USE OF DECISION TABLES TO MODEL ASSISTANCE KNOWLEDGE TO TRAIN MEDICAL RESIDENTS

Francis Real Vázquez

Francis Real Vázquez

Use of Decision Tables to Model Assistance Knowledge to Train Medical Residents

PhD Thesis
Supervised by Dr. David Riaño

DEPARTAMENT D'ENGINYERIA INFORMÀTICA I MATEMÀTICA
February 2015

UNIVERSITAT
ROVIRA i VIRGILI

UNIVERSITAT ROVIRA I VIRGILI
USE OF DECISION TABLES TO MODEL ASSISTANCE KNOWLEDGE TO TRAIN MEDICAL RESIDENTS
Francis Real Vázquez

UNIVERSITAT ROVIRA I VIRGILI

Departament d'Enginyeria Informàtica i Matemàtiques
Avinguda Països Catalans, 26
43007, Tarragona.

I STATE that the present study, entitled USE OF DECISION TABLES TO MODEL ASSISTANCE KNOWLEDGE TO TRAIN MEDICAL RESIDENTS, presented by FRANCIS REAL VAZQUEZ for the award of the degree of Doctor, has been carried out under my supervision at the Department DEPARTAMENT D'ENGINYERIA INFORMÀTICA I MATEMÀTIQUES of this university.

Tarragona, 10th January 2016.

The supervisor,

Dr. David Riaño Ramos

UNIVERSITAT ROVIRA I VIRGILI
USE OF DECISION TABLES TO MODEL ASSISTANCE KNOWLEDGE TO TRAIN MEDICAL RESIDENTS
Francis Real Vázquez

"And once the storm is over,
you won't remember how you made it through,
how you managed to survive.
You won't even be sure, whether the storm is really over.
But one thing is certain.
When you come out of the storm,
you won't be the same person who walked in.
That's what this storm's all about."

- Haruki Murakami

UNIVERSITAT ROVIRA I VIRGILI
USE OF DECISION TABLES TO MODEL ASSISTANCE KNOWLEDGE TO TRAIN MEDICAL RESIDENTS
Francis Real Vázquez

Acknowledgements

This thesis wouldn't have been possible without the aid and collaboration of people whom I wish to thank.

The hardest and longest task has been on David Riaño's shoulders. He has been always available, with an accurate advice or an encouraging word. His devotion and patience has gone far beyond what one expects from an advisor to get to what one expects from a friend.

I want to especially thanks José-Ramón Alonso, J, from Hospital Clinic of Barcelona, for his help and dedication. He provided us a new focus to this work when we were lost, and support us with his expertise and energy. Without him, this work would never have done.

I also want to thanks Antonio Collado, from SAGESSA, for his assistance in the first steps of this work.

I am particularly grateful to Pepi Vázquez, for her help to find medical bibliography.

As for the personal gratitude, I want to thanks my familiars and friends for their support and patience. I especially want to mention to my uncle Bernardo Vázquez, for his guidance, inspiration and support.

To my beloved Susana for her enthusiasm in the good moments and her encouragement and support in the hard ones. During all the time that this thesis lasted, you believed in me more than I believed in myself. This thesis is dedicated to you.

And, finally, to my new huge project, Aimund. You don't know now, but you gave the final push to finish it.

UNIVERSITAT ROVIRA I VIRGILI
USE OF DECISION TABLES TO MODEL ASSISTANCE KNOWLEDGE TO TRAIN MEDICAL RESIDENTS
Francis Real Vázquez

Contents

UNIVERSITAT ROVIRA I VIRGILI
USE OF DECISION TABLES TO MODEL ASSISTANCE KNOWLEDGE TO TRAIN MEDICAL RESIDENTS
Francis Real Vázquez

List of Figures

List of Tables

UNIVERSITAT ROVIRA I VIRGILI
USE OF DECISION TABLES TO MODEL ASSISTANCE KNOWLEDGE TO TRAIN MEDICAL RESIDENTS
Francis Real Vázquez

Abbreviations

AH	**A**rterial **H**ypertension
BP	(arterial) **B**lood **P**ressure
CA	**C**linical **A**lgorithm
CAT	**C**omputerized **A**xial **T**omography
CIG	**C**omputer **I**nterpretable **G**uideline
CPG	**C**linical **P**ractice **G**uideline
DBP	**D**iastolic **B**lood **P**ressure
DDx	**D**ifferential **D**iagnosis
DHDT	**D**iagnostic **H**ypotheses **D**ecision **T**able
DT	**D**ecision **T**able
EDT	**E**valuation **D**ecision **T**able
EHR	**E**lectronic **H**ealth **R**ecord
EM	**E**valuation **M**odule
GDT	**G**rouping **D**ecision **T**able
GP	**G**eneral **P**ractitioner
HIT	**H**ealth **I**nformation **T**echnology
ICU	**I**ntensive **C**are **U**nit
IM	**I**nterface **M**odule
IT	**I**nformation **T**echnology
MP	**M**edical **P**ractice
MPM	**M**edical **P**ractice **M**odel
PAM	**P**atient **A**nalysis **M**odule
OQM	**O**n-line **Q**uestionnaire **M**odule
PCM	**P**atient **C**reator **M**odule
PDT	**P**rognosis **D**ecision **T**able

PM	**P**atient **M**odel
RAM	**R**esult **A**nalysis **M**odule
RCT	**R**andomized **C**ontrol **T**rial
SBP	**S**ystolic **B**lood **P**ressure
SM	**S**imulation **M**odule
TDT	**T**reatment **D**ecision **T**able
TM	**T**reatment **M**odule
TSDT	**T**est **S**election **D**ecision **T**able

Chapter 1

Introduction

From its origins, one of the basic objectives of computer science has been to support the work of professionals in all areas. Medicine is one of the fields where computer science has received particular interest. The first works in this area appeared in the 1950s, and medical informatics was identified as a new speciality in the 1970s [1].

One of the main goals of medical informatics has been to provide physicians with tools to improve the quality of health care.

Every day, physicians face different tasks. Some of them are routine tasks, such as confirming roll calls, recording patient data, or visiting patients, and some other ones are cognitive tasks, such as diagnosing, proposing a therapy, or prognosticating the evolution of a patient.

Medicine is a living science in which new discoveries about diseases and drugs are continuous. Consequently, new protocols are generated while previous ones require continuous revision and adaption to the new realities. All these changes are not only a response to new discoveries, but they are also caused by social and economic changes [2]. All this implies that physicians need to be constantly trained in their clinical tasks.

Medical knowledge uses to be published as Clinical Practice Guidelines (CPG) [3]. These guidelines have been defined by the Institute of Medicine in the US as "statements that include recommendations intended to optimize patient care that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options". In practice, they are textual documents created by experts to gather all the available evidence usually with regard to a particular medical problem or disease. Evidence implies that the document or the contained recommendations have been created using an unbiased and transparent process of systematically reviewing, appraising,

and using the best clinical research findings of the highest value to aid in the delivery of optimum clinical care to patients.

Their usual focus in a single disease implies that the knowledge required by health care professionals to perform clinical practice is disseminated across multiple publications. This puts these professionals in some difficulties at the time to manage all the available knowledge, and also to keep them abreast of updates and new publications.

In order to properly address this issue, since the 1970s different information technology (IT) systems have appeared to support health care. Two of these computer-based systems that are related to this thesis are clinical decision support systems, and medical training systems.

Clinical decision support systems [4, 5] are designed to ensure the homogeneity of medical interventions, optimize the quality of care and, very often, contain the medical cost. On the other hand, medical training systems [6] aim to create new dynamics of actuation, reinforce them, and modify the existing ones for health care professionals to get adapted to new realities.

Nowadays, most of these computer-based systems are oriented to deal with someone of the previously mentioned cognitive tasks, mainly diagnosis or treatment, but few with all the tasks involved in the medical process [7–9] but none, to our knowledge, using a homogeneous technology for all these tasks in an integrated way.

Here, we define the medical process as all the cognitive medical decisions that a physician has to make since the moment a patient arrives to a health care center till the moment she is discharged. This time interval can be either short-term (e.g., critical interventions) or long-term (e.g., chronic or severe interventions). Several health care cognitive problems may occur along the medical process. These cognitive problems use to be solved when they appear, one by one. Solutions are obtained after the use the technologies that best fit each problem. So, for example, a diagnostic problem can be addressed with expert systems [10] or Bayesian networks [11] and a therapy prescription with computer-interpretable guidelines [12] or Support Vector Machines [13].

In this thesis, the medical process comprise three main tasks:

- **Diagnosis:** Medical diagnosis is defined as the identification of a disease or group of diseases by investigating the signs, the symptoms, and the history of a patient. Diagnosis provides a solid basis for the assignment of a treatment and the prognosis of the evolution of a patient [14].

- **Treatment:** A treatment is a set of actions of different kinds (e.g., surgical, physiological, pharmacological, advisory, etc.) used to relieve or cure a patient from the diseases detected after a diagnosis. The notion of treatment is often used as a synonym for therapy or intervention.

- **Prognosis:** In medicine, prognosis is the act of predicting the probable course and outcome of a disease [15, 16]. Physicians use prognosis to answer questions such as: is a patient going to improve?, what is her chance of recovery?, and how likely a relapse is?. In this work, prognosis is related to the skill of physicians to foresee the possible consequences of a treatment in the evolution of the health condition of a patient.

Although health information technology (HIT) researchers have worked in numerous fields of medicine showing the potential of these technologies to transform the delivery of health care, increasing safety, effectiveness, and efficiency, many of these studies are usually based on controlled tests [17]. Other studies exist that show the benefits of HIT to increase the adherence to guideline-based care, enhance surveillance and monitoring, and decrease medication errors [18]. All these technologies use to be build around an electronic health record (EHR) whose benefits in real medical practice has been studied [19–21], in spite that some problems still persist [22]. An EHR is defined as a systematic storage of health information about one patient in a digital format that can be shared across different health care settings. It may include heterogeneous data such as demographics, medical history, medication and allergies, immunization status, laboratory test results, radiology images, vital signs, personal statistics like age and weight, and billing information about the patient along time. [23, 24].

In routine and administrative tasks, HIT and EHRs have allowed to optimize the management of patients [25]. However, in cognitive tasks requiring HIT to support intelligent decision making, systems do not provide enough evidence of improvement [26–28]. In addition, the incorporation of HIT to health care can modify the way that medical responsibility must be applied. So for EHR, clinicians responsibilities rely on the capacities of the EHR used, that Sittig et al [29] classified into 10 key points. But for cognitive HIT, liability is difficult to set up [30] or transfer to computer systems [31].

Physicians rely on their autonomy and authority to make decisions. Even in cases where a decision support system assist their choices, they may not necessarily want to leverage this technology, because they may perceive certain technological advances as a challenge and threat to their authority, especially if they can lose the control over the conditions, processes, procedures, or content of their care work [32].

Statistical systems, often used as clinical decision support systems can obtain a very good accuracy [33], but their results use to lack of a satisfactory justification. The level of knowledge codification is an important variable that has a significant direct negative influence on the perceived usefulness of IT and on the intention to use IT products [34].

With this in mind, among the medical knowledge-based systems, we select decision table based systems because they are simple, easy to understand and well-known in medicine.

Decision tables (DTs) [35] are knowledge structures that represent knowledge as cohesive rule sets, where a rule is an *IF antecedent THEN consequent* expression. In DTs, the columns represent the rules, and the rows represent either the conditions in the antecedent or the actions in the consequent. DTs are widely used in software engineering for documenting and specifying complex decisions in a simple way, which is easy to check for consistency, completeness, and correctness [36]. In medicine, they facilitate knowledge representation, validation, maintenance, and explanation [37] and have been successfully applied to diagnose single diseases [38, 39] and to train physicians in diagnosis [40], among other medical cognitive tasks [41].

The knowledge contained in the DTs may be obtained from CPGs [42]. CPGs contain both explicit and implicit rules embedded in the text. These rules have to be extracted and translated into the DTs by experts or by automatic systems.

1.1 Objectives

The objectives of this thesis are:

1. **Integrate the whole medical process into a knowledge-based model**

   Design a knowledge representation model that allows us to homogeneously integrate knowledge structures to face the three tasks of the medical process; namely, diagnosis, treatment, and prognosis.

   To this end, we propose the construction of a functional clinical practice model, based on the medical practice model (MPM) introduced in [7]. This new model will address diagnosis, treatment, and prognosis and it will support decision making in the mentioned cognitive medical tasks.

2. **Confirm that decision tables are valid knowledge representation structures to support the medical process**

   We propose the design and the later use of specific-purpose decision tables adapted to the medical tasks of diagnosis, treatment, and prognosis. In order to validate the capability of these structures to host medical knowledge, they will be used to contain the knowledge extracted from several CPGs to support physicians in two concrete medical tasks: the diagnosis of secondary causes of hypertension, and the treatment and prognosis of shocks in the emergency unit of a tertiary hospital.

3. **Develop a clinical training system based in decision tables**

   When implemented with DTs, the clinical practice model achieved in the first objective could provide different services, as for example, decision support, supervision of medical behavior, or clinical training.

   Among these applications, we aimed to develop two frameworks for clinical training. The first framework uses the DTs about the diagnosis of secondary causes of hypertension and implements a case-based learning tool for general practitioners (GPs) to improve their skills in differential diagnosis. The second framework uses DTs about the treatment of seven different common shocks in an emergency unit. In this same framework, DTs about prognosis are used to simulate the evolution of virtual patients while the shock treatment is provided.

4. **Assess the benefits of the training frameworks in a hospital with novice physicians**

   To validate the utility of the training frameworks, we wanted to test both systems with medical residents from the *Hospital Clínic de Barcelona*. The results were evaluated for clinical quality assessment. For the first test, a pre-post study and the statistical significance of the result were done. For the second test, an experiment with a control group was conducted.

1.2 Overview

The organization of the rest of the thesis is presented in this section. After a state of the art in chapter 2, chapters 3 and 4 present the design of decision tables for diagnostic, treatment and prognostic. Chapters 5 and 6 describe the knowledge acquisition. Chapters 7 and 8 present two experiments of medical training with decision tables. Finally, chapter 9 exposes the conclusions yielded by this work and outlines some further research lines.

- **Chapter 2: Background**

  In this chapter we define the health care concepts in which this thesis are based: diagnostic, treatment, prognosis and clinical practice guidelines.

  We also overview the current trends in health information technologies related to this thesis, especially on decision tables, since it is the main issue in this work.

- **Chapter 3: Implementation of Differential Diagnosis with Decision Tables**

  In this chapter we detail the stages and necessities of differential diagnosis. In addition, we propose a design for the decision tables oriented to each one of the stages that define a differential diagnosis.

- **Chapter 4: Implementation of Treatment and Prognosis with Decision Tables**

  In this chapter we describe our approach to treatment and prognosis. We study how decision tables may be useful to describe both stages of the medical process and we present our decision table structure.

- **Chapter 5: Use of Diagnostic Decision Tables in Secondary Causes of Hypertension**

  In this chapter we present the diagnostic of the eight most common secondary causes of hypertension as a case of study. We detail the adaptation of their clinical practice guidelines to Decision Tables.

  In appendix A we describe the eight diseases considered in this experiment and we show the decision tables obtained for the diagnosis of secondary causes of hypertension.

- **Chapter 6: Use of Treatment and Prognosis Decision Tables in Emergency Shock**

  In this chapter we describe the treatment of seven types of emergency shock. We create the decision tables with the knowledge about each shock in base of clinical practice guidelines. In addition, we present how obtain the prognosis tables from expert physicians.

  In appendix B we describe the seven shocks considered in this work and we show the decision tables obtained for the treatment of emergency shock.

- **Chapter 7: Using Decision Tables to Train Residents**

  In chapter 7 we present the design of a training experiment based in the diagnostic decision tables of secondary causes of hypertension detailed in chapter 5. We detail how it was carried out as part of the training program of residents in the Hospital Clínic de Barcelona.

  Finally, we analyze and discuss the results and formative benefits obtained.

- **Chapter 8: Use of Treatment Decision Tables to Train Residents in Emergency Shock**

  In this chapter, we detail a second experiment based on the treatment of shock. This experiment includes the tables obtained in chapter 6 for treatment and prognosis. We describe how the experiment was carried out and the results obtained.

- **Chapter 9: Conclusions**

  In this chapter we summarize the research described in this thesis and we outline some future lines of research to extend the applicability and the performance of our systems.

UNIVERSITAT ROVIRA I VIRGILI
USE OF DECISION TABLES TO MODEL ASSISTANCE KNOWLEDGE TO TRAIN MEDICAL RESIDENTS
Francis Real Vázquez

Chapter 2

Background

Clinical practice, or the practice of medicine in a clinical setting, is about the combination of health care actions by a group of professionals and caregivers in order to assist a patient with her diseases and ailments. The most relevant tasks in clinical practice correspond to the diagnosis of the diseases affecting the patient, the proposal of the best treatment possible in order to prevent, cure, or palliate the diseases and ailments of the patient, and the prognosis of the evolution of the health parameters of patient. All of them are knowledge intensive tasks that require thorough and continuous training according to the constant changes and evolution of medicine.

In order to support clinicians in this complex issue, several computer-based technologies have been proposed. Closely related to this thesis, some of these technologies are clinical practice modeling, differential diagnostic generators, expert systems, and decision tables.

This chapter is about the description of clinical practice, from a medical point of view, and the presentation of some computer-based technologies to deal with clinical practice, in relation to this thesis.

2.1 Clinical Practice

In modern medicine, physicians meet patients in order to diagnose, treat, or prevent disease using their expert clinical judgment. Some of these actions are taken according to the expected evolution of the patient (i.e., prognosis). An episode of care is defined as the sequence of clinical actions performed on a patient during a short or long time period in order to solve a medical problem affecting that patient. A typical episode of care in clinical practice is composed of a set of visits or encounters between the patient an one or several health care professionals. It begins with the examination of the health

care record of the patient, followed by a medical interview and a physical examination of the signs and symptoms manifested in the patient. The doctor may order some medical tests in order to complement the information required to conclude with a diagnosis of the disease or set of diseases affecting the patient. At this point a therapy uses to be prescribed, considering the patient current condition and the expected evolution or patient prognostic. Each medical encounter is documented in the health care record of the patient for later consultations. A follow-up process may be started when the condition of the patient is stable but her cure is not confirmed.

