1999. “Legal and Political Considerations of Clinical Practice Guidelines.” *BMJ (Clinical Research Ed.)* 318 (7184): 661–64. https://doi.org/10.1136/bmj.318.7184.661.
Hysong, Sylvia J., Richard G. Best, and Jacqueline A. Pugh. 2006. “Audit and Feedback and Clinical Practice Guideline Adherence: Making Feedback Actionable.” *Implementation Science: IS* 1 (April): 9. https://doi.org/10.1186/1748-5908-1-9.
“ICHOM | ICHOM Standard Sets | View Our Collection.” n.d. *ICHOM – International Consortium for Health Outcomes Measurement* (blog). Accessed May 30, 2019. https://www.ichom.org/standard-sets/.
“IEEE Recommended Practice for Software Design Descriptions.” 1998. *IEEE Std 1016-1998*, December, 1–23. https://doi.org/10.1109/IEEESTD.1998.88828.

Inselberg, Alfred, and Bernard Dimsdale. 1990. “Parallel Coordinates: A Tool for Visualizing Multi-Dimensional Geometry.” In *Proceedings of the 1st Conference on Visualization ’90*, 361–378. VIS ’90. Los Alamitos, CA, USA: IEEE Computer Society Press. http://dl.acm.org/citation.cfm?id=949531.949588.

Isern, David, and Antonio Moreno. 2008. “Computer-Based Execution of Clinical Guidelines: A Review.” *International Journal of Medical Informatics* 77 (12): 787–808. https://doi.org/10.1016/j.ijmedinf.2008.05.010.

Jeng, Judy. 2005. “What Is Usability in the Context of the Digital Library and How Can It Be Measured.” *Information Technology and Libraries* 24 (2): 47–56. https://doi.org/10.6017/ital.v24i2.3365.

Juliusson, Gunnar, Petar Antunovic, Åsa Derolf, Sören Lehmann, Lars Möllgård, Dick Stockelberg, Ulf Tidefelt, Anders Wahlin, and Martin Höglund. 2009. “Age and Acute Myeloid Leukemia: Real World Data on Decision to Treat and Outcomes from the Swedish Acute Leukemia Registry.” *Blood* 113 (18): 4179–87. https://doi.org/10.1182/blood-2008-07-172007.

Kaiser, Katharina, and Silvia Miksch. 2005. “Modeling Computer-Supported Clinical Guidelines and Protocols A Survey.” In .

Krahn, Murray, and Gary Naglie. 2008. “The Next Step in Guideline Development: Incorporating Patient Preferences.” *JAMA* 300 (4): 436–38. https://doi.org/10.1001/jama.300.4.436.

Kumar, Anand, and Barry Smith. 2005. “Oncology Ontology in the NCI Thesaurus.” In *Artificial Intelligence in Medicine*, edited by Silvia Miksch, Jim Hunter, and Elpida T. Keravnou, 213–20. Lecture Notes in Computer Science. Springer Berlin Heidelberg.

Lanzola, Giordano, Enea Parimbelli, Giuseppe Micieli, Anna Cavallini, and Silvana Quaglini. 2014. “Data Quality and Completeness in a Web Stroke Registry as the Basis for Data and Process Mining.” Research article. Journal of Healthcare Engineering. 2014. https://doi.org/10.1260/2040-2295.5.2.163.

Larburu, Nekane, Naiara Muro, Iván Macía, Eider Sánchez, Hui Wang, John Winder, Jacques Boaud, and Brigitte Séroussi. 2019. “Augmenting Guideline-Based CDSS with Experts’ Knowledge.” In , 370–76. http://www.scitepress.org/DIGITALLIBRARY/PublicationsDetail.aspx?ID=/npWlI4UqAE=&t=1.

Latoszek-Berendsen, A., H. Tange, H. J. van den Herik, and A. Hasman. 2010. “From Clinical Practice Guidelines to Computer-Interpretable Guidelines. A Literature Overview.” *Methods of Information in Medicine* 49 (6): 550–70. https://doi.org/10.3414/ME10-01-0056.

Légaré, France, and Holly O. Witteman. 2013. “Shared Decision Making: Examining Key Elements And Barriers To Adoption Into Routine Clinical Practice.” *Health Affairs* 32 (2): 276–84. https://doi.org/10.1377/hlthaff.2012.1078.

Leonardi, Giorgio, Manuel Striani, Silvana Quaglini, Anna Cavallini, and Stefania Montani. 2018. “Leveraging Semantic Labels for Multi-Level Abstraction in Medical Process Mining and Trace Comparison.” *Journal of Biomedical Informatics* 83: 10–24. https://doi.org/10.1016/j.jbi.2018.05.012.

Liem, E. B., J. S. Obeid, E. P. Shareck, L. Sato, and R. A. Greenes. 1995. “Representation of Clinical Practice Guidelines through an Interactive World-Wide-Web Interface.” *Proceedings of the Annual Symposium on Computer Application in Medical Care*, 223–27.