The main tasks performed along clinical practice episodes of care deserve a detail consideration in the following sections.

2.1.1 Diagnosis

Diagnosis [43] is a complex process that begins with the identification of signs and symptoms of the patient and it culminates with the categorization of her illness or group of illnesses. The outcome of this process provides a solid basis for the treatment and the prognosis of patients [14].

Physicians can use several types of reasoning during the diagnostic process [44]. These are: pattern recognition, algorithmic, exhaustive, and hypothetico-deductive.

Pattern recognition is often used to diagnose conditions that have a very distinctive presentation. It entails the instant recognition of a disease, for instance the diagnosis of Down's syndrome after direct observation of the patient, that could be combined with a genetic test. This is a very efficient type of diagnosis and it is often used in busy clinical settings. The risk of using pattern recognition is that you may jump prematurely to a diagnosis without considering all possible options, or at least other concerning options.

The **Algorithmic** approach [45] is based on the use of flowcharts and clinical algorithms to determine the patient's diagnosis. The American College of Physicians define clinical algorithms as schematic models of the clinical decision pathways described in a guideline [46, 47]. The algorithmic approach to diagnosis is useful for health conditions where the information collected from patients is precise and reproducible, such as the results of a blood test or auscultation. For example, the diagnosis of chronic hypertension can be agreed after the observation of a high blood pressure in two (or three) consecutive measurements taken in intervals of 6-12 minutes.

The **Exhaustive** approach relies on gathering every possible piece of information to make the diagnosis, ordering every test the physician can think of that might provide relevant information on a patient's condition. It is only recommended for patients with

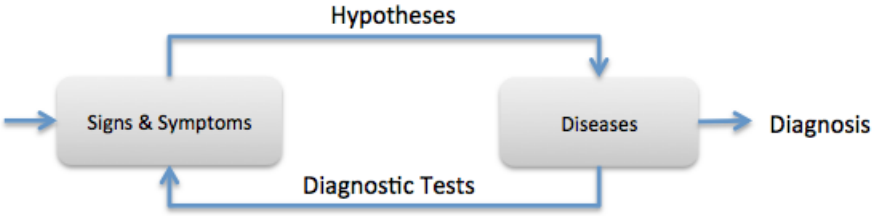


FIGURE 2.1: Overall scheme of the DDx loop

an uncommon illness and when other modes of decision-making have failed. Given the very appropriate attention paid to the cost of care and the cost containment policies, the exhaustive approach is used only rarely nowadays.

The **Hypothetico-deductive** approach is the base of differential diagnosis [48]. It consists in generating and rejecting hypotheses as more information is collected.

Differential diagnosis (DDx) is the common medical process of determining the most feasible disease, or group of diseases, affecting a patient from the observed signs and symptoms. Along a DDx process, health care professionals apply two complementary sorts of knowledge [7, 42]. On the one hand, they apply their knowledge about how to determine the diseases that could better explain the observed signs, symptoms, and test findings (see upper arrow in figure 2.1). On the other hand, they also apply their knowledge on the diagnostic tests that can better contribute to the identification of relevant signs (see lower arrow in figure 2.1).

The beginning of a DDx process starts with the gathering of all the relevant signs and symptoms. This is not easy, because the patient's presentation of her symptoms can be affected by her experiences and her understanding of these symptoms [43]. Moreover, a vector of symptoms merely listed seldom leads to a diagnosis because the same vector of symptoms can overlap between several diseases [49]. Still, this individual diversity does not use to mislead the experienced doctor, as it is seen from the fact that in more that 70% of cases, medical diagnosis is based on the patient's history alone [50].

After the observation of the initial signs and symptoms, the physician applies the first mentioned sort of knowledge to identify a group of diseases that could explain the patient condition. These remain as diagnostic hypotheses. At this point, the physician applies the second sort of knowledge explained in order to deduce which are the diagnostic test that more evidence could bring to accept or to reject the different hypotheses. The resources available and the cost of the tests also condition this selection. After performing the decided test, more findings are obtained that can modify the set of hypotheses till a single hypothesis remains which becomes the final diagnosis.

The DDx process can be affected by the simultaneous observation of several coexisting diseases in the patient or comorbidity.

2.1.2 Treatment

In medicine, the treatment [51] is based on the application of therapeutic measures to improve the condition of a patient.

Usually, a treatment is started once the diagnosis is completed and the diseases affecting the patient are already identified. However, it is possible to start symptomatic treatments independent to the diagnostic process. These treatments are aimed at relieving some of the patient's symptoms without affecting the basic underlying cause. Alternatively, therapy trials exist that help physicians to accept or refuse diagnostic hypotheses during the differential diagnosis.

In Harrison's [51], two approaches to clinical treatment are identified. The first one is based on the point of view of the physician, who generally uses objective parameters of the patient condition that are easily measurable to judge the results of a feasible therapeutic intervention. These parameters could be the findings in a physical examination (e.g., a pupil abnormal size), a numerical measure (e.g., the central venous pressure), the findings in a X-ray exploration (e.g., an unexpected mass located in a organ), etc.

The second approach is focused on the point of view of the patient. In this approach, the treatment is based on subjective parameters as pain relieving, functional maintenance or recovery, or a healthy life enjoyment. This is the central issue of the modern so called patient-centered approach to clinical practice [52].

The components of health status or quality of life of a patient in the patient-centered approach may include physical well-being, the capacity for physical activity, personal and professional performance, sexual function, intellectual function, and overall health perception [53–55].

All these quality of life indicators can be assessed by structured interviews or questionnaires to the patient, that will give the physician an idea of the subjective evolution of the treatment.

Regardless of Harrison's approaches [51], a clinical treatment is composed of health care actions that can be pharmacological (e.g., analgesic or antibiotic), surgical (e.g., neurosurgery or oral and maxillofacial surgery), therapeutic (e.g., psychotherapy or occupational therapy), or rehabilitative (e.g., physiotherapy), among others.

According to their purpose, treatments can be classified into curative or healing, chronic, palliative, symptomatic, and preventive.

Curative or Healing treatment aims to eradicate the disease from the patient. Either if it is a short- or a long-term treatment, the patient is relieved from her diseases at the end of her curative treatment.

Chronic treatment applies to chronic diseases. A chronic disease is a long-lasting condition that can be controlled but not cured. The evolution of chronic diseases uses to be ranked in stages or levels, that go from mild to severe conditions. A chronic treatment aims to stabilize the patient in her current stage or to move her to lower stages. Patient follow-up is part of this sort of treatment.

Palliative treatment or palliative care is applied when the disease can not be cured and it is in a very advanced or final stage [56]. Its main aim is focused on increasing the quality of life of the patient and relieving her from suffering.

Symptomatic treatment pursues to alleviate the symptoms of a unknown disease or if a curative treatment for a known disease has not a short term application. Physicians apply this sort of treatment to increase the comfort and well being of the patient, but also when a suspicion that some symptom can imply undesired organic consequences.

Preventive treatment is applied before a disease is diagnosed and the main purpose is to prevent the disease to appear. According to the American College of Preventive Medicine, preventive medicine aims at keeping patients healthy. It is a medical specialty that focuses on the health of individuals, communities, and defined populations. Its goal is to protect, promote, and maintain health and well-being and to prevent disease, disability, and death [57]. Preventive treatments encompass all Leavell's preventive levels: primary, secondary, and tertiary. Primary prevention keeps the disease process from becoming established by eliminating causes of disease or increasing resistance to disease. Secondary prevention interrupts the disease process before it becomes symptomatic. Tertiary prevention limits the physical and social consequences of symptomatic disease.

The concepts comorbidity and multimorbidity are also narrowly related to clinical treatment. Comorbidity [58] is defined as the presence of two or more diseases simultaneously in the same patient. One of the diseases is primary (or index) and the rest are secondary. Comorbid treatment consists in the treatment of the primary disease, conditioned to the presence of the secondary diseases. Conversely, multimorbidity defines all the concurrent diseases to be primary and the treatments of all the single diseases are combined to provide a unique treatment of the whole patient condition.

2.1.3 Prognosis

A definition of the prognosis of a certain disease [59] is the typical course of the illness in response to a certain treatment, together with the spectrum of random deviations from that course. But medical prognosis covers different meanings that go from the prognosis of the disease (i.e., how is the disease expected to evolve?) to the prognosis of a treatment (i.e., how long is the treatment expected to be or to have an observable effect?), or the prognosis of therapeutic actions (i.e., which changes the application of certain action will cause in the patient condition?), or the prognosis of the patient (i.e., when will a patient start feeling better?) [60].

All these prognostic questions can be answered from the perspective of the evidence-based medicine or the personal experience of the physician [61]. Evidence-based medicine proposes the utilization of prospective or retrospective data and statistics in order to construct prognostic values. Personal experience exploit the skill of physicians to foresee the prognostic results on the basis of past cases the physician has been aware of.

Prognosis is a very important issue when making medical decisions. Both the doctor and the patient must be informed about future possibilities in order to make the best expected decisions [62]. Due to this importance, many modern decision support systems integrate, albeit sometimes implicit, prognosis models for the selection of the best diagnosis or treatment [15].

2.1.4 Clinical Practice Guidelines and Clinical Algorithms

In health care, all the available knowledge of a disease is reported on a Clinical Practice Guideline (CPG). Clinical practice guidelines are defined as systematically developed statements to assist practitioners and patient decisions about appropriate heath care for specific circumstances [63].

Guidelines help clinicians translate best evidence into best practice. A well-crafted guideline promotes quality by reducing healthcare variations, improving diagnostic accuracy, promoting effective therapy, and discouraging ineffective or potentially harmful interventions [64].

One of the main objectives of Clinical Practice Guidelines is to standardize health care. This improves the overall quality of service and makes CPGs to serve as baseline for new strategies. CPGs contribute to sharing and extending medical expertise, being the base of medical training, and providing the dissemination of clinical processes and procedures. Also, by following clinical practice guidelines, physicians may have legal reasons to be protected against possible lawsuits.

Clinical practice Guidelines use to be specific for one disease, which is called the primary disease, and it may contain indications on how to act if the patient has other diseases, which are considered secondary in the CPG. Although CPGs are text based, they include different resources, such as clinical algorithms or tables, to describe clinical knowledge.

A Clinical Algorithm (CA) [65] is a flowchart specially suited for representing a sequence of clinical decisions, for teaching clinical decision making, and for guiding patient care [66]. They are schematic models of the clinical decision pathway that combine health care actions with decision points in a sequential process. CAs are a powerful tools to summarize sequential clinical interventions (i.e., diagnostic or therapeutic process). They can be found as part of CPGs, and they can be the result of a knowledge engineering process [67], or derived from health-care data [68].

Medical knowledge registered in CPGs uses to be labeled with the sort of evidence supporting that knowledge. Several classification of the levels of evidence exist [69], among which table 2.1 describe the evidences for diagnostic knowledge, table 2.2 describe the evidences for treatment knowledge, and table 2.3 describe the evidences for prognostic knowledge.

Level	Meaning
I	High-quality, multi-centered or single-centered, cohort study validating a diagnostic test (with "gold‘" standard as reference) in a series of consecutive patients; or a systematic review of these studies.
II	Exploratory cohort study developing diagnostic criteria (with "gold" standard as reference) in a series of consecutive patient; or a systematic review of these studies.
III	Diagnostic study in nonconsecutive patients (without consistently applied "gold" standard as reference); or a systematic review of these studies.
IV	Case-control study; or any of the above diagnostic studies in the absence of a universally accepted "gold" standard.
V	Expert opinion developed via consensus process; case report or clinical example; or evidence based on physiology, bench research or "first principles".

TABLE 2.1: Evidence Rating Scale for Diagnostic Studies

According to these classifications of clinical evidences, we can conclude that while the evidence of diagnostic knowledge in CPGs is based on retrospective cohort studies of data about diagnosed patients and "gold" standard comparison, the source of evidence for therapeutic knowledge in CPGs are the randomized control trials, and for prognostic knowledge the prospective cohort analysis of data.

Level	Meaning
I	High-quality, multi-centered or single-centered, randomized controlled trial with adequate power; or systematic review of these studies.
II	Lesser-quality, randomized controlled trial; prospective cohort or comparative study; or systematic review of these studies.
III	Retrospective cohort or comparative study; case-control study; or systematic review of these studies.
IV	Case series with pre/post test; or only post test.
V	Expert opinion developed via consensus process; case report or clinical example; or evidence based on physiology, bench research or "first principles".

TABLE 2.2: Evidence Rating Scale for Therapeutic Studies

Level	Meaning
I	High-quality, multi-centered or single-centered, prospective cohort or comparative study with adequate power; or a systematic review of these studies.
II	Lesser-quality prospective cohort or comparative study; retrospective cohort or comparative study; untreated controls from a randomized controlled trial; or a systematic review of these studies.
III	Case-control study; or systematic review of these studies.
IV	Case series with pre/post test; or only post test.
V	Expert opinion developed via consensus process; case report or clinical example; or evidence based on physiology, bench research or "first principles".

TABLE 2.3: Evidence Rating Scale for Prognostic Studies

Cohort studies are based on the isolation of a group of patients (cohort) with a common conditions that we want to study against the general population from which the cohort was extracted. This is a sort of data analysis to find out medical evidences, but there are other approaches such as pre-post studies that compare the same population before and after a clinical action takes place (e.g., drug prescription), or randomized control trials that reduce the risks of bias that can exist in cohort and pre-post studies.

Randomized control trials (RCT) find clinical evidence by randomly allocating the subjects of the study in one or another group representing the different topics of study. For treatment analysis, two groups are defined the group that will receive the treatment and the group that will not receive the treatment but a placebo (control group). For treatment comparison, a separate group is defined for each one of the treatments to be compared (and maybe a placebo group, if the null treatment wants to be compared too). RTCs can be single-blind if the subjects involved in the study do not know which treatment they receive, double-blind if both the studied subjects and the researchers

performing the trial do not know which treatment a subject is receiving, so avoiding on-purpose or purposeless treatment differentiation between groups.

RTCs can be classified according to their design [70, 71] in parallel group trials, or studies where each participant is randomized to one of the groups; crossover trials, where each participant is exposed to each intervention in a random sequence over time; cluster trials, in which predefined clusters of individuals are randomly allocated to different groups; factorial trials, where participants are randomly assigned to individual interventions or a combination of interventions, and split body trials where separate body parts within each participant are randomized.

Another common practice to find out medical evidences is the use of questionnaires and surveys in medicine [72]. The aim is to gather valid, reliable, unbiased and discriminatory data from a representative sample of respondents [73] . However, the information yielded is subject to error and bias from a range of sources. To solve these problems, there are different issues to consider when applying a clinical questionnaire, for example the questionnaire design and the survey administration [74, 75].

There are different areas where the difficulties to obtain data makes the questionnaires and surveys suitable. For example, due the difficulties to obtain predictive data from CPGs, questionnaires have been successfully used in prognosis [76]. Questionnaires and surveys have been used to assess the satisfaction and the performance of medical procedures or computer medical tools. They have also been used in medical education [77].

2.1.5 Patient Simulation

Simulation is a technique to "replace or amplify real experiences with guided experiences that evoke or replicate aspects of the real world" [78]. The term "simulator" used in health care usually refers to a device that presents a simulated patient and interacts appropriately with the actions taken by the simulation participant [79]. Barjis et al. [80] classified health care simulation in four areas: clinical simulation, operational simulation, managerial simulation, and educational simulation.

Clinical Simulation is used to study, analyze and replicate the behavior of diseases and biological processes in the human body.

Operational Simulation is used for capturing, analyzing and studying health care operations, service delivery, scheduling, health care business processes, and patient flow.

Managerial Simulation is used as a tool for managerial purposes, decision making, policy implementation, and strategic planning.

Educational Simulation is used for training and educational purposes. In medical education, there are multiple simulation modalities, among which standardized patients, patient simulators, and virtual patients are the most prominent. Standardized patients are actors or actresses trained to portray patients with specific clinical symptoms and conditions. They are out of the scope of this thesis. Patient simulators [81] are complex systems integrating a manikin and a control unit. The manikin contains sensors that are able to detect external actions performed by the learner. Sensor signals are sent to the control unit to produce reaction in the manikin according to the patient model simulated. Patient simulators can interact with medical devices for the sake of medical realism. Virtual patients (VPs) are a 'specific type of computer program that simulates real-life clinical scenarios; learners emulate the roles of health care providers to obtain a history, conduct a physical exam, and make diagnostic and therapeutic decisions' [82]. They play out on the computer screen.

Operational and managerial simulations are closely interrelated and correspond to the components for health care process management. Conversely, clinical and educational simulations are more related to the care of the patient.

Kneebone [83] identified four advantages of simulator-based training: (1) the learning agenda is determined by the learner availability rather than the patient availability, (2) learner's failures have no direct impact in real patients, (3) simulators can objectively calculate the evolution of the learner's performance, and (4) simulators can enhance both collaborative and individual learning. The list can be extended with additional advantages, such as gaining experience through the exposure to an as-large-as-required number of cases or training sessions [82], or the learners preference of simulators in comparison to paper-based cases.

A review of health-care simulation studies [84] concludes that, for different areas of medical training, simulation has been demonstrated to lead to clinical improvement and melioration of procedural performance, medical knowledge, comfort in procedures, and improvements in performance during retesting in simulated scenarios. Simulation has also been shown to be a reliable tool for assessing learners and for teaching topics such as teamwork and communication.

On the contrary, several limitations of simulator-based medical training have been identified [81]: (1) risk of limited clinical realism, (2) lack of empathy towards the case, (3) reticence of health care professionals to participate in simulated training programs, and (4) danger of an ineffective use of cases causing a meaningful learning

In [85], McGaghie et al identified twelve features and best practices that simulator-based medical training should fulfill: feedback (the learner's decisions should have a formative

response), deliberate practice (the training system should implement means to engage the learner in the learning of medical practice), curriculum integration (simulation-based training should be only one part of the education and duties of the medical student), outcome measurement (reliable data should exist to assess whether the learner is achieving professional skills or not), simulation fidelity (educational goals should direct the simulation technology used and not the other way around), skill acquisition and maintenance (simulation should permit the evaluation of the skills learned and their durability along time), mastery learning (the system should ensure that all learners accomplish all educational objectives with little or no outcome variation), transfer to practice (the skills acquired during the training are useful in real clinical settings), team training (simulators should have to train learners in a work-team typical context where internal communication is essential), high-stakes testing (simulation-based training should allow confirmation that a learner has reached a proficiency level that qualifies her as competent enough to jump to a new learning stage), instructor training (education of the instructor in the correct application of the simulation tool is essential to obtain the best possible outcomes of the learner), and educational and professional context (the learner's context in which the learning takes place has profound effects on the quality of the learning outcomes).