Lobach, D. F., and W. E. Hammond. 1997. “Computerized Decision Support Based on a Clinical Practice Guideline Improves Compliance with Care Standards.” *The*

American Journal of Medicine 102 (1): 89–98. https://doi.org/10.1016/s0002-9343(96)00382-8.

Lucas, Peter J. F., and Felipe Orihuela-Espina. 2015. "Representing Knowledge for Clinical Diagnostic Reasoning." In *Foundations of Biomedical Knowledge Representation*, edited by Arjen Hommersom and Peter J. F. Lucas, 35–45. Lecture Notes in Computer Science 9521. Springer International Publishing. https://doi.org/10.1007/978-3-319-28007-3_3.

Luckmann, Roger. 2001. "Evidence-Based Medicine: How to Practice and Teach EBM, 2nd Edition: By David L. Sackett, Sharon E. Straus, W. Scott Richardson, William Rosenberg, and R. Brian Haynes, Churchill Livingstone, 2000." *Journal of Intensive Care Medicine* 16 (3): 155–56. https://doi.org/10.1177/088506660101600307.

May, P. Hanrahan, D. A. Keim, B. Shneiderman, and S. Card. 2010. "The State of Visual Analytics: Views on What Visual Analytics Is and Where It Is Going." 2010.

Maymone, Mayra Buainain de Castro, Stephanie D. Gan, and Michael Bigby. 2014. "Evaluating the Strength of Clinical Recommendations in the Medical Literature: GRADE, SORT, and AGREE." *The Journal of Investigative Dermatology* 134 (10): 1–5. https://doi.org/10.1038/jid.2014.335.

Miguel, Jose P., David Mauricio, and Glen Rodriguez. 2014. "A Review of Software Quality Models for the Evaluation of Software Products." *International Journal of Software Engineering & Applications* 5 (6): 31–53. https://doi.org/10.5121/ijsea.2014.5603.

Miksch, Silvia. 1999. "Plan Management in the Medical Domain." *AI Commun.* 12 (4): 209–235.

Muro N., Larburu N., Kerexeta J., Arrue M. 2018. "D5.4 - Design and Development of the Experience Model, Reasoning Engine and Visualization Module." Tech. Rep.

Muro, Naiara, Eider Sanchez, Carlos Toro, Eduardo Carrasco, Sebastián A. Ríos, Frank Guijarro, and Manuel Graña. 2016. "Experience-Based Electronic Health Records." *Cybernetics and Systems* 47 (1–2): 126–39. https://doi.org/10.1080/01969722.2016.1128774.

Noble, Simon, Axel Matzdorff, Anthony Maraveyas, Majbrit V. Holm, and Giovanni Pisa. 2015. "Assessing Patients' Anticoagulation Preferences for the Treatment of Cancer-Associated Thrombosis Using Conjoint Methodology." *Haematologica* 100 (11): 1486–92. https://doi.org/10.3324/haematol.2015.127126.

Panzarasa, Silvia, Silvana Quaglini, Lucia Sacchi, Anna Cavallini, Giuseppe Micieli, and Mario Stefanelli. 2010. "Data Mining Techniques for Analyzing Stroke Care Processes." *Studies in Health Technology and Informatics* 160 (Pt 2): 939–43.

Patkar, Vivek, Dionisio Acosta, Tim Davidson, Alison Jones, John Fox, and Mohammad Keshtgar. 2011. "Cancer Multidisciplinary Team Meetings: Evidence, Challenges, and the Role of Clinical Decision Support Technology." Research article. International Journal of Breast Cancer. 2011. https://doi.org/10.4061/2011/831605.

Payne, Thomas H. 2000. "Computer Decision Support Systems." *Chest* 118 (2, Supplement): 47S-52S. https://doi.org/10.1378/chest.118.2_suppl.47S.

Peleg, Mor. 2013. "Computer-Interpretable Clinical Guidelines: A Methodological Review." *Journal of Biomedical Informatics* 46 (4): 744–63. https://doi.org/10.1016/j.jbi.2013.06.009.

Peleg, Mor, Pnina Soffer, and Johny Ghattas. 2008. "Mining Process Execution and Outcomes – Position Paper." In *Business Process Management Workshops*, edited by Arthur ter Hofstede, Boualem Benatallah, and Hye-Young Paik, 395–400. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-78238-4_40.

Pruski, Cédric, Rodrigo Bonacin, and Marcos Da Silveira. 2011. “Towards the Formalization of Guidelines Care Actions Using Patterns and Semantic Web Technologies.” In *Artificial Intelligence in Medicine*, edited by Mor Peleg, Nada Lavrač, and Carlo Combi, 302–6. Lecture Notes in Computer Science. Springer Berlin Heidelberg.

Quaglini, Silvana, Yuval Shahar, Mor Peleg, Silvia Miksch, Carlo Napolitano, Mercedes Rigla, Angels Pallàs, Enea Parimbelli, and Lucia Sacchi. 2013. “Supporting Shared Decision Making within the MobiGuide Project.” *AMIA Annual Symposium Proceedings* 2013 (November): 1175–84.

Sackett, David L., William M. C. Rosenberg, J. A. Muir Gray, R. Brian Haynes, and W. Scott Richardson. 1996. “Evidence Based Medicine: What It Is and What It Isn’t.” *BMJ* 312 (7023): 71–72. https://doi.org/10.1136/bmj.312.7023.71.

Sanchez, Eider, Wang Peng, Carlos Toro, Cesar Sanin, Manuel Graña, Edward Szczerbicki, Eduardo Carrasco, Frank Guijarro, and Luis Brualla. 2014. “Decisional DNA for Modeling and Reuse of Experiential Clinical Assessments in Breast Cancer Diagnosis and Treatment.” *Neurocomput.* 146 (C): 308–318. https://doi.org/10.1016/j.neucom.2014.06.032.

Say, Rebecca E., and Richard Thomson. 2003. “The Importance of Patient Preferences in Treatment Decisions—Challenges for Doctors.” *BMJ* 327 (7414): 542–45. https://doi.org/10.1136/bmj.327.7414.542.

Seidling, H. M., S. P. W. Schmitt, T. Bruckner, J. Kaltschmidt, M. G. Pruszydlo, C. Senger, T. Bertsche, I. Walter-Sack, and W. E. Haefeli. 2010. “Patient-Specific Electronic Decision Support Reduces Prescription of Excessive Doses.” *Quality and Safety in Health Care* 19 (5): e15–e15. https://doi.org/10.1136/qshc.2009.033175.

Séroussi, B., C. Laouénan, J. Gligorov, S. Uzan, F. Mentré, and J. Bouaud. 2013. “Which Breast Cancer Decisions Remain Non-Compliant with Guidelines despite the Use of Computerised Decision Support?” *British Journal of Cancer* 109 (5): 1147–56. https://doi.org/10.1038/bjc.2013.453.

Séroussi, Brigitte, Gilles Guézennec, Jean-Baptiste Lamy, Naiara Muro, Nekane Larburu, Booma Devi Sekar, Coralie Prebet, and Jacques Bouaud. 2017. “Reconciliation of Multiple Guidelines for Decision Support: A Case Study on the Multidisciplinary Management of Breast Cancer within the DESIREE Project.” *AMIA ... Annual Symposium Proceedings. AMIA Symposium* 2017: 1527–36.

Shafiq, Syed Imran, Cesar Sanin, Edward Szczerbicki, and Carlos Toro. 2014. “Implementing Virtual Engineering Objects (VEO) with the Set of Experience Knowledge Structure (SOEKS).” *Procedia Computer Science*, Knowledge-Based and Intelligent Information & Engineering Systems 18th Annual Conference, KES-2014 Gdynia, Poland, September 2014 Proceedings, 35 (January): 644–52. https://doi.org/10.1016/j.procs.2014.08.146.

Shapiro, Charles L., and Abram Recht. 2001. “Side Effects of Adjuvant Treatment of Breast Cancer.” *New England Journal of Medicine* 344 (26): 1997–2008. https://doi.org/10.1056/NEJM200106283442607.

Shekelle, Paul G., Steven H. Woolf, Martin Eccles, and Jeremy Grimshaw. 1999. “Clinical Guidelines: Developing Guidelines.” *BMJ : British Medical Journal* 318 (7183): 593. https://doi.org/10.1136/bmj.318.7183.593.

Sherman, Rachel E., Steven A. Anderson, Gerald J. Dal Pan, Gerry W. Gray, Thomas Gross, Nina L. Hunter, Lisa LaVange, et al. 2016. “Real-World Evidence — What Is It and What Can It Tell Us?” *New England Journal of Medicine* 375 (23): 2293–97. https://doi.org/10.1056/NEJMsb1609216.

Silberstein, SD. 2005. “Clinical Practice Guidelines.” *Cephalalgia* 25 (10): 765–66. https://doi.org/10.1111/j.1468-2982.2005.01014.x.

Sim, Ida, Paul Gorman, Robert A. Greenes, R. Brian Haynes, Bonnie Kaplan, Harold Lehmann, and Paul C. Tang. 2001. "Clinical Decision Support Systems for the Practice of Evidence-Based Medicine." *Journal of the American Medical Informatics Association* 8 (6): 527–34. https://doi.org/10.1136/jamia.2001.0080527.

Simpao, Allan F., Luis M. Ahumada, Jorge A. Gálvez, and Mohamed A. Rehman. 2014. "A Review of Analytics and Clinical Informatics in Health Care." *Journal of Medical Systems* 38 (4): 45. https://doi.org/10.1007/s10916-014-0045-x.

Sioutos, Nicholas, Sherri de Coronado, Margaret W. Haber, Frank W. Hartel, Wen-Ling Shaiu, and Lawrence W. Wright. 2007. "NCI Thesaurus: A Semantic Model Integrating Cancer-Related Clinical and Molecular Information." *Journal of Biomedical Informatics*, Bio*Medical Informatics, 40 (1): 30–43. https://doi.org/10.1016/j.jbi.2006.02.013.