2.2 Computer Technology Support to Clinical Practice

Clinical practice was introduced in chapter 1 as a combination of several routine and cognitive tasks. Among these, in section 2.1 we described some features about three of the most outstanding clinical cognitive tasks: diagnosis, treatment, and prognosis. Along the last five decades, medical informatics has been dealing with the development of multiple computer technologies to support health care professionals in these three important tasks [86, 87], but also delivering medical education tools [88]. The approaches can be mainly classified in three groups: problem-oriented, task-oriented, and holistic modeling.

Problem-oriented technologies address concrete clinical situations, as for example diagnosing diabetes [89] or cancer [90]. Less specific, task-oriented technologies have to do with the support to health care professionals when dealing with a clinical situation such as diagnosis, therapy suggestion, or prognosis, but these technologies are applicable to several medical problems [4]. Holistic modeling is the most generic approach and it pursues the definition of formal models to cover a broad area of medical practice.

In the next sections, we introduce some of these technologies that are directly related to this thesis. For holistic clinical practice models we describe the medical practice model

(MPM) and computer-interpretable guidelines (CIG). For task-oriented technologies we consider differential diagnosis (DDx) generators and decision tables (DT).

2.2.1 Clinical Practice Models

Modeling is the engineering act of producing a representation of a system for understanding, gaining an insight into the properties of that system, and predicting future outcomes. In medicine, modeling is useful in multiple tasks such as education, standardization, dissemination, innovation, and decision support [51]. Moreover, good systematic modeling of medical systems concludes with a formal model that, when applied, it may serve to improve quality, equity, optimization, and automation of processes within the modeled systems.

In heath-care, *medical practice* (MP or the practice of medicine) is a varied and complex process that combines actions in a concrete health care setting that are performed by a group of professionals and caregivers in order to assist patients with their illnesses and ailments. This definition introduces MP at four different levels: functional (i.e., medical actions involved in MP), clinical (i.e., physical and technical requirements implementing MP), human resources (i.e., agents participating in MP), and medical case (i.e., sort of patient or disease addressed).

Modeling these MP levels is seen not only as something beneficial, but also as a need in health care [51, 91–93]. This is reflected in multiple models representing specific medical services (e.g., ER [94] or ICU [95]), tasks (e.g., diagnosis [96], treatment [46, 97, 98], and prognosis [99–101]), or diseases. However, only few studies exist that model MP at the functional level as a combination of diagnostic, therapeutic, and prognostic tasks [7–9, 91, 92, 102].

Among them, the *Medical Practice Model* (MPM) [7] is a holistic functional model of MP integrating diagnosis, treatment, and prognosis tasks.

Extracted from [7], figure 2.2 shows a representation of the modules of the MPM, and their interactions. The modules represent the tasks (as squares) and decisions (as diamonds) that patients and physicians have to perform during the medical process.

This model can be seen as a generalization and extension of other existing models [8, 9, 91, 92, 102]. It contains fifteen subtasks and nine decision points. Subtasks are involved in diagnostic, therapeutic, or prognostic procedures.

Diagnostic subtasks implement a differential diagnosis process and they (1) address the elicitation of patient information, (3) the generation of diagnostic hypothesis, (5) the

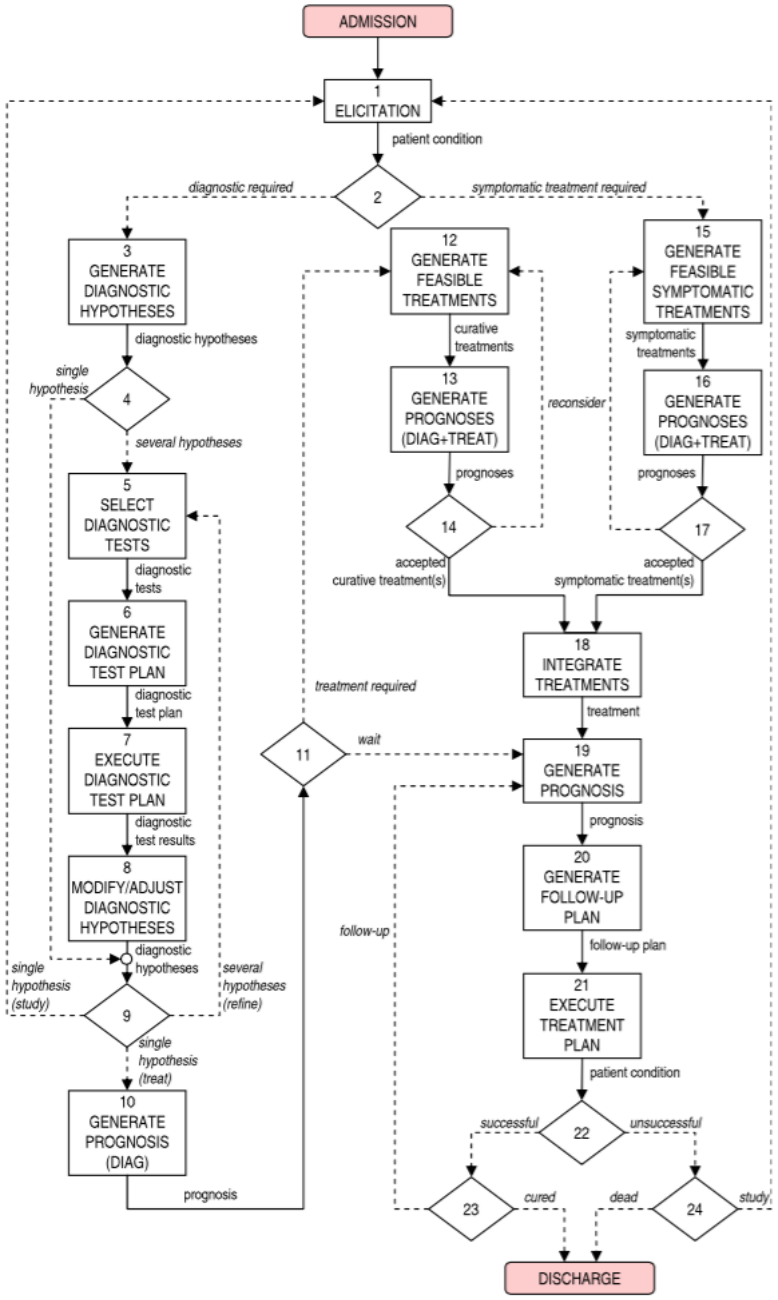


FIGURE 2.2: The MPM functional model.

selection of appropriate diagnostic tests, (6) the optional definition of a plan to perform the required diagnostic tests and (7) the execution of this plan, and (8) the modification of the diagnostic hypotheses according to the results obtained by the diagnostic tests performed.

Therapeutic subtasks represent both symptomatic and curative treatments. Symptomatic treatment is described by (15) the generation of feasible symptomatic treatments, which is the process of determining possible alternative treatments. Similarly, curative treatment is described by (12) the generation of feasible (curative) treatments.

Whenever a symptomatic treatment, a curative treatment, or both are available, (18) they have to be integrated and (21) executed after (20) the generation of a follow-up plan.

Three prognostic subtasks are identified in this model: (10) foresee the evolution of a diagnosed disease or group of diseases in order to determine if a treatment is required, (13,16) anticipate the consequences of a curative or symptomatic treatment before it is applied, in order to decide the most promising treatment, and (19) anticipate the evolution of a patient when a treatment is started in order to decide the follow-up plan.

The model also introduces several decision points. These are: (2) to decide whether a diagnosis, a symptomatic treatment, or both are required, (4) to differentiate if one or several diagnostic hypotheses are possible, and (9) if a single hypothesis is accepted or not. Once the patient is diagnosed we have also (11) to decide whether a treatment is deserved or if it is better to wait for the patient evolution. Whenever a curative or a symptomatic treatment is recommended, we have to decide (14,15) whether to apply them or continue looking for more promising alternative treatments. During patient follow-up three additional decisions have to be made with regard to whether (22) the treatment is being successful or not, and (23,24) if the patient cures or dies.

2.2.2 Computer-Interpretable Clinical Guidelines

In section 2.1.4 the importance of clinical practice guidelines in the medical decision-making process was discussed. CPGs can improve the quality of patient care and reduce costs. However, the access to the correct guidelines and to the information contained in textual guidelines at the point of care may entail unacceptable costs in terms of time, especially during the encounters between the health care professional and the patient requiring the guidelines [103, 104].

In order to overcome this drawback, the medical knowledge contained in textual CPGs can be represented as computer formal structures that, once they are integrated in computer systems, they can be accessed whenever and wherever they are needed at an acceptable time cost. These computer structures representing CPGs, also called computer interpretable guidelines (CIG) [105] and the formal languages to represent them are the basis to develop CIG-based decision-support systems which have a better chance of impacting clinician behavior than narrative guidelines [12].

A number of groups are actively developing CIG representation languages for this purpose [106–109], each one using different approaches. They include document-centric

models, decision trees, probabilistic models, task-network models, ontologies, and extended clinical algorithms. Some of the most referred languages for CIGs repersentation are: Asbru [110, 111], EON [112], GASTON [113, 114], GLARE [115, 116], GLIF3 [117, 118], HELEN [119], NewGuide [120–122], PROforma [123], SAGE [124], and SDA [125].

These languages have been subject to detailed analysis in order to detect the basic common components and requirements. In [126], [127] and [105], it is argued that CIG languages must comprise, at last, three basic representation primitives which are states, decisions, and actions. States are representations of health patient conditions. Decisions are points of the clinical procedures (either diagnostic, therapeutic, or prognostic) where several courses of action may begin and a decision must be taken on which one to follow. Actions correspond to descriptions of pharmacological and non pharmacological interventions.

Moreover, CIGs have also to be able to represent scheduling constraints such as sequences, concurrences, alternatives, and loops. Most of the CIG languages also implement time constraints and dosage indications. The SDA modeling language [125] is a simplistic modeling system addressing all these basic requirements plus the possibility of adding non-deterministic clinical procedures. The interaction of CIGs with electronic health records and care flows is usually described as necessary for the good progression and integration of these computer systems in health care [12], but this still remains an open challenge.

Implementing guidelines in computer-based decision support systems are increasingly applied not only for clinical practice, but also for policy development, utilization management, education, clinical trials, or workflow facilitation [114]. Related issues to CIGs when they are seen as clinical knowledge representation languages are to develop methods and frameworks for guideline modeling and representation, guideline acquisition, guideline verification and testing, and guideline execution [12, 114].

The representation capacity of these languages have also been analyzed in comparison to traditional control-flow theory [128]. Control-flow theory is a mature field with extensive application and validation in the real world and in multiple domains such as business or industry. This theory establishes 43 different control-flow patters that are able to represent any process. In their work, Mulyar et al [128] reach several interesting conclusions such as that CIG languages are able to perfectly represent clinical guidelines with many fewer workflow patterns than control-flow theory allows. Interestingly, CIG languages offer a flexibility that allows them to represent decisions that are not covered by standard workflow management systems, but these can be suitable to represent guidelines.

A critical point of CIG languages is that they are conceived to represent CPGs, but there is an open discussion on whether CPGs are valid in modern medicine nowadays. This discussion is founded on two main facts:

- As discussed in section 2.1.4, CPG highest evidences come from randomized controlled trials, but the subjects participating in these clinical trials use to be homogeneous in terms of their care condition and not necessarily similar to the typical patient attended in clinical settings where the CPGs have to be applied [129]. So, the knowledge acquired for RCT patients is applied to patients that can be completely different in terms of medical needs.
- Another source of discussion of CPGs have to do with their focus on a single disease (primary or index disease) with considerations about additional diseases (secondary diseases or comorbidities). In modern societies, the aging of the population combined with the increment of the prevalence of chronic conditions are giving raise to an enlargement of cases with more than one simultaneous disease, who require a combined treatment of all these diseases (multimorbidity). In a recent study, Bähler et al [130] reported that 76.6% of patients over 65 are multimorbid, they require 11.4 more consultations, and their costs are 5.5 higher in average than non-multimorbid patients. So, the validity of the sort of knowledge contained in single-disease oriented CPGs is called into question [131].

Although CPGs, and consequently CIGs, are focused on single diseases, there are a few studies based in the combination of CIGs for the treatment of comorbid and multimorbid patients [132–136].

2.2.3 DDx Generators

Over the years, several computer tools have emerged to support physicians in differential diagnosis (DDx) [137]. In the last decades a group of these tools have appeared under the name of *DDx generators*. These are "computer programs that assist the clinician by combining symptoms, findings, and other factors to suggest a list of possible diagnoses for consideration" [138].

Some of the most outstanding DDx generators are CADUCEUS, DiagnosisPro, DxMate, DxPlain, Esagil, Illiad, Isabel, Meditel, PEPID, QMR, and ZeroMD, but there are others.

As these systems evolve, the number of health care professionals using them grows. So, DxPlain moved from 4,772 registered users in 2000 to 11,411 in 2004 [139] and Isabel

registered over 12,000 new users between July 2001 and August 2003 [140], with more than 25,000 users in 2004 [141]. The benefits of these systems have been reported in terms of medical error reduction [142, 143], decision support [144, 145], training capability [146, 147], and cost reduction [148].

Last versions of these tools use to be accessible through the Internet. Under significantly different interfaces, all of them provide a similar service. Broadly speaking, DDx generators allow the user to introduce general information about the case to diagnose (e.g., gender and age) together with medical information (e.g., signs, symptoms, risk factors, or comorbidities). Different internal knowledge bases are then used by the generators to rank the feasible diseases this case can have. Some generators also display information about recommended diagnostic tests to reduce the rank of diagnostic hypotheses. The user can continue providing additional inputs, and the generator recalculates the rank in a continuous loop.

Several studies exist comparing DDx generators. Hammersley et al [149] compared Meditel and DxPlain, and conclude that the lists provided by Meditel are more accurate and complete, and it is faster than DxPlain, in 1988. Berner et al [150] compared Dxplain, Iliad, Meditel, and QMR in terms of the diseases considered (knowledge base), the correct diagnosis, the correct diagnosis in the top ten positions of the ranking, the length of the ranking list, the relevance of the top twenty results for a group of experts, and the comprehensibility of these results for these same experts. DxPlain emerged as the best knowledge base and comprehensive results (tied to Meditel), but Meditel remained as one with more correct results, overcame by QMR when only the ten first diseases in the rank are considered. The shorter lists were provided by QMR that was also the best in relevance.

Two years later, Berner et al published an extension of this study [151] with more case studies and making a distinction between the results whether the disease of the case was or was not contained in the knowledge base of the generator. Previous results were confirmed again, but a new quality measure was analyzed replacing the rank list length. In this study the average distance of the correct diagnosis to the first position of the list was calculated. QMR obtained significantly better results.

Bond et al [138] compared DiagnosisPro, DxPlain, Isabel, and PEPID. As opposed to the previous studies, the results measured the subjective opinion of a group of experts in the classification of twenty cases from the New England Journal of Medicine (2010 editions) and from the Medical Knowledge Self Assessment Program (version 14) of the American College of Physicians. Experts scored the accuracy of the DDx generators in a 0-5 range (5 being the higher score). DxPlain and Isabel obtained the highest score, with average value 3.45.

In [152], we performed an analysis of the results of DiagnosisPro, DxMate, DxPlain, and Isabel when they were confronted to multimorbid cases. Our study concluded that all these generators showed a equivalent diagnostic sensibility for multimorbid cases than they have for single-disease patients. DiagnosisPro was the exception, with 35% better sensitivity in front of single-disease cases. DxPlain and DiagnosisPro are the best ones diagnosing single-disease cases, whereas DxMate and DxPlain outperform the rest diagnosing multimorbid cases.

2.2.4 Decision Tables

A decision table [153] is a matrix that relates a set of decision input variables with a set of output actions. It is divided into four areas (see figure 2.3): the condition stub, the action stub, the condition entry, and the action entry. The *condition stub* is the upper left side of the table and it contains the decision input variables (and their respective cardinalities m_i) as a column. The lower part of this column describes the *action stub* as a list of the feasible output actions. In the right hand side of the table, each column represents a *decision rule* that relates the values of a subset of decision input variables (*condition entry*) with a subset of the output actions (*action entry*).

Decision tables have been applied in different domains and for different purposes. Some of the most frequent uses have been [154]: as expert systems [155], for software development [156], with programming languages [157], as control systems [158], in applied mathematics [159], and in economics [160].

In medicine, depending on the sort of decision that a table is made for, decision variables in the condition stub can be signs, symptoms, findings, laboratory results, etc. whereas action variables in the action stub can be diagnostic hypothesis, diagnoses, interventions, etc. These medical concepts can be found organized in hierarchical classifications of international coding systems [161] such as ICD10CM [162], ICPC [163], and ATC [164], or in medical ontologies as SNOMED CT [165] or CPO [166].

As knowledge structures in general computer science, decision tables have a number of advantages at different levels: interpretation, creation, application, and maintenance.

At the **interpretation level**, decision tables are explicit and easy to understand and to interpret, so increasing the clarity of the systems. They have the ability to clearly organize the complexity of a system's logic [39].

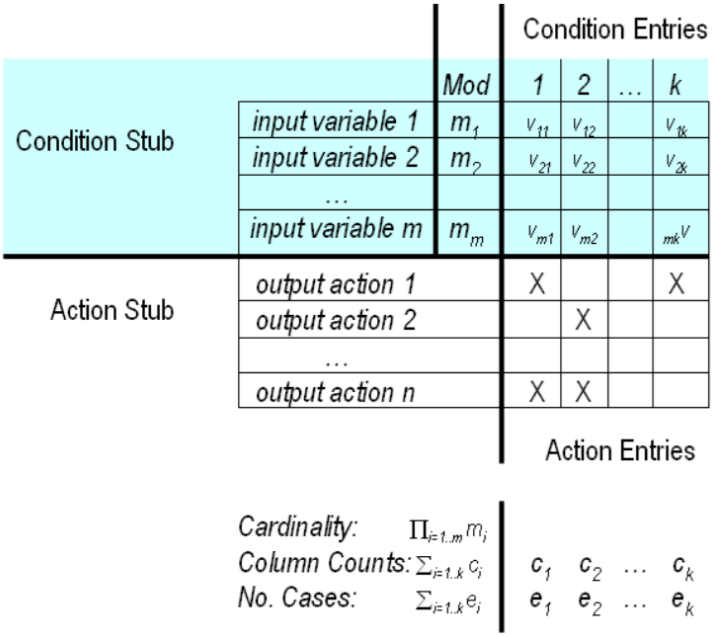


FIGURE 2.3: Conventional Decision Table Structure

At the **creation level**, decision tables stand out as simple and appropriate mechanisms to acquire expert knowledge [167]. Since decision tables are linguistically and structurally understandable, knowledge engineers become less necessary to fill in the tables, in part because domain experts can perform this labor with minor complication [168].