Sittig, Dean F., Adam Wright, Jerome A. Osheroff, Blackford Middleton, Jonathan M. Teich, Joan S. Ash, Emily Campbell, and David W. Bates. 2008. "Grand Challenges in Clinical Decision Support." *Journal of Biomedical Informatics* 41 (2): 387–92. https://doi.org/10.1016/j.jbi.2007.09.003.

Sun, Jimeng, and Chandan K. Reddy. 2013. "Big Data Analytics for Healthcare." In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1525–1525. KDD '13. New York, NY, USA: ACM. https://doi.org/10.1145/2487575.2506178.

Sutton, David R., and John Fox. 2003. "The Syntax and Semantics of the PROforma Guideline Modeling Language." *Journal of the American Medical Informatics Association : JAMIA* 10 (5): 433–43. https://doi.org/10.1197/jamia.M1264.

Taylor, Cath, Alastair J. Munro, Rob Glynne-Jones, Clive Griffith, Paul Trevatt, Michael Richards, and Amanda J. Ramirez. 2010. "Multidisciplinary Team Working in Cancer: What Is the Evidence?" *BMJ* 340 (March): c951. https://doi.org/10.1136/bmj.c951.

Terenziani, Paolo, Stefania Montani, Alessio Bottrighi, Gianpaolo Molino, and Mauro Torchio. 2008. "Computer-Based Medical Guidelines and Protocols: A Primer and Current Trends." In .

Therasse, Patrick, Susan G. Arbuck, Elizabeth A. Eisenhauer, Jantien Wanders, Richard S. Kaplan, Larry Rubinstein, Jaap Verweij, et al. 2000. "New Guidelines to Evaluate the Response to Treatment in Solid Tumors." *JNCI: Journal of the National Cancer Institute* 92 (3): 205–16. https://doi.org/10.1093/jnci/92.3.205.

Thomas, Lois. 1999. "Clinical Practice Guidelines." *Evidence Based Nursing* 2 (2): 38–39. https://doi.org/10.1136/ebn.2.2.38.

Toussi, Massoud, Jean-Baptiste Lamy, Philippe Le Toumelin, and Alain Venot. 2009. "Using Data Mining Techniques to Explore Physicians' Therapeutic Decisions When Clinical Guidelines Do Not Provide Recommendations: Methods and Example for Type 2 Diabetes." *BMC Medical Informatics and Decision Making* 9: 28. https://doi.org/10.1186/1472-6947-9-28.

Trotti, Andy, A. Dimitrios Colevas, Ann Setser, Valerie Rusch, David Jaques, Volker Budach, Corey Langer, et al. 2003. "CTCAE v3.0: Development of a Comprehensive Grading System for the Adverse Effects of Cancer Treatment." *Seminars in Radiation Oncology* 13 (3): 176–81. https://doi.org/10.1016/S1053-4296(03)00031-6.

Tu, S. W., and M. A. Musen. 1999. "A Flexible Approach to Guideline Modeling." *Proceedings of the AMIA Symposium*, 420–24.

Tunis, Sean R. 1994. "Internists' Attitudes about Clinical Practice Guidelines." *Annals of Internal Medicine* 120 (11): 956. https://doi.org/10.7326/0003-4819-120-11-199406010-00008.

Vaitsis, Christos, Gunnar Nilsson, and Nabil Zary. 2014. “Visual Analytics in Healthcare Education: Exploring Novel Ways to Analyze and Represent Big Data in Undergraduate Medical Education.” *PeerJ* 2 (November): e683. https://doi.org/10.7717/peerj.683.

Vincent, Charles, Graham Neale, and Maria Woloshynowych. 2001. “Adverse Events in British Hospitals: Preliminary Retrospective Record Review.” *BMJ* 322 (7285): 517–19. https://doi.org/10.1136/bmj.322.7285.517.

Votruba, Peter, Silvia Miksch, and Robert Kosara. 2004. “Facilitating Knowledge Maintenance of Clinical Guidelines and Protocols.” *Studies in Health Technology and Informatics* 107 (Pt): 57–61. https://doi.org/10.3233/978-1-60750-949-3-57.

Wang, Dongwen, Mor Peleg, Samson W. Tu, Aziz A. Boxwala, Robert A. Greenes, Vimla L. Patel, and Edward H. Shortliffe. 2002. “Representation Primitives, Process Models and Patient Data in Computer-Interpretable Clinical Practice Guidelines:” *International Journal of Medical Informatics* 68 (1): 59–70. https://doi.org/10.1016/S1386-5056(02)00065-5.

Weijden, Trudy van der, France Légaré, Antoine Boivin, Jako S Burgers, Haske van Veenendaal, Anne M Stiggelbout, Marjan Faber, and Glyn Elwyn. 2010. “How to Integrate Individual Patient Values and Preferences in Clinical Practice Guidelines? A Research Protocol.” *Implementation Science : IS* 5 (February): 10. https://doi.org/10.1186/1748-5908-5-10.

West, S., V. King, T. S. Carey, K. N. Lohr, N. McKoy, S. F. Sutton, and L. Lux. 2002. *Systems to Rate the Strength of Scientific Evidence: Summary*. Agency for Healthcare Research and Quality (US). https://www.ncbi.nlm.nih.gov/books/NBK11930/.

Woolf, Steven H., Richard Grol, Allen Hutchinson, Martin Eccles, and Jeremy Grimshaw. 1999. “Potential Benefits, Limitations, and Harms of Clinical Guidelines.” *BMJ* 318 (7182): 527–30. https://doi.org/10.1136/bmj.318.7182.527.

Zegers, M., M. C. de Bruijne, C. Wagner, L. H. F. Hoonhout, R. Waaijman, M. Smits, F. a. G. Hout, et al. 2009. “Adverse Events and Potentially Preventable Deaths in Dutch Hospitals: Results of a Retrospective Patient Record Review Study.” *BMJ Quality & Safety* 18 (4): 297–302. https://doi.org/10.1136/qshc.2007.025924.

Zhou, Hong, Xiaoru Yuan, Huamin Qu, Weiwei Cui, and Baoquan Chen. 2008. “Visual Clustering in Parallel Coordinates.” *Computer Graphics Forum* 27 (3): 1047–54. https://doi.org/10.1111/j.1467-8659.2008.01241.x.

Annex 1 – List of ONK patient IDs corresponding to non-compliant decisions from retrospective cases dataset

Desims ID	Scenario	Age	Recommendation	CRITERIA	Executed Treatment
10001	D	47	Adjuvant CT Adjuvant HT RT standard fractionation Boost irradiation Lymph node irradiation.	Surgery: Mastectomy Preferred Patient preferences	Neoadjuvant CT Surgery • Bilateral Mastectomy Endocrine Therapy
10004	A	67	Surgery • Lumpectomy Sentinel node	Tumour non-traceable restriction Patient preferences	Endocrine Therapy
10017	A	51	Neoadjuvant CT	Surgery: breast and tumor restriction Genetic data	Surgery • Removal of axillary sentinel node • Tumorectomy to lesion
10023	A	77	Neoadjuvant CT	Age	Surgery • Simple lumpectomy • Axillary sentinel lymph node biopsy
10028	A	71	Surgery • Lumpectomy Sentinel node	Others (Comorbidity Scale Value)	Surgery • Tumorectomy to Lesion
10030	B	91	Surgery • Lumpectomy Sentinel node	Age	Surgery • Simple lumpectomy
10031	D	67	Adjuvant HT Breast hypofractionation	Age	Breast RT hypofractionation Endocrine Therapy

			Boost hypofractionation		
10037	A	61	Neoadjuvant HT or CT	Others	Surgery • Removal of axillary sentinel node • Tumorectomy to lesion
10058	A	70	Surgery • Lumpectomy Sentinel node	Others (Clinical Trial)	Endocrine Therapy
10062	A	78	Neoadjuvant CT	Age	Endocrine therapy
10069	A	56	Surgery • Lumpectomy Sentinel node	Patients preferences (avoid surgery)	Endocrine Therapy