At the **application level**, decision tables can be used both in textual documents as a means to clarify ideas (e.g., clinical practice guidelines), and also incorporated in computer knowledge systems (e.g., decision support systems) [169]. Computer systems based on decision tables have been proved to have an efficiency similar to other systems [170]. In addition, the knowledge stored in decision tables remains separated from the program logic, this causing that changes in the knowledge does not necessarily affect the program. Decision tables can also act as an intermediate structure to easily move from a knowledge representation model to another different model [37].

At the **maintenance level**, DTs are easy to maintain and grow with new information. They are reusable for different systems and easy to validate.

Decision tables also have drawbacks. In the past, they have been criticized for being susceptible to redundancy, ambiguity, contradiction, and conflict, apart of their feasible space exponential growth respect to the number of conditions [41].

Redundancy happens when the knowledge represented in one rule is already embedded in the rest of the rules. **Ambiguity** is an undesired property of the terms used in the table, when these are too generic or having several meanings. **Contradiction** is the situation in which, for the same clinical case, two or more rules in the table are applicable

and they conclude on opposite actions. **Conflict** is when two or more rules are applicable and their actions intersect.

All these advantages and drawbacks of conventional decision tables in computer science are inherited by decision tables when they are used in clinical practice [41]. However, some of the main drawbacks have been addressed by introducing modifications to the conventional decision tables. These modifications have given rise among others to the so called augmented decision tables [39] that can fight redundancy, contraction and conflict; the fuzzy decision tables [169, 170] that formalize ambiguity; the second order decision tables [171] that reduce the size of the table; the semantic decision tables [172] that eliminate ambiguity, and the weighted decision tables [173] that can solve contradictions and conflicts.

Noticeably, semantic decision tables are able to express knowledge in terms of semantic medical concepts. These concepts belong to an ontology as for example SNOMED CT [165], or to the semantic network.

In the past, decision tables have been successfully applied to multiple immediate decisions in health care. For example, we may find applications to perinatal transmission of hepatitis B by immunization [174], coronary artery bypass graft in acute myocardial infarction [175], diagnosis of appendicitis [39], treatment of hypercholesterolemia [176], organ allocation [36], evaluation of chest pain [38], epileptic discharges [177] and many others.

However their application to medical decision procedures has been limited. Two counterexamples can be found in [39, 178] where decision tables were used to decide clinical practice processes such as determining the appropriate intervention of patients with suspected appendicitis, and others.

In relation to the application of decision tables to clinical practice, an interesting antecedent exists that proposed a family of decision table structures able to implement all the components of the Medical Practice Model (MPM) [7] introduced in section 2.2.1. In this work [41], Riaño proposed a family of eighteen different sorts of decision tables able to support health care professionals to answer fifteen diagnostic, therapeutic, and prognostic questions that the MPM model was expected to answer. Figure 2.4 shows the structure of these tables in terms of the medical concepts to be stored in the condition and action stubs (i.e., a patient condition as a list of signs and symptoms, a diagnostic hypothesis as a list of alternative diseases, a diagnosis as a list of diseases, a diagnostic test as a list of tests, a treatment as a list of clinical actions or interventions, and a prognosis as a pair (time, condition) representing the expected time for that patient condition to be reached).

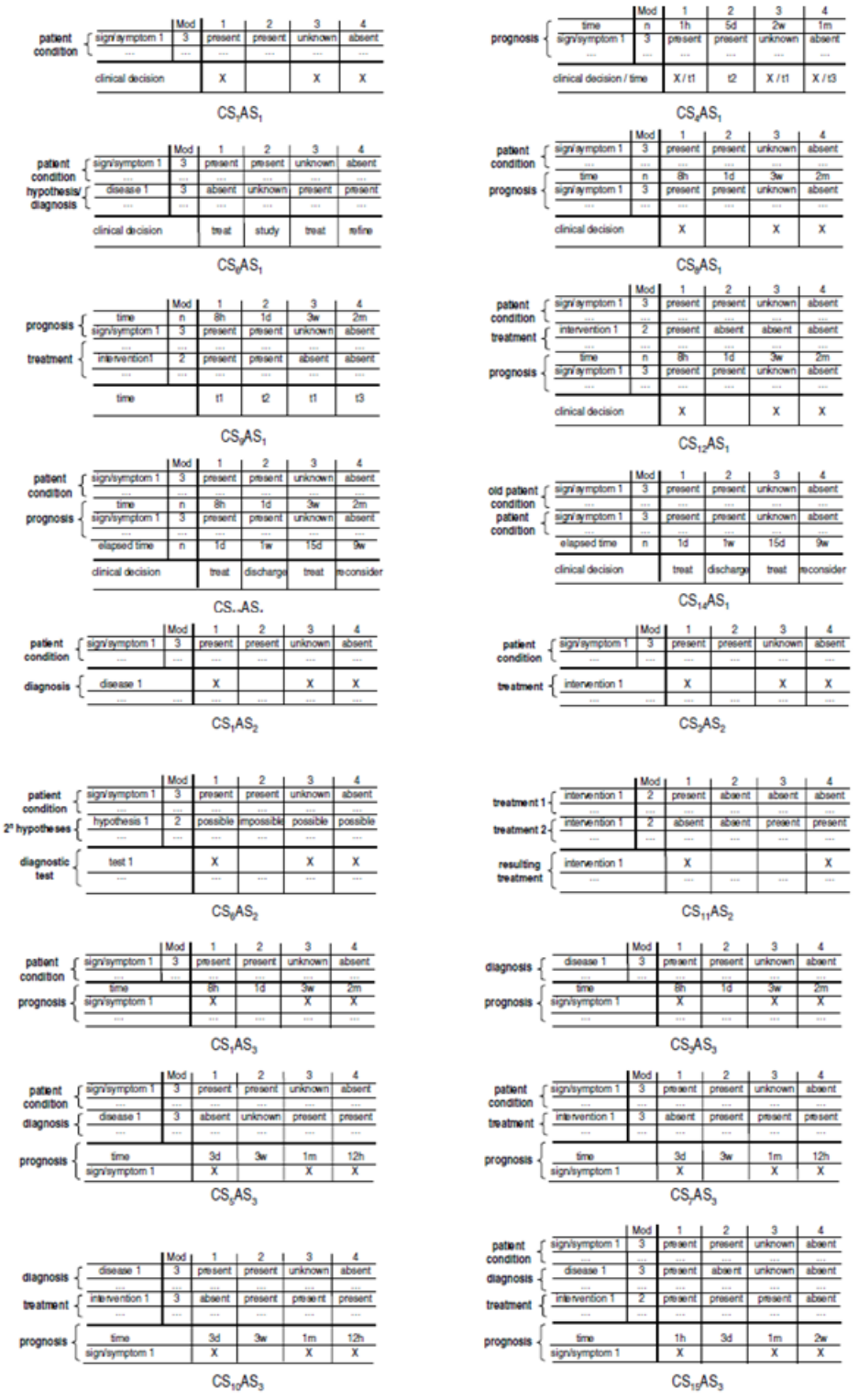


FIGURE 2.4: Family of eighteen decision table structures to implement the MPM model.

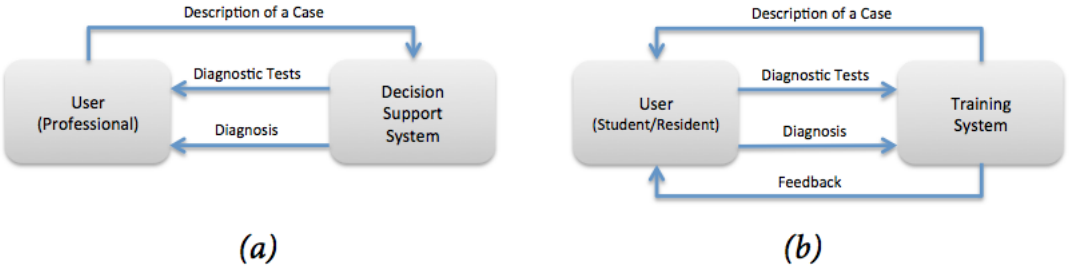


FIGURE 2.5: Operational differences between (a) a DDx decision Support System tool and (b) a DDx Training System tool

2.2.5 Medical Education Tools

One of the main applications of decision tables in DDx are as clinical decision support systems, or as training systems. Between these kinds of systems there are important operational differences that figure 2.5 outlines.

DDx Decision Support Systems (see figure 2.5(a)), as for example DDx generators, receive the description of a case and propose a ranking of the feasible diseases. Feasibility is computed in terms of the internal knowledge contained in the tool. Optionally, they can also recommend a set of diagnostic tests to better arrive to a final diagnosis. This is the case, for example, of Diagnosis Pro and ESAGIL.

On the contrary, the DDx Training Tools (see figure 2.5(b)) are expected to work the other way around. That is to say, they provide clinical cases to the trainee [179], who proposes either a set of possible diagnoses (i.e., diagnostic hypotheses), or a set of diagnostic tests aiming to increase the evidence about disease acceptance or refutation. This behavior mimics reality [7], where patients have a group of signs and symptoms that the physician must interpret to reach a final diagnosis. Finally, the training tool provides a feedback with the correct answers to support the learning process.

Training systems is one of the modalities of e-learning integrated in medical education. E-learning refers to the use of Internet technologies to deliver a broad array of learning modes that enhance learners' knowledge and performance [180].

E-learning technologies offer learners control over content, learning sequence, pace of learning, time, and often media, allowing them to tailor their experiences to meet their personal learning objectives. In diverse medical education contexts, e-learning appears to be at least as effective as traditional instructor-led methods such as lectures.

Students do not see e-learning as replacing traditional instructor-led training but as a complement to it, forming part of a blended-learning strategy [180]. A developing infrastructure to support e-learning within medical education includes repositories, or digital

libraries, to manage access to e-learning materials, consensus on technical standardization, and methods for peer review of these resources.

Some of the e-learning modalities are: simulation technology, synchronous learning delivery, and web-based for standardized patients [181].

The centers specialized in **simulation technologies**, use full mannequins or models connected to various display units that guide the trainees' performance during simulation sessions. Evaluating trainees' skills are usually conducted via observation by faculty members, who complete checklist forms for assessing trainees' psychomotor and technical skills.

The **synchronous learning modality**, such as Webcast, consists of a live video/audio broadcast of training sessions and archival of training materials for later access by participants. Many benefits include: connecting learners from distant sites to live training sessions, creating opportunities for trainers and participants to interact in real time, fostering peer-to-peer feedback, interacting with learning resources such as lecture notes or simulated cases, and accessing training materials for self-paced review.

The use of **standardized patients** has been an integral part of medical education for both teaching and assessment purposes. Web-based and video technology have been piloted to test whether performance-based skills, such as decision making or error disclosure skills, can be taught and evaluated. Clever, et al. [182] concluded that standardized patients-physician interaction was feasible in long distance assessment. Videoconferencing was also effective in assessing a physician's communication skills [183].

2.3 Conclusions

In this chapter we have considered only the elements that are relevant to the thesis. We have left out of the explanation important medical informatics concepts and technologies that are not directly related to our work, such as electronic health care records, ontologies, health care standards, or artificial intelligence topics such as expert systems, decision support systems, or machine learning techniques.

We have described clinical practice as a complex tasks that has to deal, at least, with three important issues which are diagnosis, treatment, and prognosis, and have analyzed how these issues are decomposed into smaller tasks in the Medical Practice Model (MPM). This analysis of the basic components of clinical practice is important to be prepared to define, in the next two chapters, formal knowledge structures to support medical decision making and training.

We have also introduced differential diagnosis (DDx) as one of the key cognitive processes in clinical practice and how DDx generators are computer-based tools supporting physicians in this sort of diagnosis. Clinical practice guidelines (CPGs) have also been presented as the regular means to gather all the available evidence-based knowledge about concrete single-diseases, particularly procedural knowledge related to therapeutic treatments. CPGs are actionable by means of computer-interpretable guidelines (CIGs) which can be described with multiple formal languages, among which some of the most important ones have been mentioned and their representation patterns described. The adaptation of CPGs and CIGs to multimorbid patients is a challenge for future investigations, that we have partially address for diagnosis in chapters 3 and 5.

Prognosis is another cognitive tasks participating in clinical practice. It is hard to find a single definition of prognosis since it uses to encompass multiple meanings. Here, we have adopted one of them which is the anticipation of the future evolution of the health condition of one patient either if she is subject to a perfectly defined treatment or if she receives a null treatment. As we have seen in section 2.2.1, this kind of prognosis participates in three concrete points of the MPM model implementing clinical practice. But not only this, our interpretation of prognosis in this thesis also shares important coincidences with the concept of virtual patient (VP), a computer-based medical education paradigm to train novel physicians to make diagnostic and therapeutic decisions. VP represents the sort of prognosis that we are describing in chapters 4 and 6, and testing in chapter 8.

All this will be achieved under one homogeneous knowledge representation formalism, decision tables (DTs), whose pros and cons have been discussed. Along this thesis different types of DTs will be designed to host assorted types of medical knowledge for decision support with regard to diagnostic, therapeutic, and prognostic & simulating tasks in clinical practice.

Chapter 3

Implementation of Differential Diagnosis with Decision Tables

In this work, we consider differential diagnosis as the clinical process of determining the set of diseases affecting a patient. It uses to be based on the continued refinement of a set of candidate diseases or diagnostic hypotheses, by means of getting new findings that reinforce or weaken the evidences of these candidate diseases and driving some of them to rejection or to a final acceptance. We consider the process divided into three steps (see figure 3.1): making diagnostic hypotheses, selecting appropriate diagnostic tests, and discarding negligible hypotheses.

Making diagnostic hypotheses: All the available relevant signs and symptoms about the patient under study are used in the first step of the differential diagnosis process to identify the set of feasible diseases that could be affecting that patient. The feasible diseases are called diagnostic hypotheses.

Selecting appropriate diagnostic tests: Then, during the second step and according to the current diagnostic hypotheses, a set of diagnostic tests are identified that could bring new evidences to accept or to reject some of these hypotheses. The goal is to employ these evidences to progressively reduce as much as possible the set of hypotheses.

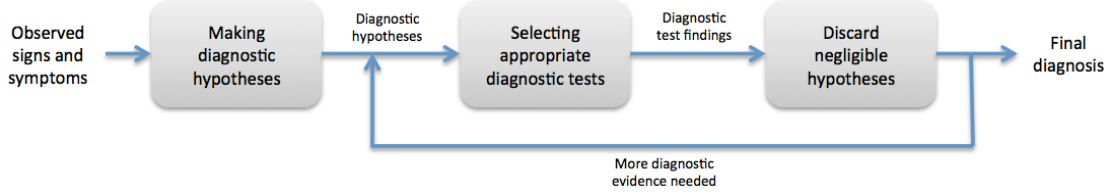


FIGURE 3.1: Subprocesses of the DDx loop

Discarding negligible hypotheses: After performing the selected diagnostic tests, the third step is deserved to evaluate the results obtained and to use the possible findings to refine the current set of diagnostic hypotheses.

Sometimes, the results of the tests do not provide enough evidence to conclude the diagnostic process and new diagnostic tests are needed. In this case, the second and third steps are repeated until a final diagnosis is accepted. This diagnosis can be composed of one or several diseases (multimorbidity).

The application of differential diagnosis (DDx) in medical practice entails a set of additional considerations:

1. Sometimes a diagnostic test needs to be repeated before a finding is taken for sure, as a means of confirmation or second check. For example, the observation of high blood pressure in a patient uses to be followed by a second, or even a third measurement before an elevated value is accepted.

2. Some diagnostic hypotheses can be inferred from the observation of an undefined number of signs, among several. For example, the symptoms of a common cold may include: cough, itchy or sore throat, runny or stuffy nose, congestion, slight body aches or a mild headache, sneezing, watering eyes, low-grade fever and mild fatigue. If a patient presents cough, sore throat, stuffy nose and headache may be enough to suspect about common cold even if not all the symptoms are detected.

3. Some other times the selection of a diagnostic test depends on the need to confirm some symptoms. For example, if a patient refers dysthermia feeling, the physician may measure the temperature to confirm if the patient presents fever.

4. Also, discarding or accepting a diagnostic hypothesis may depend on the observation of some concrete findings. For example, if the measurement of blood pressure is low, the physician can discard hypertension.

5. It could also happen that the absence or the negation of some results in a test is relevant enough to accept or to discard a hypothesis. For example, almost all people who develop severe deep vein thrombosis have an elevated blood level of a clot-dissolving substance called D dimer. If a patient do not present this substance in a blood test, the physician can discard severe deep vein thrombosis.

6. Sometimes a test can be optional. The optional tests may supplement the diagnostic tests in order to evaluate complementary information. For example, to diagnose different causes of hypertension, a physician may request a urine test, but she optionally may request a blood test to assess the severity of the damages produced in the body by the disease.

7. It is also important to distinguish between irrelevant findings and relevant unknown findings. A finding is considered irrelevant to a concrete disease or diagnostic hypothesis if it is not related to that disease or hypothesis. That is to say, the presence or absence of the finding does not provide additional information about the targeted disease or hypothesis. Conversely, a relevant unknown finding is an observation that is positively related to the disease (i.e., its presence or absence may provide some evidence), but no information about it is available at the moment.

In order to represent the knowledge related to DDx with decision tables, we must provide decision tables with a structure able not only to describe the three steps of the process, but also capable to deal with these considerations.

3.1 Decision Tables in Diagnosis

In this section we present an alternative to conventional decision tables, adapted to host the knowledge required in medical diagnosis. These new tables are called grouping decision tables. We also detail how this sort of tables are adapted to manage each one of the steps that we have identified in DDx and that figure 3.1 shows.