Annex 2 – Experience-based rule generated for patient 10107

```
<?xml version="1.0" encoding="UTF-8"?>
<Rule id="ONK-E/_VIC-RL01" context="PatientEntity" name="The non-compliance criteria is
    ClinicalParameterRelatedClinicianRestriction with justifications equal to OtherClinicalRestriction. ">
   <Context att="" kind="global" ref="" type="PatientEntity" var="thePatient" />
   <Context att="" kind="global" ref="" type="SideEntity" var="theSide" />
   <Context att="" kind="global" ref="" type="LesionEntity" var="theLesion" />
   <Context att="hasSide" kind="alias" ref="theLesion" type="" var="theSide" />
   <Context att="hasPatient" kind="alias" ref="theSide" type="" var="thePatient" />
   <Conditional>
      <If>
         <BinaryOperator op="AND">
            <Condition op="GREATER">
               <ModelRef att="ERResult" var="theLesion" />
               <Literal type="integer">1</Literal>
            </Condition>
            <BinaryOperator op="AND">
               <Condition op="GREATER">
                  <ModelRef att="PRResult" var="theLesion" />
                  <Literal type="integer">1</Literal>
               </Condition>
               <BinaryOperator op="AND">
                  <Condition op="EQUAL">
                     <ModelRef att="PRPositivePatient" var="thePatient" />
                     <Literal type="boolean">true</Literal>
                  </Condition>
                  <BinaryOperator op="AND">
                     <Condition op="EQUAL">
                        <ModelRef att="MultipleTumoursRestriction" var="thePatient" />
                        <Literal type="boolean">false</Literal>
                     </Condition>
                     <BinaryOperator op="AND">
                        <Condition op="EQUAL">
                           <ModelRef att="RadiotherapyContraindicated" var="thePatient" />
                           <Literal type="boolean">false</Literal>
                        </Condition>
                        <BinaryOperator op="AND">
                           <Condition op="EQUAL">
                              <ModelRef att="MultifocalTumourRestriction" var="thePatient" />
                              <Literal type="boolean">false</Literal>
                           </Condition>
                           <BinaryOperator op="AND">
                              <Condition op="EQUAL">
                                 <ModelRef att="MastectomyPreferredRestriction" var="thePatient" />
                                 <Literal type="boolean">false</Literal>
                              </Condition>
                              <BinaryOperator op="AND">
                                 <Condition op="EQUAL">
                                    <ModelRef att="ExtensiveCDISRestriction" var="thePatient" />
                                    <Literal type="boolean">false</Literal>
                                 </Condition>
                                 <BinaryOperator op="AND">
                                    <Condition op="EQUAL">
                                       <ModelRef att="BreastAndTumorRestriction" var="thePatient" />
                                       <Literal type="boolean">false</Literal>
                                    </Condition>
                                    <BinaryOperator op="AND">
                                       <Condition op="EQUAL">
                                          <ModelRef att="BreastTNM.7cN" var="theSide" />
                                          <ModelRef elem="cN0" />
                                       </Condition>
                                       <BinaryOperator op="AND">
                                          <Condition op="LESS">
                                             <ModelRef att="TumorSize" var="theLesion" />
                                             <Literal type="integer">20</Literal>
                                          </Condition>
                                          <BinaryOperator op="AND">
                                             <Condition op="EQUAL">
                                                <ModelRef att="Her2NegativePatient" var="thePatient" />
                                                <Literal type="boolean">true</Literal>
                                             </Condition>
```

```
                                    <BinaryOperator op="AND">
                                        <Condition op="EQUAL">
                                            <ModelRef att="HRPositivePatient" var="thePatient" />
                                            <Literal type="boolean">true</Literal>
                                        </Condition>
                                        <BinaryOperator op="AND">
                                            <Condition op="EQUAL">
                                                <ModelRef att="BUScenario" var="thePatient" />
                                                <ModelRef elem="ScenarioA" />
                                            </Condition>
                                            <BinaryOperator op="AND">
                                                <IsSubtypeOf elem="InvasiveBreastCarcinoma">
                                                    <ModelRef att="HistologicType" var="theLesion" />
                                                </IsSubtypeOf>
                                                <BinaryOperator op="AND">
                                                    <Condition op="GREATER_OR_EQUAL">
                                                        <FunctionalExpression function="numberOf">
                                                            <ModelRef elem="theLesion" />
                                                        </FunctionalExpression>
                                                        <Literal type="integer">1</Literal>
                                                    </Condition>
                                                    <Condition op="EQUAL">
                                                        <ModelRef att="OtherClinicalRestriction" var="thePatient" />
                                                        <Literal type="boolean">true</Literal>
                                                    </Condition>
                                                </BinaryOperator>
                                            </BinaryOperator>
                                        </BinaryOperator>
                                    </BinaryOperator>
                                </BinaryOperator>
                            </BinaryOperator>
                        </BinaryOperator>
                    </BinaryOperator>
                </BinaryOperator>
            </BinaryOperator>
        </BinaryOperator>
        <Then>
            <Seq>
                <BuildContainer elem="RecoEntity">
                    <Seq>
                        <BuildObject conformance="SHOULD" elem="OrderEntity" evidenceLevel="exp" step="">
                            <BuildContent boostDose="" breastDose="" cycles="" duration="" durationUnit="" fractions="" fractionsPeriodicity="">
                                <ModelRef elem="Lumpectomy" />
                            </BuildContent>
                        </BuildObject>
                    </Seq>
                </BuildContainer>
                <BuildContainer elem="RecoEntity">
                    <Seq>
                        <BuildObject conformance="SHOULD" elem="OrderEntity" evidenceLevel="exp" step="">
                            <BuildContent boostDose="" breastDose="" cycles="" duration="" durationUnit="" fractions="" fractionsPeriodicity="">
                                <ModelRef elem="SentinelAxillaryLymphNodeBiopsy" />
                            </BuildContent>
                        </BuildObject>
                    </Seq>
                </BuildContainer>
            </Seq>
        </Then>
    </If>
</Conditional>
</Rule>
```