3.1.1 Grouping Decision Tables

One of the main features of decision tables (DTs) is their exponential growth when new variables are considered. The variables involved in DDx are related to diagnostic hypotheses and diseases, patient signs and symptoms, and diagnostic tests and findings. In medicine, the huge amount of possible diseases, signs and symptoms, and diagnostic tests makes conventional DTs not to be an efficient representation of DDx knowledge. It is necessary to modify the structure of the DT rules and therefore increase the expressiveness of the tables and consequently reduce their growth to manageable sizes when dealing with DDx.

With this objective, we propose a new model of DT, the *grouping decision tables* (GDT) that extend semantic decision tables [172]. GDTs are decision tables whose condition and action stubs contain concepts that correspond to a semantic codification of terms. In our implementation this semantic codification is SNOMED CT [165, 184]. Their condition entries can take the single values **yes**, **no**, **?**, **void**, or the grouping values **Y#*n*** or **N#*n***, and their action entries can take the single values **X**, **void**, or **opt**.

For condition entries, a value **yes** means that the concept in the corresponding condition stub is satisfied, **no** means that the concept is not satisfied, **?** is when it is an unknown

but still relevant concept, and the void symbol when the concept is not relevant (either if it is known or not) to make the diagnostic decision represented by the rule.

For action entries, the value **X** means that the concept in the corresponding action stub must be applied, the **void** value means that the action concept must not be applied, and the **opt** value means that the application of the action concept is optional (i.e., the user of the table must decide whether to apply it or not).

For each rule, several decision entries can be of the type **Y#*n***, where ***n*** represents an identification number. For the same rule, several grouping values **Y#*n*** are possible with the same or a different ***n***. A grouping value **Y#*n*** with a given same ***n*** value represents the statement that some of the condition stubs related to the condition entries containing such grouping value are satisfied. This is useful, for example, to determine that in order to suspect of a concrete disease (i.e., hypothesis) it is enough to observe a subset of their possible signs and symptoms, but not necessarily all of them. In this case, the condition entries of all the possible signs and symptoms of the disease d_i should be given the value **Y#*n*** with the same ***n***, in the GDT rule, and an **X** value in the action entry corresponding to the action stub of the disease d_i.

A GDT rule can also have one or several grouping values **N#*n***, with one or several identification numbers ***n***. For a given ***n***, the grouping value **N#*n*** in a rule is satisfied if all the condition stubs of such grouping value in the rule are **no** or **?**. This sort of grouping is useful, for example, to build a condition that we have not observed a set of signs or symptoms in a patient.

Another difference with conventional decision tables is that GDTs permit the simultaneous activation of more that one rule. This may be useful in domains like medicine, where several conditions can be observed that activate alternative clinical actions, all of them medically correct. If only one of them needs to be applied, and there is no evidence on which one, then the choice should correspond to the physician.

3.1.2 Sorts of Grouping Decision Tables for Differential Diagnosis

GDTs are used to represent the knowledge required to support DDx. The above mentioned three steps of DDx involve different decision problems that need to be addressed with different knowledge structures. For this reason, we have designed several sorts of GDTs. These sorts of GDTs complement one each other in order to fully support the DDx process depicted in figure 3.1.

3.1.2.1 Deciding the Diagnostic Hypotheses

The first step of diagnosis is to collect all possible information about the patient. This information can come from the patient's own explanation, from the observations of the physician, or from the patient's medical history.

All this information, summarized as signs and symptoms, provides the input required to decide about the diseases that can possibly affect the patient at the diagnostic time. Experienced physicians may categorize these hypotheses according to their probability to the point of ignoring some of them. Our automation of the DDx process requires the consideration of all the possible diagnoses, before this expert filter of the most probable diseases is applied. Initially, we consider that all the hypotheses are feasible.

A Diagnostic Hypotheses Decision Table (DHDT) is designed to answer the question about which the possible diseases affecting our targeted patient can be. Formally speaking, if D_0 and D_1 are two subsets of the SNOMED CT concepts *disease* $\sqcup$ *disorder*, and S a subset of the SNOMED CT concept *clinical finding*, then the condition stub of DHDT is $D_0 \cup S$, and the action stub is D_1. See table 3.1 for a structural representation.

	Rule 1	**...**	**Rule n**
disease$_{01}$		...	
...		...	
disease$_{0i}$		...	
sign$_1$		...	
...		...	
sign$_j$		...	
disease$_{11}$		...	
...		...	
disease$_{1k}$		...	

TABLE 3.1: Diagnostic Hypotheses Decision Table Structure

The condition stub of a DHDT contains the patient past and current diseases (these can be obtained from the patient history record or after questioning the patient) and the possible signs and symptoms observed in the patient. The action stub of a DHDT represents possible diseases affecting the patient who satisfies the description in the corresponding condition stub.

Each rule of a DHDT can represent a type of patient (expressed in terms of a set of signs and symptoms) with none disease currently diagnosed, but possibly having one or several diseases (hypotheses), or a patient already diagnosed of one or more diseases that, according to the current signs and symptoms, is suspected of suffering some other diseases (i.e., multimorbid case).

At the time of execution of a DHDT, none, one, or several rules can be activated, and none, one, or several diseases can be suspected. Suspected diseases are those in the action stub with an **X** value in the corresponding action entry of the activated rules. The union of the suspected diseases of all the activated rules represent the initial set of diagnostic hypotheses of the DDx process.

3.1.2.2 Deciding the Diagnostic Tests

When the diagnostic hypothesis is not conclusive about the patient diagnosis, the physician must decide which evidence needs to be increased in order to reach acceptance or refusal of the current set of hypotheses. This increment is obtained after the application of diagnostic tests that provide additional findings to support these acceptances or refusals. One same diagnostic test may provide information about one or more than one hypotheses, and each hypothesis may require one or more diagnostic tests to get confirmed or refuted.

The selection of diagnostics tests to validate or refuse a disease takes into account different considerations [185]: the feasibility of the tests, the medical costs, the patient comfort and safety, the availability of the tests in the health care setting, etc.

When alternative diagnostic tests are possible, physicians may choose tests giving priority to the time of response, costs and patient comfort and safety, over the tests with high precision. These considerations are reflected in clinical practice guidelines, that give priority to some tests over others, providing one or more sequences of tests to assess the diseases. For example, the diagnosis of acute appendicitis [186] is primarily based on clinical findings. There are several diagnostic tests that may be helpful: ultrasound, computerized tomography (CT) scan, and laparoscopy. CT scan should be preferred over ultrasounds in the diagnosis of appendicitis because it has a higher accuracy, but ultrasound is usually preferred because it lacks of radiation, it has a better cost-effectiveness, and it is more available, if compared to CT scan. Similarly, in spite that laparoscopy is also more accurate than ultrasound, it is an invasive procedure that requires anesthesia and that has a similar risk to performing an appendectomy. Consequently, it should be avoided in favor of ultrasound test, and only utilized in highly justified cases.

Test Selection Decision Tables (TSDT) are a new sort of GDT that we have designed to answer the following question: with the current information, which group of tests will provide better new insights to accept or refuse the suspected diseases? Formally speaking, if D is a subset of SNOMED CT *disease* $\sqcup$ *disorder* concepts, and T_0 and T_1 subsets of SNOMED CT *procedure* concepts, representing diagnostic tests, then the

condition stub of TSDT is $D \cup T_0$, and the action stub is T_1. See table 3.2 for a structural representation.

	Rule 1	**...**	**Rule n**
disease$_1$		...	
...		...	
disease$_i$		...	
test$_{01}$		...	
...		...	
test$_{0j}$		...	
test$_{11}$		...	
...		...	
test$_{1k}$		...	

TABLE 3.2: Test Selection Decision Table Structure

The condition stub includes the possible suspected diseases and the diagnostic tests that have been possibly applied. The action stub of the table contains the list of new possibly suggested tests.

For all the rules in the table, when a test in the condition stub has the value **yes** in the corresponding condition entry, it means that the tests is done and the resulting finding is positive to validate the disease. If the value is **no**, it means that the test is done and the result of the test refuses the disease or diseases affected in the considered rule. A value of **?** in the condition entry means that we have no information about the result of the test or the test was not made.

At the time of execution of a TSDT, none, one, or several rules can be activated, and therefore none, one, or several diagnostic tests might be recommended. The recommended tests will be those in the action stub with a value **X** in the corresponding action entry of some of the activated rules. A rule can be activated if the current diagnostic hypotheses satisfy the condition entry of the rule and either none diagnostic test has been performed yet (i.e., the condition entries related to the tests are all **void**, or **?**), or some diagnostic tests have been performed and they appear with value **no** in the corresponding condition stub if the result of the test is negative, or with value **yes** if the diagnostic test is positive.

Some diagnostic tests might be optional to validate a disease. These tests may not provide enough evidence to take a decision, but they can suggest the physician the direction to take in next decisions.

3.1.2.3 Modifying the Set of Hypotheses

After the diagnostic test results are obtained, physicians decide if the previous hypotheses are maintained, accepted as final diagnosis, or rejected.

With this purpose we have designed a new sort of GDT under the general name of Evaluation Decision Tables (EDT). These tables have been conceived to contain the sort of knowledge required to decide on whether there is enough evidence for a concrete hypothesis to become a final diagnosis or, if new diagnostic tests are recommended.

Table 3.3 shows the structure of an evaluation decision table. Formally, if D_0 and D_1 are subsets of the SNOMED CT *disease* $\sqcup$ *disorder* concepts, and T a subset of SNOMED CT *procedure* concepts, representing diagnostic tests, then the condition stub of a EDT is $D_0 \cup T$ and the action stub D_1.

	Rule 1	**...**	**Rule n**
disease$_{01}$		...	
...		...	
disease$_{0i}$		...	
test$_1$		...	
...		...	
test$_j$		...	
disease$_{11}$		...	
...		...	
disease$_{1k}$		...	

TABLE 3.3: Evaluation Decision Table Structure

An EDT rule is activated when a patient is suspected of having a set of diseases that satisfy the condition stub of the rule and all of the diagnostic test that have obtained a negative result appear as **no** or **void** in the condition entries of the corresponding tests in the decision stub, and all the tests performed and obtaining a positive result appear as **yes** or **void** in the condition entries of the corresponding condition stubs of the tests, in the rule.

When a disease in the action stub has the value **X** in the corresponding action entry of an active rule, it means that the disease is confirmed by the ruke and it becomes part of the final diagnosis. If the value is the **void**, it means that the disease is discarded as diagnostic hypothesis but, if the disease has the value **?** this indicates that the disease in is still possible but not confirmed, and therefore part of the diagnostic hypothesis.

Notice that the condition stub contains the target disease (or diseases) and the possibly performed tests. Sometimes a positive result in a test does not carry enough evidence to accept the hypothesis. Some other times, the diagnostic tests are used to discard the

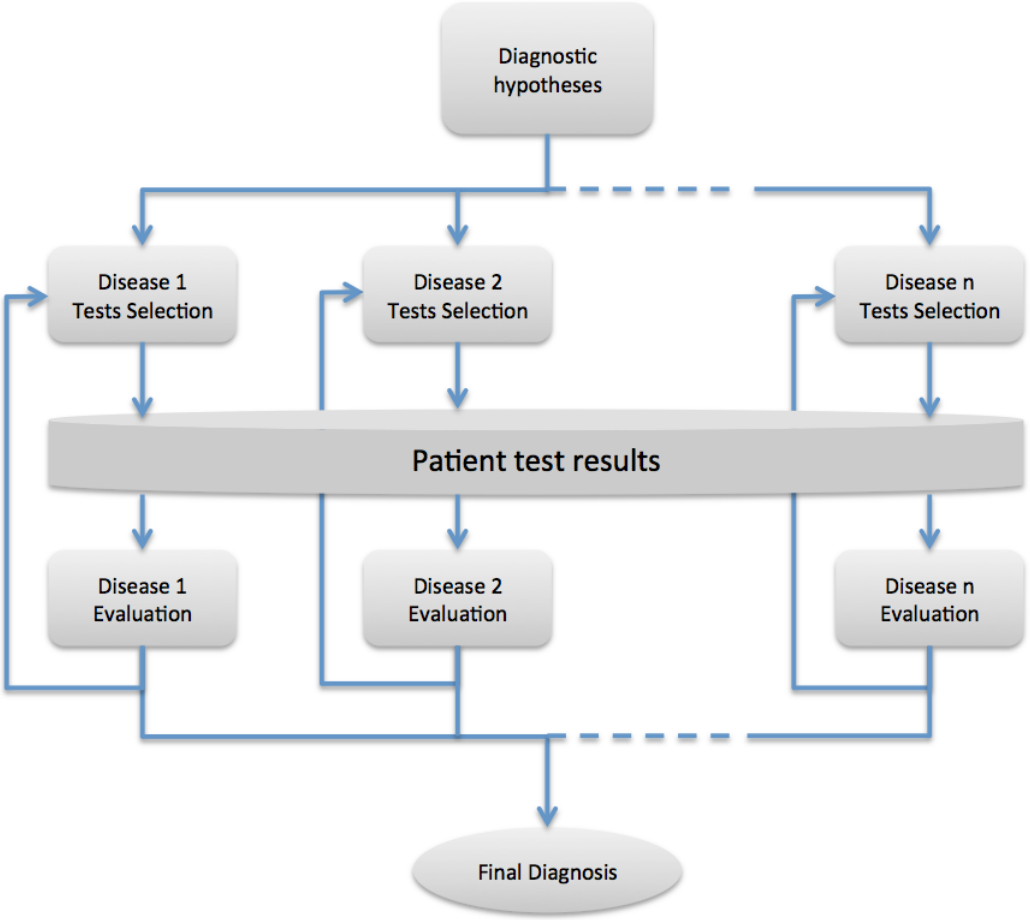


FIGURE 3.2: DT-based implementation of DDx in medical practice

hypothesis, but the precision to predict a disease is very low. Other tests have a high rate of false negatives. In these cases, if the physician suspects that the hypothesis is feasible, new orders to repeat the same sort of diagnostic test can be necessary.

3.2 The Differential Diagnostic System

The three models of diagnostic decision tables described in the previous sections are integrated to implement the DDx process depicted in figure 3.1. This integration is outlined in figure 3.2 to represent our proposed model to implement DDx with DTs.

The model requires as input the available information about the targeted patient. This information includes the signs, the symptoms, and the known diseases affecting the patient. From this information a set of diagnostic hypotheses is obtained after the application of the Diagnostic Hypothesis Decision Table (DHDT) in section 3.1.2.1.

This table recommends a set of diagnostic hypotheses. Each one of the diseases in this set is evaluated separately, creating an independent hypothesis line for each one of the diagnostic hypothesis. For each line, a Test Selection Decision Table (TSDT)

The system select the appropriate tests applying the Test Selection Decision Tables (see section 3.1.2.2) for each disease. As result, the system obtains a set of tests to perform after the combination of the results of all hypothesis lines. It is possible that several diseases need the same diagnostic test (e.g. radiography) or different tests with the same analysis (e.g. glucose test and cholesterol test using the same blood test). The set of tests proposed to the patient considers these possibilities to avoid duplicate tests and to optimize the number of diagnostic tests.

When all the tests have been performed, the new findings provided by the tests are evaluated for each hypothesis disease with the Evaluation Decision Table (see section 3.1.2.3).

For each disease, if the evaluation table rejects the hypothesis, the system discards this investigation line. If the evaluation table accepts the hypothesis, the system includes the disease in the final diagnosis. Otherwise, if the system needs additional information, it restart the new test selection applying the new information provided by the previous tests.

Finally, when all the hypotheses are accepted or rejected, the final diagnosis is concluded with all the accepted hypotheses. If there are more that one disease in the final diagnosis, the system concludes that it is a comorbid case.

3.3 Conclusions

Differential diagnosis (DDx) is a complex cognitive task involved in clinical practice. Our review of clinical publications describing DDx allowed us to identify three main subtasks of DDx, namely making diagnostic hypothesis, selecting appropriate diagnostic tests, and discarding negligible hypotheses. The knowledge underlying in each one of these subtasks was analyzed and three sorts of decision table structures were proposed, all of them under a same DT model that was called grouping decision table (GTD). The knowledge to make diagnostic hypothesis can be represented with diagnostic hypothesis decision tables (DHDT). The knowledge required to select the appropriate diagnostic tests was represented with the test selection decision tables (TSDT). And the knowledge employed to discard negligible hypotheses was captured in the evaluation decision tables (EDT).

All these tables integrate under a differential diagnostic system to provide a complete model of implementation of DDx with decision tables. The capability of the model to represent the diagnostic knowledge contained in clinical practice guidelines will be shown in chapter 5 for the diagnosis of secondary causes of hypertension. The utility of the model will be subject to analysis in chapter 7 with an application to train residents in a tertiary hospital.

UNIVERSITAT ROVIRA I VIRGILI
USE OF DECISION TABLES TO MODEL ASSISTANCE KNOWLEDGE TO TRAIN MEDICAL RESIDENTS
Francis Real Vázquez

Chapter 4

Implementation of Treatment and Prognosis with Decision Tables

In this thesis, we consider treatment as the process of carrying out therapeutic actions aimed at normalizing the situation of the patient. This represents an approach based on the point of view of the doctor (as seen in section 2.1.2), where some objective parameters, which are measurable, are used to assess the patient's condition.

This approach is the same as the one used in clinical practice guidelines. CPGs provide accurate details about the procedures to diagnose a disease, but with the treatment they are less concise. In CPGs, the treatment uses to be focused on the objectives and they provide different alternatives to achieve these objectives.

Usually, it is possible to reach the objectives of the treatment in different ways, using different groups of drugs or applying different treatment procedures. This means that physicians have a variety of possibilities that may work better or worse depending on each concrete patient. It is the doctor who has to decide which actions are taken among the ones proposed in the CPG and she is allowed to modify them as the treatment evolves. In order to represent all this way of doing, when decision tables are applied to health-care treatment, they must contain all the treatment alternatives appearing in the CPG and be flexible enough to allow changes along the treatment.

When a physician decides a treatment action, she uses to foresee the *a priori* consequences of this action in the patient. This kind of prediction is the sort of prognosis that we address in this work, in spite that other interpretations of medical prognosis exist [100].