Table of figures

Figure 1: Approach diagram. 6
Figure 2: An example of a parallel coordinates plot representation of the Iris Dataset. 20
Figure 3: Eclipse editor overview with all the defined packages. The classes under the *decisionalEventObjects* package, and that will be described in this chapter, are deployed on the left part of the application. 28
Figure 4: The *DecisionalEvent* Java class and its attributes declaration. 30
Figure 5: The *NonComplianceCriteria* Java class and its attributes declaration. 30
Figure 6: The *Outcomes* Java class and its attributes declaration. 31
Figure 7: The *Guideline*, *RuleGroup,* and *Rule* Java classes and their attributes declaration. 33
Figure 8: The *ContextEntity* Java class and its attributes declaration. 34
Figure 9: The *ConditionTriplet* Java class and its attributes declaration. 35
Figure 10: The *Condition* Java class and its attributes declaration. 35
Figure 11: The *Recommendation* Java class and its attributes declaration. 36
Figure 12: The *Order* java class and its attributes declaration. The first two attribute classes are defined for the specific clinical domain of Breast Cancer. 37
Figure 13. The relational view of the database containing all the information related to the decision-making process. 39
Figure 14: The relational view of the database containing all the clinical data related to the patient following the FHIR messaging definition and structuration. 40
Figure 15. A simple example of a complete Condition Triplet instance, composed of a precedent condition, a binary operator and a consequent condition, following the proposed CIG formalization. 41
Figure 16: Screenshot of the DRL language-based rule template, used for CIG formalization. 44
Figure 17: The authoring tool GUI for the definition of new rules. 47
Figure 18: Example of the possible values that a variable within a condition of a rule could take according to the knowledge in gestational diabetes mellitus domain modeled in an ontology. 48
Figure 19: Example of the definition of a condition within a rule, which states that a patient's age must be greater than 50 years old. 48
Figure 20: Example of a second condition definition linked by the “AND” binary operator. . 49
Figure 21: Example of a simplified rule for non-metastatic breast cancer with infiltrating tumor management. 52
Figure 22: Summary of the representation of the data that composes a Decisional Event for Patient 1 in a time t=0. The non-compliance related variables are highlighted in red. 53
Figure 23: Example of a simplified rule for non-metastatic breast cancer with infiltrating tumor management, matching the Final Decision made by the BU for Patient 1. The conflictive antecedent is highlighted in orange since Tumor Size >20 ϵ CR and Tumor Size ≤20 ϵ CR’. 54
Figure 24: The new experience-based rule generated to cope with this CPG non-compliant clinical case 55
Figure 25: The representation of the breast cancer standard set of ICHOM. 60
Figure 26: The strength of the recommendation computation approach based on GRADE using the toxicities, the treatment response, and usability values. 62
Figure 27: The Decisional Events history visualization dashboard. 65
Figure 28: A zoomed overview of all graphs for the specific cases treated with chemotherapy treatments. 66

Figure 29: The Parallel Coordinates Plot (PCP) for outcome expectancy showing the results of 67 patients. 68
Figure 30: The Parallel Coordinates Plot (PCP) for outcome expectancy when applying some ranges restrictions among variables to highlight patients with higher toxicity grades. 69
Figure 31: Example of the “Histologic Type” class definition within the DESIREE BCKM. 73
Figure 32: NRL representation of a simple rule. 75
Figure 33: DGDL representation of a simple rule. 76
Figure 34: Non-compliance justifications options provided to BUs during the decision-making process. 77
Figure 35: System architecture with all modules and communications. 91

Table of tables

Table 1: Description of the tables composing the relational database for storing all the information related to the decision-making process. 38
Table 2: Set of clinical parameters values defining the Patient 1 and Patient 2. 51
Table 3: The selected ISO/ IEC CD 25010: 2007 model characteristics to be tested for the software quality evaluation. 82
Table 4: Classification of all of the subsystems or modules of the experience-based CDSS. . 84
Table 5: Table describing the measured features of each of the functionalities of the experience-based CDSS for each performed unit test. 90
Table 6: Definition of all integrations and communications exchanges within the system. 94
Table 7: Other interactions defined for the correct functioning of the system 94
Table 8: Number of recorded retrospective decisions taking into account all clinical scenarios (i.e. A, B, C, D, defined in Chapter 4) 96
Table 9: Number of non-compliant decisions identified for ONK retrospective patients' cases dataset depending on the scenarios 97
Table 10: The 13 non-compliant prospective cases reported from the second clinical validation phase. 99

MURO AMUCHASTEGUI, Naiara – Thèse de doctorat - 2019

Développement d'un système avancé d'aide à la décision clinique : enrichir la connaissance issue des guides de pratique clinique avec l'expérience

Résumé :

La médecine fondée sur les preuves a permis de formaliser des guides de pratique clinique qui définissent des flux de travail et des recommandations à suivre pour un domaine clinique concis. Ces guides se sont construits dans le but de standardiser les soins de santé et d'obtenir les meilleurs résultats possibles pour les patients. Néanmoins, les médecins n'adhèrent pas toujours à ces directives en raison de diverses limitations cliniques et de mise-en-œuvre. D'une part, les médecins n'ont pas toujours familiarisés ou en accord avec les lignes directrices des guides de pratique clinique, doutant ainsi de leur efficacité et des résultats attendus par rapport aux pratiques antérieures. D'autre part, maintenir ces guides à jour en incluant les dernières preuves établies requiert une gestion continue d'une documentation établie sur support papier. Les systèmes d'aide à la décision clinique sont ainsi proposés comme aide durant le processus de prise de décision clinique, par la mise en œuvre informatisée des guides pour promouvoir leur consultation et l'adhésion des médecins. Bien que ces systèmes aident à améliorer la conformité des guides, il subsiste certains obstacles hérités des guides sur support papier qui ne sont pas résolus avec leur mise en œuvre informatisée, comme le traitement des cas complexes non-définis dans les directives ou le manque de représentation d'autres facteurs externes qui peuvent influer sur les traitements fournis et faire dévier des recommandations des guides (c.-à-d. les préférences du patient).