Clinical actions are different depending on the patient, and the consequences of these clinical actions can be different also depending on the patient. So, for example, a

patient may require a certain dosage according to her age, weight, body mass, co-morbid conditions, risk factors, interaction with other drugs that the patient is taking, etc. and the effect of this drug (i.e., benefits and harms) might vary depending on the patient's natural resistance to the drug, frailty, or the normality parameters of the signs of the patient.

4.1 Treatment Decision Tables

The process of deciding which treatment is recommended for a patient is complex and it involves many issues to consider. In order to make proper decisions, it is important to know the general condition of the patient. This condition is described by generic parameters (e.g., sex, age, or weight), health state parameters (e.g., heart rate or blood pressure), comorbid diseases (e.g., hypertension or diabetes), and other treatment actions in course (e.g., taking beta-blockers or ACE inhibitors).

The factors that are related to a patient but not to her health status, can also be relevant at the time of deciding a treatment, e.g. the economic cost can preclude the selection of some treatments, the patient's medical insurance policy defining the catalog of treatments that are covered or uncovered, the religious believes of the patient can also ban some clinical actions, or the simple will of the patient may favor some clinical actions in front of other ones less comfortable.

For inpatients, other conditioning factors exist that are related to the health care center where they are tended, its facilities and available resources; e.g., some emergency treatment such as CAT scans cannot be possible in some health care centers, or centers in rural areas may have a reduce list of health care services.

Finally, the experience and the expertise of the physicians attending the patient may also condition the selection of one treatment among others.

All these determinants have been considered at the time of designing decision table structures to represent the medical knowledge concerning the prescription of a treatment. These tables are called Treatment Decision Tables (TDT), and they are broadly described as a sort of decision tables that are able to decide upon the question: which treatment actions are recommended for a target patient under a certain given health condition?

The aim of these tables is to record all the possible treatments that clinical practice guidelines suggest for patients in concrete states and to recover this information when a concrete patient needs to be treated. The physician has the final responsibility of

choosing the appropriate treatment having into account the patient condition, the socio-economic factors and preferences, the clinical facilities, and her experience and expertness. Consequently, TDTs are computer actionable tools to help health care professionals in the task of deciding proper treatment prescriptions.

Table 4.1 shows the general structure of TDTs. Formally, if D, S are subsets of the respective SNOMED CT concepts *disease* $\sqcup$ *disorder* and *clinical finding*, and A_0 and A_1 subsets of *pharmaceutical/biologic product* $\sqcup$ *procedure*, then the condition stub of TDTs is $D \cup S \cup A_0$ with D a non-empty set, and the action stub is A_1.

	Rule 1	**...**	**Rule n**
disease$_1$		...	
...		...	
disease$_i$		...	
sign$_1$		...	
...		...	
sign$_j$		...	
action$_1$		...	
...		...	
action$_k$		...	
action$_1$		...	
...		...	
action$_n$		...	

TABLE 4.1: Treatment Decision Table Structure

Each Treatment Decision Table represents the possible treatments of a single disease, which is considered the primary (or index) disease. The condition stub of the table contains this disease, but it may also contain other secondary diseases in order to allow the representation of treatments of the primary disease with co-morbidities.

This sort of tables can also describe the health condition of the patient which a treatment is recommended for. This is accomplished with the introduction of signs and symptoms in the condition stub, and the current treatment in terms of the clinical actions followed, if there are any. For example, TDTs are able to represent that all the patients suffering from arterial hypertension in grade 1 require diuretics [187], but also that all the patients suffering from arterial hypertension but that have diabetes mellitus should take ACE inhibitor instead of diuretics [188], and also that if the patient is already taking ACE inhibitor, then she has to complement this drug with angiotensin receptor blockers.

TDTs extend semantic decision tables [172]. Their condition and action entries need to be more descriptive than the ones of Grouping Decision Tables because of the high number of possible alternatives to achieve the treatment objectives. This amount of possible alternatives causes tables to grow and to lose clarity and easy handling. In order

to overcome this issue, TDTs are able to incorporate a new type of value generalizations to their condition entries which will reduce the size of the tables and will make them more readable.

Their condition entries can be boolean, numeric, ranges, dosages, repetitions, or empty. Boolean values (i.e., yes or no values) are useful to represent, for example, the presence or not of a risk factor, a secondary disease, if a concrete clinical action was done, or if the patient is taken a concrete drug. Numeric values are useful to represent punctual numeric information, as the grade of the disease. For example, the New York Heart Association (NYHA) Functional Classification distinguishes between four levels of heart failure, 1-4. Range values are used to describe intervals with respect a concrete sign, for example if the patient has a blood pressure above 140 mmHg. Dosage values (e.g., low, medium, high) are useful to describe the dose that the patient is currently taken with regard to a particular drug. Repetition values are useful to represent the number of times that a treatment action has been done, for example, several shocks are treated with epinephrine bolus, but giving more than 2-3 bolus may be dangerous for the patient.

TDT action entries values can take positive, negative, optional, dosage, and logic values. A positive value (represented by **X**) means that a treatment action must be applied. A negative value (representing with a void table cell) means that the treatment action is not recommended. Optional values (identified with the term **Opt**) are used when the application of the action may complement the other treatment actions; for example, if a patient needs non-steroidal anti-inflammatory drugs such as aspirin or ibuprofen for a long time, and we suspect that this may affect her gastrointestinal tract, some optional stomach-protective drugs could be necessary to complement the treatment. Dosage values (represented by exact values or value groups like **low**, **medium**, or **high**) are useful to represent treatment drug prescriptions. It means that a drug is recommended with a specific dosage or dosage group. For example, the administration of acebutolol for hypertension may start with 400 mg/day, or with a low dosage (between 300-500 mg/day) and be increased until it achieves the treatment objectives, or in emergency arrhythmia, start with 1000 mg/day or a high 900-1200 dosage. Finally, logic values (represented with the terms **OR** or **XOR**) are useful to describe different alternatives where at least one of them must be chosen (**OR**), or one and only one of them must be chosen (**XOR**). For example, it is possible to use XOR when different drugs have a similar effect but only one has to be prescribed (e.g., beta-blocker or ACE-inhibitor). In a similar way, the logic value OR is useful when a physician has to prescribe changes in the patient's lifestyle, not all the possible changes are applicable, but at least one of them is required (e.g., walk or bike exercise).

4.2 Prognosis Decision Tables

When a physician prescribes a treatment, she uses to use her experience to foresee the expected evolution of the patient condition as a consequence of that treatment. These changes can be known in a generic way, but the concrete changes in a specific patient are not always easy to anticipate. For this reason, when a physician recommends a treatment, it is necessary to control periodically the patient condition and her response to the treatment.

Often, the evolution of the patient differs from the medical expectations and the treatment decisions have to be reassessed. Sometimes, the effect of a treatment action is the one expected, but non desired secondary effects may happen. As Dwight D. Eisenhower said [189]: "plans are useless but planning is indispensable".

Prognosis Decision Tables (PDT) try to solve the question: What is the expected evolution of a patient when a group of treatment actions is applied?

Table 4.2 shows the general structure of PDTs. Formally, if D and A are subsets of the respective SNOMED CT concepts *disease* $\sqcup$ *disorder* and *pharmaceutical/biologic product* $\sqcup$ *procedure*, and S_0 and S_1 are set of concepts of *clinical finding*, then the condition stub of PDTs is $D \cup A \cup S_0$ with D a non-empty set, and the action stub is S_1.

	Rule 1	**...**	**Rule n**
disease$_1$		...	
...		...	
disease$_i$		...	
action$_1$		...	
...		...	
action$_j$		...	
sign$_1$		...	
...		...	
sign$_k$		...	
sign$_1$		...	
...		...	
sign$_n$		...	

TABLE 4.2: Prognosis Decision Table Structure

PDTs are able to represent anticipation rules about possible evolutions of a patient for a concrete single disease which is called the primary disease. Other secondary diseases are also possible to be part of the table as co-morbid conditions of the primary disease.

Treatment actions may affect patients with different diseases in a different way. All these conditionals are captured as rules in the PDT. The condition stub contains the target

disease, other diseases that may interact with the treatment actions, the actual patient condition expressed in terms of constraints about a subset of signs and symptoms, and the treatment actions that are going to be applied. The action stub of the table contains the modification in the signs that the patient is expected to experience.

The condition entries for the PDT rules may include boolean values, range label values, and empty values. A Boolean value (i.e., **yes** or **no**) are useful to represent, for example, the presence or not of a disease, or if a treatment action is in course. Range label values (with ranges **low**, **medium**, or **high**) are used to represent both pharmacological dosage as for example amount of drugs in infusions (e.g., high dosage, when it is above 500 mg/day), and the qualification of vital signs as for example the heart rate (e.g., low heart rate, when it is under 50 beats per minute).

The values in the action entries of the rules can be boolean, increment, decrement, finish, and empty. A boolean value (i.e., **yes** or **no**) are useful to describe true-false signs as for example the emergence of a jugular venous engorgement. The increment and decrement values are useful to describe the behavior of numerical signs as for example the variation in the central venous pressure. The finish values are used when an action is not recommended and potentially harmful for a patient. This may happen for example when a drug is prescribed and the patient is allergic to that drug.

In our case, PDTs could not be obtained from clinical practice guidelines because of the lack of information about prognosis for each sign. Conversely, these tables had to be filled by experts or obtained after the computer analysis of health care data.

4.3 Patient Simulation with Prognosis Decision Tables

The structure of PDTs makes them not only useful for prognostic purposes as the one defined in this thesis; that is to say, the anticipation of possible future evolution of the signs and symptoms of a patient according to her current diseases, signs and symptoms, and the treatment applied, but also as a key component for the simulation of artificial patients.

Our proposal for patient simulation [190] considers that under the same health conditions, different patients can have different normality parameters for their vital signs, and their response to a same treatment can vary. For example, a systolic blood pressure (SBP) of 90 mmHg could be considered normal for a certain patient, but very low for a patient with hypertension because the normality parameter of these two patients for SBP are different. Also, some patients may present resistance to certain drugs or hypersensitivity to some treatments. Sometimes, the general health condition of a patient

or her risk factors (which are not necessarily related to the disease under consideration) can make certain clinical actions not to be recommended or even counter-indicated. Additionally, the same dosages may have different effects depending on each clinical condition.

These are some of the reasons why the application of PDTs alone is incomplete to implement realistic simulation. So, we complemented PDTs with a Patient Model (PM) that allows the description of different sorts of patients. See table 4.3. This way, patient simulation in this thesis is defined as the personalization of the predictions of a PDT to the PM describing the patient under consideration.

```
Patient ID : ##
AGE: ##          WEIGHT: ## Kg
SEX: ##          HEIGHT: ## m
...
PRIMARY DIAGNOSIS: ##
SECONDARY DIAGNOSES: ##

DESCRIPTION: ##

VITAL SIGN RANGES:     MIN     LOW     HIGH    MAX     UNITS
sign-1                 ##      ##      ##      ##      mmHg
...
sign-M                 ##      ##      ##      ##      bmp

ACTION SENSITIVITIES:
action-1  ==>  ## %
...
action-N  ==>  ## %

INITIAL SIGNS:
  <sign-1 = ##>
  ...
  <sign-M = ##>
```

TABLE 4.3: Patient Model representation structure

While a PDT can be seen to contain the knowledge describing the evolution of a standard patient diagnosed of a concrete disease in front of a standard treatment and, therefore, proposing standard prognoses, a PM represents the knowledge about a concrete sort of patient or clinical case. As figure 4.1 depicts, the PMs allow the customization of the results produced by the PDTs in accordance to the features of single patients.

In the PM we determine the special behavior of a type of patient for each treatment action by defining her sensitivity/resistance with a percentage: 0% representing full

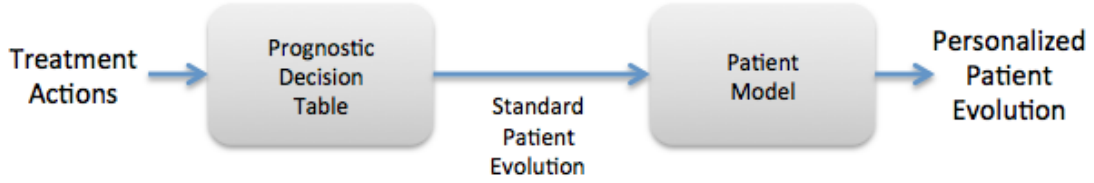


FIGURE 4.1: Patient Simulation Architecture

resistance, values between 0% and 100% partial resistances, 100% the standard effect, and values above 100% crescent sensitivities. See the section `ACTION SENSITIVITIES` in table 4.3 and how a subset of SNOMED concepts in *pharmaceutical/biologic product* ⊔ *procedure*, representing clinical treatment actions, can define different patient sensitivity percentages.

Furthermore, comorbidities and physical conditions may vary vital signs normal references and their limits, in every single patient. Clinicians use this sort of variations to determine the treatment goals. In the PM we are allowed to specify normality parameters of cases by means of ranges that will be used not only to assess the effects of clinical actions over tolerable limits but also to decide on patient discharges. See these ranges under `VITAL SIGN RANGES` in table 4.3.

In PM, cases are allowed to contain sensitivity/resistance percentages for all the relevant clinical actions related to the considered diseases, and vital sign ranges for all the signs and symptoms that are relevant to the diseases under consideration. Table 4.3 shows the basic template to define PMs. Patient normality parameters are defined under the section `VITAL SIGN RANGES`, with boundaries `MIN`, `LOW`, `HIGH` and `MAX`. These values define the ranges for unacceptably low (below `MIN`) and unacceptably high (above `MAX`) causing the simulation to stop, at risk (between `MIN` and `LOW` or between `HIGH` and `MAX`) requiring urgent intervention, and normal (between `LOW` and `HIGH`) to consider discharge. The section `INITIAL SIGNS` describe the values for all the sings of the case at the time of admission.

4.4 Long Term Treatment Decision System

PDTs or the combined use of PDTs with PMs can be used to support decision making about medical treatment within a short term simulation. They provide a prognostic of the patient evolution when several actions are applied. When it is interleaved with a step of therapy adjustment, this prognostic may be repeated several times in order to plan long term treatment, so supporting clinical follow-up.

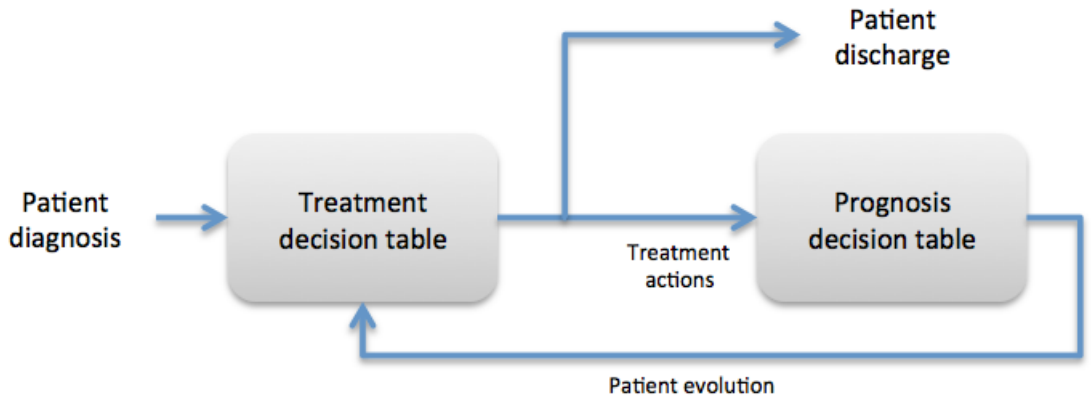


FIGURE 4.2: Treatment and Prognosis Architecture

Figure 4.2 shows how to iterate the treatment and prognosis decision tables in a long term follow-up decision system. If DM is used, it have to be considered together to the prognostic decision table.

First, when the patient is diagnosed of one primary disease and possibly some secondary diseases, the treatment decision table (TDT) is used to suggests none, one, or several treatment actions, according to the current condition of the patient (i.e., current values of relevant signs and symptoms). Then, the prognosis decision table (PDT) acts as a virtual patient, providing the expected evolution. With this new patient status, the treatment decision table (TDT) can suggest a new set of treatment actions. This process can continue for a short- or a long-term until the treatment decision table (TDT) concludes that the patient should be discharged.

Each loop of this process can be considered to define a short-term treatment, but when this process finishes after several loops, the resulting patient condition at the discharge time represents the patient condition after the application of a long-term treatment. This new condition may be compared with the treatment objectives to assess the potential quality of this therapy both at short- and long-term.

4.5 Conclusions

Treatment is a medical process in which the physician establishes some therapeutic objectives and, through different clinical actions, she tries to get them. The clinical practice guidelines provide the different alternatives to approach these objectives, but the best option depends on the patient and on the medical criterion. Here, we present the treatment decision tables (TDT) to abstract the knowledge provided by clinical practice guidelines.

Prognosis is the medical process mainly based on evidence studies or medical experience, that determines the medical decisions along the medical treatment. This thesis focuses on the prognosis of the patient evolution when the patient is subject to concrete treatment actions. The prognosis decision table (PDT) is introduced to model medical experience in prognosis knowledge.

PDTs are also introduced to make patient simulation possible. But this simulation only represents the evolution of standard patients in front of standard treatments. In order to personalize the simulation with different types of patients, a patient model representation was introduced.

In addition, a system that combines treatment decision tables and prognosis decision tables are presented to simulate automated long term treatments.

Chapter 5

Use of Diagnostic Decision Tables in Secondary Causes of Hypertension

In chapter 3, three types of decision tables to implement a differential diagnostic process were presented. Having these knowledge structures is only the first step of the generation of medical knowledge to support health care professionals in the differential diagnostic process. In a second step, these computer structures need to be filled with the required medical knowledge. Capturing this knowledge from CPGs and representing it inside the proposed decision tables is an important knowledge acquisition task that requires further explanation. In this chapter we describe the usage of these tables to represent differential diagnosis knowledge related to a particular medical problem: the secondary causes of hypertension. The methodological procedure aiming to implement DDx knowledge with DTs is also described.

5.1 Secondary Causes of Hypertension

Blood pressure is the pressure of the blood against the inner walls of the blood vessels. Arterial blood pressure (BP) refers to pressure in arteries. BP uses to be represented by two measures: systolic blood pressure and diastolic blood pressure. Systolic blood pressure (SBP) is the blood pressure during contraction of the ventricles, whilst diastolic blood pressure (DBP) is the minimum blood pressure measured during relaxation and dilatation of the ventricles. Normal values of BP (or SBP-DBP) can vary for adults and children. So, Mayo Clinic describe BPs below 120-80 mmHg as normal for adults, and upper values till 139-89 mmHg as pre-hypertension that requires the adoption of

healthy lifestyle habits. Higher values are considered the patient to suffer from Arterial Hypertension (AH), an abnormal high blood pressure in the patient's arteries [191] that may require pharmacological treatment. AH is the most common condition seen in primary care [192] and it may lead to serious problems such as myocardial infarction, stroke, renal failure, or death, if it is not detected promptly and treated appropriately.

AH is considered a disease (called primary or essential hypertension) by many medical standards such as ICD9-CM, ICD10, and SNOMED-CT, but also a clinical condition induced by other causes or diseases (called secondary causes). Secondary AH is the presence of a specific condition known to cause hypertension, which may be the primary cause or a contributing factor in a patient who already has primary hypertension. Although uncommon, only 5-10% of hypertension cases are due to secondary causes [193], it may cause major morbidity for a subset of patients [194]. The clinical practice guideline of AH [191] identifies *acromegaly*, *adrenal Cushing's syndrome*, *coarctation of the aorta*, *glomerulonephritis*, *hyperparathyroidism*, *pheochromocytoma*, *renovascular disease*, and *sleep apnea* as eight of the main secondary causes of arterial hypertension.

This guideline [191] and the eight CPGs corresponding to the above mentioned secondary causes of AH [195–202] were analyzed with the help of senior general practitioners of the health care centers *Hospital Clinic de Barcelona* and *SAGESSA*. The diagnostic knowledge available in these guidelines was converted into clinical algorithms representing a decision procedure for a common and easy understanding by both the physicians and the computer scientists. The clinical algorithms obtained for the secondary causes of hypertension are showed in Appendix A. These clinical algorithms [203] describe each one the sequences of clinical tests that are needed in order to accept or refuse one concrete secondary cause of hypertension.

Figure 5.1 shows the clinical algorithm obtained from the CPG about acromegaly [195]. Acromegaly (commonly called giantism) is a hormonal disorder that develops when the pituitary gland produces too much growth hormone during adulthood. When this happens, bones increase in size, including those of hands, feet and face.

The algorithm starts with the suspicion of a disease, in this example acromegaly. Initially, it proposes one test, the *plasma IGF 1 measurement*, and depending on the result the disease is rejected or a new test is suggested. This time the measurement of the *plasma growth hormone* in order to gain more evidence about the suspected disease. Even if this second test has a positive result a third test is recommended by the CPG, which is a *magnetic resonance imaging of the head*. If all tree tests are positive, acromegaly is accepted. If the first two tests are positive, but the last one turn out to be negative, still a last test is indicated (*computed tomography of the chest and abdomen*) before acromegaly is accepted or rejected.

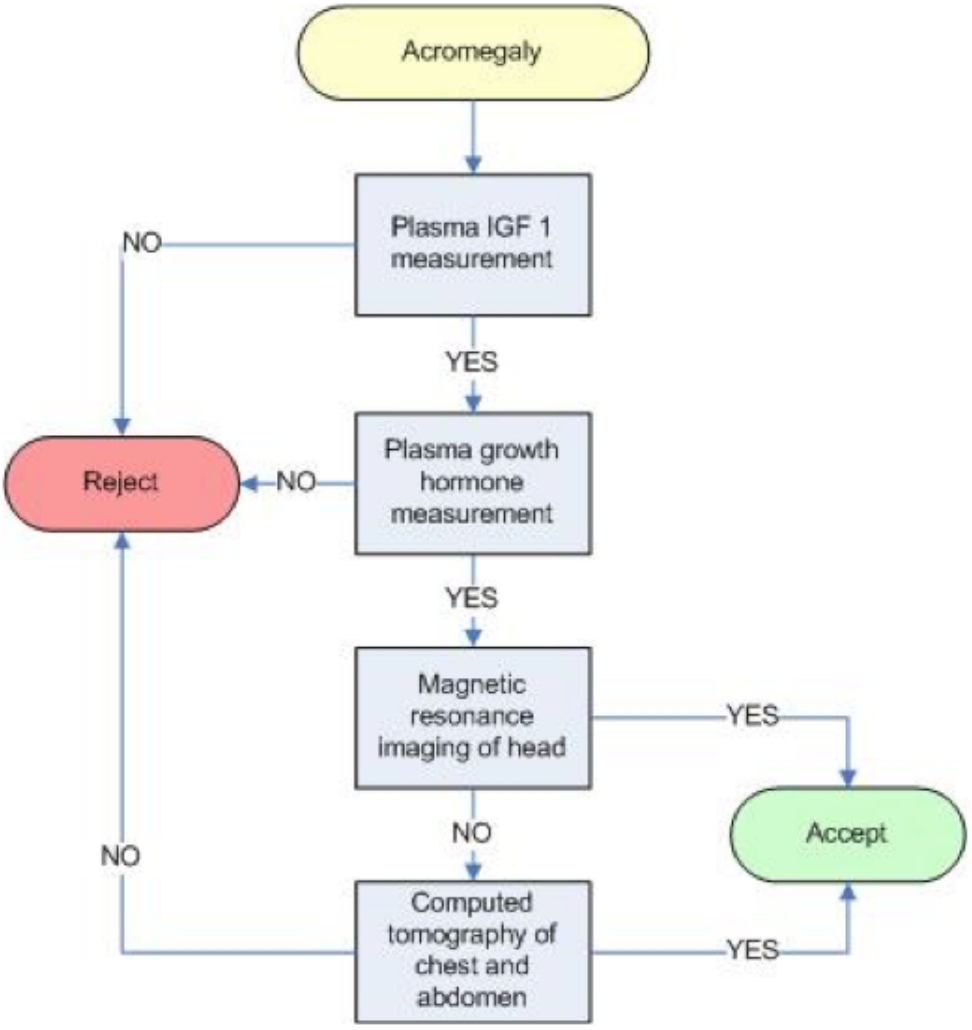


FIGURE 5.1: Acromegaly Algorithm

This sort of clinical algorithms is useful to build the Test Selection Decision Tables explained in section 3.1.2.2 and the Evaluation Decision Tables introduced in section 3.1.2.3. Moreover, a differential diagnosis process also requires a Diagnostic Hypotheses Decision Table to determine which among all the possible diseases are feasible according to the current condition of the patient.

5.2 Making a Diagnostic Hypotheses Decision Table

The Diagnostic Hypotheses Decision Table (DHDT), presented in section 3.1.2.1, is a common table for all the diseases participating in the differential diagnosis process. In our example of diagnosing secondary causes of hypertension, the eight possible secondary causes define the rules of the DHDT.

This table must be built from the CPGs of the diseases we want our table to be able to suspect of. Figure 5.2 summarizes a methodology to make DHDT. For each disease, all the relevant signs and symptoms have to be detected and then mapped into SNOMED CT concepts, before they are classified in four categories: mandatory, alternative, irrelevant, and impossible.

Mandatory signs are those that define the disease. Patients with the disease must have these signs and if some of them is not observed, there is no reason to suspect of the disease. In the case of secondary causes of hypertension, *arterial hypertension* is a mandatory sign.

Alternative signs are those that can provide evidence about the disease. The presence of one or more of these signs in a patient represents a solid basis both to suspect that the patient may have the disease and to consider it in the differential diagnosis process. In the example of acromegaly, the sign *acquired skeletal deformity* is alternative.

Irrelevant signs are those that can appear with the disease sometimes, but they alone do not provide a basis to suspect the disease. For example, in acromegaly *headaches* or

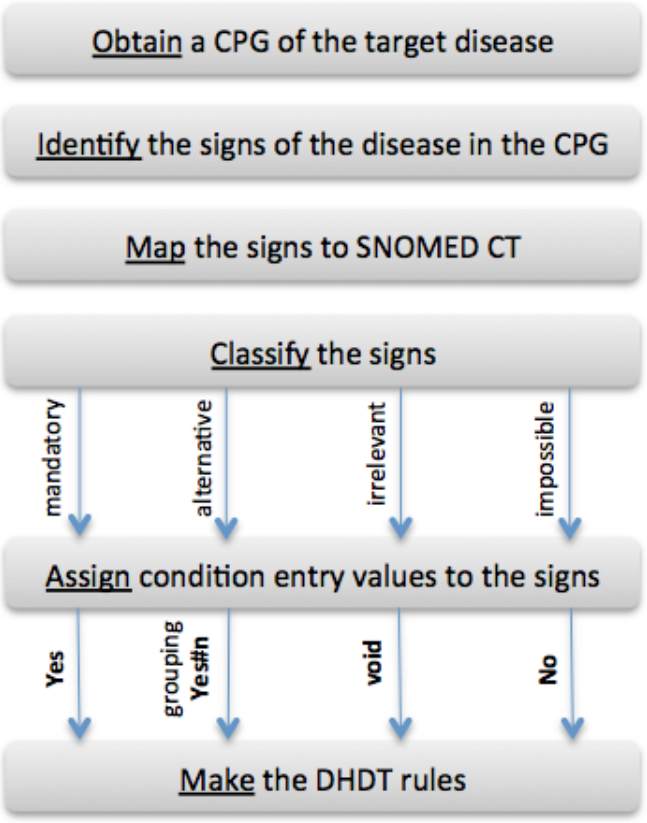


FIGURE 5.2: Methodology to make Diagnostic Hypotheses Decision Tables

fatigue can appear sometimes, but according to the medical experts, they not provide evidence that the diagnosis might be acromegaly.

Impossible signs are those that when observed, the disease is not possible.

Once classified, each sign receives a condition entry value. Mandatory signs are given the value **Yes**, all the alternative signs are combined under the same grouping positive value **Yes#1** (if several alternatives are possible, groupings **Yes#*n*** with increasing n are determined), irrelevant signs are assigned a **void** value, and impossible signs a **No** value.

Then a new rule for the disease is introduced in the DHDT with the identified signs in the condition stub and the corresponding values in the condition entry of the rule. The disease is appended to the action stub and an **X** is placed in the corresponding action entry. Several rules can be possible if the condition to suspect the disease is complex.

This is a simple incremental methodology by which new diseases can be included for differential diagnosis consideration with no need to reconsider the rules previously contained in the DHDT.

Figure 5.3 shows the DHDT for differential diagnosis of secondary causes of hypertension obtained after the application of this methodology to the clinical guidelines of *acromegaly* [195], *adrenal Cushing's syndrome* [196], *coarctation of the aorta* [197], *glomerulonephritis* [198], *hyperparathyroidism* [199], *pheochromocytoma* [200], *renovascular disease* [201], and *sleep apnea* [202]. For this task we counted with the appreciated support of two senior medical experts from the Hospital Clínic de Barcelona and SAGESSA.

In the condition stub of the table we can observe the mandatory, the alternative, and the irrelevant signs after the SNOMED-CT codification. None of the considered CPGs describe impossible signs, so the table does not contain such sort of signs. Some of the signs are relevant only for one rule, for example *snoring*, for more that one rule, for example *muscle weakness*.

The action stub contains the eight secondary causes of hypertension. Each column to the right represents a rule that defines the condition to consider one disease as possible cause of secondary hypertension. For example, the first rule describes the patient condition to suspect pheochromocytoma. This rule shows that a patient needs to have an hypertensive disorder (mandatory sign), and also one or more additional signs among abdominal pain, acute necrosis, constipation, fever, or tachycardia (alternative signs), in order to include pheochromocytoma in the list of diagnostic hypotheses.

Abdominal pain	Yes#1				Yes#1			
Acquired skeletal deformity						Yes#1		
Acute necrosis	Yes#1							
Amenorrhea		Yes#1						
Anxiety		Yes#1						
Apnea							Yes#1	
Bone pain		Yes#1						
Constipation	Yes#1							
Edema		Yes#1	Yes#1					
Epistaxis								Yes#1
Excessive sleepiness					Yes#1		Yes#1	
Fever	Yes#1							
Headache								Yes#1
Heart murmur								Yes#1
Hirsutism		Yes#1						
Hypertensive disorder	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Insomnia		Yes#1						
Muscle weakness		Yes#1			Yes#1			Yes#1
NOT Age more than 50 years				Yes#1				
NOT Progress satisfactory				Yes#1				
Nausea and vomiting			Yes#1		Yes#1			
Oliguria			Yes#1					
Polyuria					Yes#1			
Skin striae		Yes#1						
Snoring							Yes#1	
Tachycardia	Yes#1							
Unequal pulse								Yes#1
Acromegaly						X		
Adrenal Cushing's syndrome		X						
Coarctation of aorta								X
Glomerulonephritis			X					
Hyperparathyroidism					X			
Pheochromocytoma	X							
Renovascular hypertension				X				
Sleep apnea							X	

FIGURE 5.3: Secondary Hypertension Diagnostic Hypotheses Decision Table

5.3 Making the Test Selection Decision Tables

According to the differential diagnostic process described in chapter 2, after the identification of the diagnostic hypotheses, a set of diagnostic test might be performed in order to increase the evidences for accepting or rejecting some of these hypotheses. The information about what the correct diagnostic tests for each hypothesized disease is contained in the clinical algorithm of that disease, as we introduced in section 5.1 for the diagnosis of secondary hypertension. Our proposal is to transform these clinical algorithms into Test Selection Decision Tables (TSDT) and use the TSDTs of all the diseases involved in the hypotheses to determine the diagnostic test that will better contribute to the validation or refutation of these hypotheses.

The proposed methodology to obtain TSDTs from clinical algorithms starts with the identification of the diagnostic tests in the algorithm, and their codification in SNOMED-CT. See figure 5.4 for a summary of the methodology. For example, in the clinical

algorithm shown in figure 5.1 the identification of diagnostic tests in the clinical algorithm concludes with *plasma IGF 1 measurement*, *plasma growth hormone measurement*, *magnetic resonance imaging of head*, and *computed tomography of chest and abdomen*.

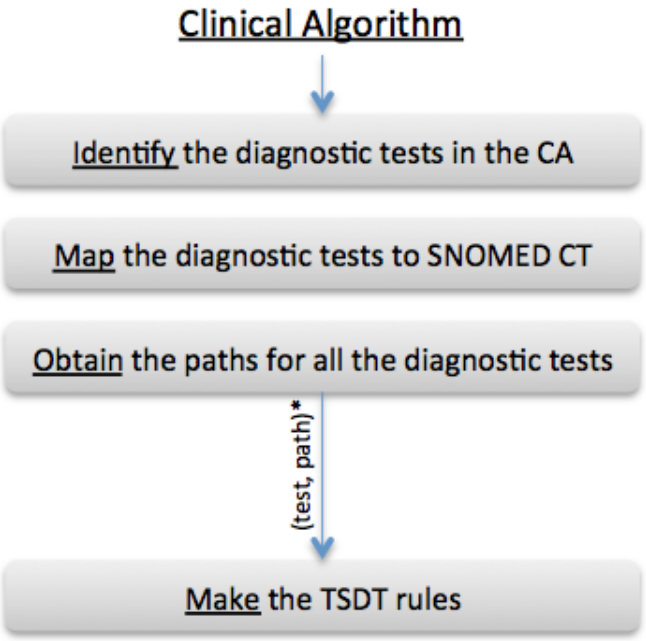


FIGURE 5.4: Methodology to make Test Selection Decision Tables

The hypothesis and all the tests recovered will define the condition stub of the TSDT, and the test alone will formalize the action stub. For each recovered test, we obtain all the possible paths in the clinical algorithm from the start to the end. Observe that, in medicine the diagnostic process can not last for extremely long periods or involve a huge number of steps. For this reason, the number and lengths of all the possible paths described in the clinical algorithms are contained to a manageable size. Each path is transformed into a new rule of the TSDT. In this rule, all the no-leaf tests that are in the path will appear in the condition entry with the value described in the corresponding connectors of the path, and the test at the end of the path (leaf tests) with the unknown value **?**.

In the action entry of the rule, the test at the end of the path is marked with an **X**, indicating its recommendation.

This methodology to construct TSDTs from clinical algorithms can be automated and it is easy to validate.

Figure 5.5 shows the TSDT obtained for acromegaly from the clinical algorithm presented in figure 5.1, after the application of the previously explained methodology.

Acromegaly	Yes	Yes	Yes	Yes
Computed tomography of chest and abdomen				?
Magnetic resonance imaging of head			?	No
Plasma IGF 1 measurement	?	Yes	Yes	Yes
Plasma growth hormone measurement		?	Yes	Yes
Computed tomography of chest and abdomen				X
Magnetic resonance imaging of head			X	
Plasma IGF 1 measurement	X			
Plasma growth hormone measurement		X		

FIGURE 5.5: Acromegaly Test Selection Decision Table

In the condition stub of the table, the first variable is the suspected disease (i.e., acromegaly) and the rest corresponds to the tests associated to this disease, after a SNOMED-CT codification. The action stub contains all the diagnostic tests found in the clinical algorithm.

The acromegaly algorithm in figure 5.1 produces the following paths:

1. Acromegaly $\rightarrow$ Plasma IGF 1 measurement
2. Acromegaly $\rightarrow$ Plasma IGF 1 measurement $\overset{Yes}{\rightarrow}$ Plasma growth hormone measurement
3. Acromegaly $\rightarrow$ Plasma IGF 1 measurement $\overset{Yes}{\rightarrow}$ Plasma growth hormone measurement $\overset{Yes}{\rightarrow}$ Magnetic resonance imaging of head
4. Acromegaly $\rightarrow$ Plasma IGF 1 measurement $\overset{Yes}{\rightarrow}$ Plasma growth hormone measurement $\overset{Yes}{\rightarrow}$ Magnetic resonance imaging of head $\overset{No}{\rightarrow}$ Computed tomography of chest and abdomen

These four paths are then transformed into the rules described in figure 5.5. For example, path 3 assigns value **Yes** to the hypothesis *acromegaly* but also to the diagnostic tests *plasma IGF 1 measurement*, and *plasma growth hormone measurement*, and **?** to the test *magnetic resonance imaging of head*. This means that in order to recommend the magnetic resonance imaging of head, the previous tests must have been done and a positive result obtained, and the hypothesis for acromegaly persist.