La présente thèse propose un système avancé d'aide à la décision clinique pour faire face aux limitations du soutien purement basé en guides et aller au-delà des connaissances formalisées en analysant les données cliniques, les résultats et les performances de toutes les décisions prises au fil du temps. Pour atteindre ces objectifs, une approche de modélisation des connaissances et performances cliniques de manière sémantique validée et informatisée a été présentée, en s'appuyant sur une ontologie et avec la formalisation du concept d'Événement Décisionnel. De plus, un cadre indépendant du domaine a été mis en place pour faciliter le processus d'informatisation, de mise à jour et de mise en œuvre des guides de pratique clinique au sein d'un système d'aide à la décision clinique afin de fournir un soutien clinique à pour chaque patient interrogé. Pour répondre aux limites des guides, une méthodologie permettant d'augmenter les connaissances cliniques en utilisant l'expérience a été présentée ainsi qu'une évaluation de la performance clinique et de la qualité au fil du temps, en fonction des différents résultats cliniques étudiés, tels que l'utilisabilité et la fiabilité clinique derrière les connaissances cliniques formalisées. Enfin, les données du monde réel accumulées ont été explorées pour soutenir les cas futurs, promouvoir l'étude de nouvelles hypothèses cliniques et aider à la détection des tendances et des modèles sur les données à l'aide d'outils d'analyse visuelle.

Les modules présentés ont été développés et mis en œuvre dans leur majorité dans le cadre du projet européen Horizon 2020 DESIREE, dans lequel le cas d'utilisation était axé sur le soutien des unités de soins du sein au cours du processus décisionnel pour la prise en charge des patientes atteintes d'un cancer du sein primaire, en effectuant une validation technique et clinique sur l'architecture présentée, dont les résultats sont présentés dans cette thèse. Néanmoins, certains des modules ont également été utilisés dans d'autres domaines médicaux tels que le développement des guides de pratique clinique pour le diabète gestationnel, mettant en évidence l'interopérabilité et la flexibilité du travail présenté.

Mots clés : guides de pratique clinique, Événement Décisionnel, systèmes d'aide à la décision clinique, formalisation de l'expérience, techniques d'analyse visuelle, données du monde réel.

Development of an advanced clinical decision support system: enriching the guideline-based knowledge with experience

Abstract:

Evidence-Based Medicine has been formalized as Clinical Practice Guidelines, which define workflows and recommendations to be followed for a given clinical domain. These documents were formalized aiming to standardize healthcare and seeking the best patient outcomes. Nevertheless, clinicians do not adhere as expected to these guidelines due to several clinical and implementation limitations. On one hand, clinicians do not feel familiar, agree with and or are unaware of guidelines, hence doubting their self-efficacy and outcome expectancy compared to previous or more common practices. On the other hand, maintaining these guidelines updated with the most recent evidence requires continuous versioning of these paper-based documents. Clinical Decision Support Systems are proposed to help during the clinical decision-making process with the computerized implementation of the guidelines to promote their easy consultation and increased compliance. Even if these systems help improving guideline compliance, there are still some barriers inherited from paper-based guidelines that are not solved, such as managing complex cases not defined within the guidelines or the lack of representation of other external factors that may influence the provided treatments, biasing from guidelines' recommendations (i.e. patient preferences). Retrieving observational data and patients' quality of life outcomes related to the provided healthcare during routine clinical practice could help to identify and overcome these limitations and would generate Real World Data representing the real population and going beyond the limitations of the knowledge reported in the Randomized Clinical Trials.

This thesis proposes an advanced Clinical Decision Support System for coping with the purely guideline-based support limitations and going beyond the formalized knowledge by analyzing the clinical data, outcomes, and performance of all the decisions made over time. To achieve these objectives, an approach for modeling the clinical knowledge and performance in a semantically validated and computerized way has been presented, leaning on an ontology and the formalization of the Decisional Event concept. Moreover, a domain-independent framework has been implemented for easing the process of computerizing, updating and implementing Clinical Practice Guidelines within a Clinical Decision Support System in order to provide clinical support for any queried patient. For addressing the reported guideline limitations, a methodology for augmenting the clinical knowledge using experience has been presented along with some clinical performance and quality evaluation over time, based on different studied clinical outcomes, such as the usability and the strength of the rules for evaluating the clinical reliability behind the formalized clinical knowledge. Finally, the accumulated Real World Data was explored to support future cases, promoting the study of new clinical hypotheses and helping in the detection of trends and patterns over the data using visual analytics tools.

The presented modules had been developed and implemented in their majority within the European Horizon 2020 project DESIREE, in which the use case was focused on supporting Breast Units during the decision-making process for Primary Breast Cancer patients management, performing a technical and clinical validation over the presented architecture, whose results are presented in this thesis. Nevertheless, some of the modules have been also used in other medical domains such as Gestational Diabetes guidelines development, highlighting the interoperability and flexibility of the presented work.

Keywords: Clinical Practice Guidelines, Decisional Event, Clinical Decision Support Systems, Experience formalization, Visual Analytics, Real World Data.