5.4 Making the Evaluation Decision Tables

The last tables in the diagnostic process are the Evaluation Decision Tables. See section 3.1.2.3 for more detail. This sort of tables are designed to decide whether a hypothesis is confirmed (**X** value), still active (**?** value), or rejected (**void** value), after the results of the diagnostic test.

The procedure to build this sort of tables from a clinical algorithm is similar to the procedure explained in the previous section. See a diagram in figure 5.6.

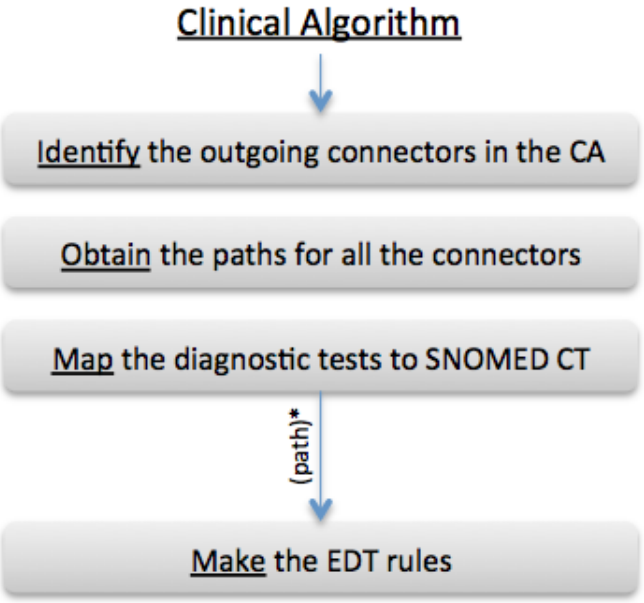


FIGURE 5.6: Methodology to make Evaluation Decision Tables

First, we have to detect all the connectors going out of a diagnostic test in the clinical algorithm. These connectors are labeled with their respective values. For each one of these connectors all the possible paths from the start to them are extracted. The path is concluded with the test acceptance or rejection following the last connector. The clinical tests contained in the paths are mapped into SNOMED-CT terms. Each path represents a rule in the EDT whose condition stub is composed of the diagnostic hypothesis and the observed diagnostic tests, and the action stub contains the hypothesis. Each rule will introduce as condition entries the values contained in the respective outgoing connectors of the tests in the path. The action entry of the rule will contain **X** if the last connector of the path drives to an acceptance of the hypothesis, a **void** value if it drives to a rejection of the hypothesis, and **?** if the connector points to a new test.

For acromegaly in figure 5.1 there are eight connectors emerging from a diagnostic test, and there is only one possible way. Therefore, the following paths are identified:

1. Acromegaly $\rightarrow$ Plasma IGF 1 measurement $\overset{No}{\rightarrow}$ **reject**
2. Acromegaly $\rightarrow$ Plasma IGF 1 measurement $\overset{Yes}{\rightarrow}$ Plasma growth hormone measurement
3. Acromegaly $\rightarrow$ Plasma IGF 1 measurement $\overset{Yes}{\rightarrow}$ Plasma growth hormone measurement $\overset{No}{\rightarrow}$ **reject**
4. Acromegaly $\rightarrow$ Plasma IGF 1 measurement $\overset{Yes}{\rightarrow}$ Plasma growth hormone measurement $\overset{Yes}{\rightarrow}$ Magnetic resonance imaging of head
5. Acromegaly $\rightarrow$ Plasma IGF 1 measurement $\overset{Yes}{\rightarrow}$ Plasma growth hormone measurement $\overset{Yes}{\rightarrow}$ Magnetic resonance imaging of head $\overset{Yes}{\rightarrow}$ **accept**
6. Acromegaly $\rightarrow$ Plasma IGF 1 measurement $\overset{Yes}{\rightarrow}$ Plasma growth hormone measurement $\overset{Yes}{\rightarrow}$ Magnetic resonance imaging of head $\overset{No}{\rightarrow}$ Computer tomography of chest and abdomen
7. Acromegaly $\rightarrow$ Plasma IGF 1 measurement $\overset{Yes}{\rightarrow}$ Plasma growth hormone measurement $\overset{Yes}{\rightarrow}$ Magnetic resonance imaging of head $\overset{No}{\rightarrow}$ Computer tomography of chest and abdomen $\overset{Yes}{\rightarrow}$ **accept**
8. Acromegaly $\rightarrow$ Plasma IGF 1 measurement $\overset{Yes}{\rightarrow}$ Plasma growth hormone measurement $\overset{Yes}{\rightarrow}$ Magnetic resonance imaging of head $\overset{No}{\rightarrow}$ Computer tomography of chest and abdomen $\overset{No}{\rightarrow}$ **reject**

Each path is then transformed in a rule of the EDT in figure 5.7. For example, the forth path about the connector with label 'yes' exiting from the diagnostic test *plasma growth hormone measurement* is transformed into the rule that fires when acromegaly is suspected (**Yes** value), plasma IGF 1 measurement and plasma growth hormone measurement have been done and positive results were obtained (**Yes** values), and a magnetic resonance imagining of head has not been done or unknown result (**?** value). In this case, acromegaly is still suspected (it is part of the diagnostic hypotheses) as it is indicated by the **?** value in the action entry. This is the forth rule in the EDT.

Acromegaly	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Computed tomography of chest and abdomen						?	Yes	No
Magnetic resonance imaging of head				?	Yes	No	No	No
Plasma IGF 1 measurement	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Plasma growth hormone measurement		?	No	Yes	Yes	Yes	Yes	Yes
Acromegaly		?		?	X	?	X	

FIGURE 5.7: Acromegaly Evaluation Decision Table

5.5 Conclusions

Differential diagnosis is a cognitive tasks of medical practice. In chapter 3 three complementary grouping decision tables were designed to gather the knowledge required along differential diagnosis. The way these computer structures have to be filled with the evidence-based knowledge contained in clinical practice guidelines remained an open issue that we addressed in this chapter with the introduction of a methodology that figure 5.8 depicts. For a target disease that we consider during DDx, we search a clinical practice guideline. Then the manual construction of a clinical algorithm by health care experts. This algorithm summarizes the decision procedure about the order in which diagnostic tests must be performed.

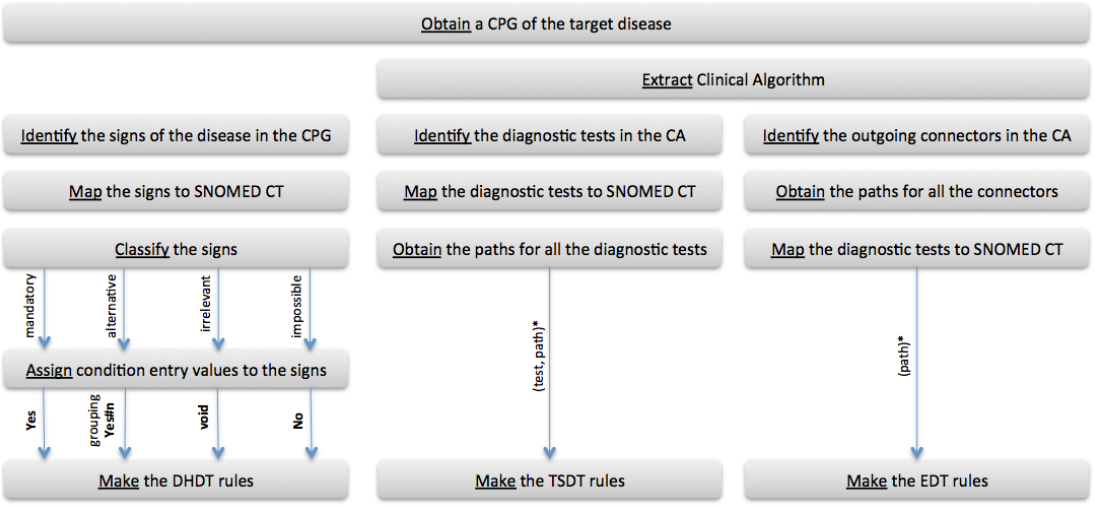


FIGURE 5.8: Methodology to implement DDx with decision tables

The signs and symptoms related to the identification of the disease are classified and organized as rules of a unique Diagnostic Hypotheses Decision Table. Simultaneously, the clinical algorithm is transformed into both a Test Selection Decision Table and a Evaluation Decision Table per disease. The TSDT helps to decide which diagnostic tests have to be performed at each particular moment, according to the clinical guideline. The EDT allows the modification of the suspected diseases.

The main applications of these methodologies and decision tables are as clinical decision support systems for differential diagnosis, or as training systems to improve differential diagnosis skills and the adherence of their decisions to the indications contained in the clinical practice guidelines. In chapter 7 we describe an application to the training of residents in the Hospital Clínic de Barcelona with regard to secondary causes of hypertension.

UNIVERSITAT ROVIRA I VIRGILI
USE OF DECISION TABLES TO MODEL ASSISTANCE KNOWLEDGE TO TRAIN MEDICAL RESIDENTS
Francis Real Vázquez

Chapter 6

Use of Treatment and Prognosis Decision Tables in Emergency Shock

Chapter 5 showed how to get the knowledge of medical differential diagnosis from clinical practice guidelines and clinical algorithms, and how to apply this knowledge to build decision tables through secondary causes of hypertension. Now, this new chapter shows in a similar way how to build treatment and prognosis decision tables. A new methodology for knowledge acquisition addressed to capture treatment and prognosis knowledge in decision tables. This methodology is applied to the clinical problem of treating and prognosticating seven different types of emergency shocks.

6.1 Emergency Shock

Shock is a common condition in critical care, affecting about one third of patients in the intensive care units (ICU). It is described as the clinical expression of circulatory failure that results in inadequate cellular oxygen utilization [204].

Some of the most common shocks are cardiogenic shock, anaphylactoid shock, cardiac tamponade, hemorrhagic shock, neurogenic shock, shock due to acute pulmonary embolism, and septic shock (see descriptions in appendix B).

Clinical reaction to shocks in ICU must be fast and precise because of the vital consequences on the patient and to prevent worsening organ dysfunction and failure. These reactions entail the combined application of ventilatory support, fluid resuscitation, and vasoactive agents [204].

All these actions have a direct and sometimes immediate consequence in some internal hemodynamic parameters. These parameters combine under the name of cardiac output, and they are: volemia (or the amount of fluids), heart rate (or frequency of heart beats), contractility (or heart strength), and vasoconstriction (or the weight of the vessels).

Since the most of these parameters are not directly observable by the physician who is attending the patient, medical decisions must be taken in terms of some observable vital signs such as: heart rate, central venous pressure, or arterial blood pressure.

The main objective of the treatment actions in the emergency shock is stabilize the hemodynamic parameters. This clinical actions can be continuous or discrete depending on the duration of they effects on the hemodynamic parameters. Continuous actions have an effect while they are applied, but the effect disappears when the action is disrupted. For example, drug infusions [205] such as dobutamine infusion are used when there is a cardiac decompensation and they are given in order to maintain a predictable pharmacodynamic action. When an infusion is started, the effect in the patient increases initially quickly, but then more slowly until a maximum effect. If the infusion supply is stopped, the effect decay in a short time. On the contrary, discrete actions have persisting effects over time. From a medical point of view, in an emergency context (few hours), we can assume that the effect of discrete actions persists along the whole patient treatment at the ICU. For example, one of the actions required for anaphylactoid shock is to take antihistamine. The antihistaminics differ in duration, but usually have an effect of 12 hours or more. This covers all the shock emergency treatment.

6.2 Making Treatment Decision Tables

Treatment Decision Tables (TDT) were introduced in section 4.1. These were computer structures able to contain the description of treatment actions as condition-action rules where the condition part represents both the current state of the patient and the current treatment (if there is any), and the action part describes the new actions to start. In order to obtain the knowledge necessary to fill TDTs, medical experts have to analyze the clinical practice guidelines of the disease whose treatment knowledge we want to include in the table. Usually, treatment sections in clinical practice guidelines are goal-oriented. This means that they point out a set of target parameters and variables, and then describe one or more strategies to achieve these targets.

To build a TDT we propose the new methodology that figure 6.1 summarizes.

In this methodology, we start selecting a CPG that describes the treatment of the target disease. Then, we extract the relevant signs an actions involved in the treatment. These

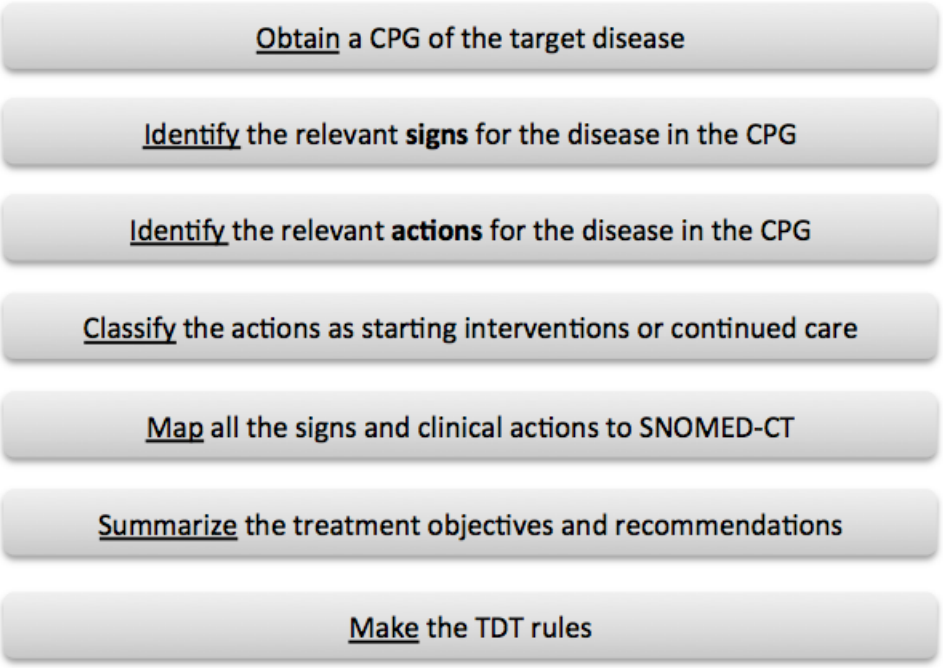


FIGURE 6.1: Methodology to make Treatment Decision Tables

are indicated in the corresponding treatment section of the guideline. These terms are mapped into SNOMED CT concepts. Actions can be classified as starting interventions or continued care. Starting interventions describe treatments that are conditioned to the patient state, while continued care describe actions resulting from the evolution of the patient when she is already receiving a concrete treatment, and a change in this treatment is required.

The text in the clinical practice guideline is interpreted by a medical expert, who summarizes the treatment objectives and the recommendations to achieve these objectives. Table 6.1 shows the summary of objectives and recommendations of the anaphylactoid shock, extracted for the guideline [206].

```
If patient not stabilized (arterial blood pressure <= 70)
     first, attempt it with 1-2 doses of adrenaline
     if (BP<=70), pharmacological treatment
          epinephrine infusion
          norepinephrine infusion
when (BP>70)
     antihistamine
     hydrocortisone (opt)
```

TABLE 6.1: Summary of Anaphylactoid Shock Treatment

These objectives and recommendations are used as a guide to build the rules. For example, if the arterial blood pressure is under 70 provide adrenaline bolus to the patient, one or two doses. For this recommendation we obtain to rules:

1. Anaphylactoid Shock + Epinephrine bolus = NO → Epinephrine bolus
2. Anaphylactoid Shock + BP <= 70 + Epinephrine bolus = 1 dose → Epinephrine bolus

These rules could be modified with the exceptions in the guidelines or the general knowledge of the experts. For example, although the general rule is to provide two bolus of epinephrine, if the patient is nearly stabilized after the first bolus, the second bolus can be replaced by other less aggressive treatment action.

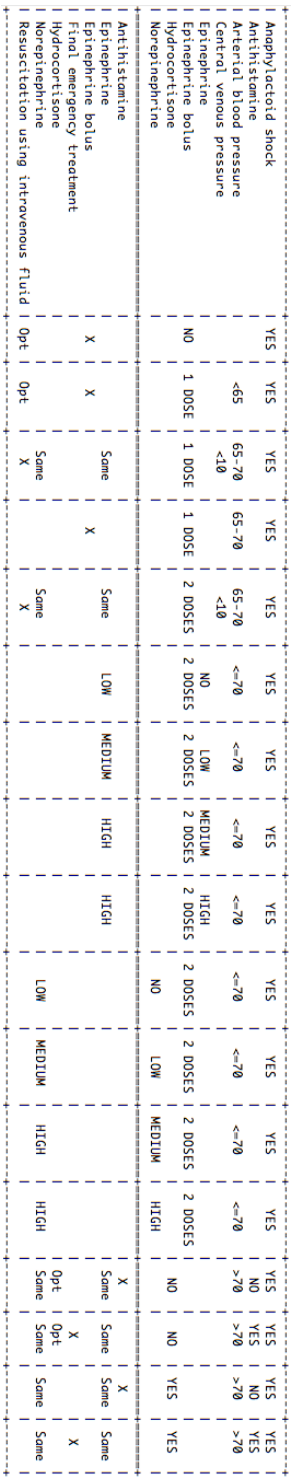
In the case of the treatment of shock in a UCI, medical experts from the Hospital Clínic de Barcelona identified seven signs and symptoms that the CPGs described as relevant for the management of shock [206–214]: heart rate, central venous pressure, arterial blood pressure, systolic blood pressure (SBP), diastolic blood pressure (DBP), finding of hematocrit, and superior vena cava oxygen saturation.

They also identified eighteen different treatment actions: antibiotic therapy, antihistamine, atropine, diuretic, epinephrine bolus, thrombolytic, hydrocortisone, insertion of intra-aortic balloon counterpulsation, pericardiocentesis, reperfusion (KT), resuscitation using intravenous fluid, transfusion of plasma, transfusion of red blood cells, vasodilators, dobutamine infusion, dopamine infusion, epinephrine infusion, and norepinephrine infusion.

Figure 6.2 shows an example of TDT obtained after the application of the above described methodology to the clinical practice guideline about the anaphylactoid shock [206]. The terms in the table are expressed under the SNOMED-CT codification.

The condition stub of this table has the target disease (i.e., anaphylactoid shock), the signs and symptoms marked in the CPG as objectives of this kind of shock (i.e., arterial blood pressure and central venous pressure), and some of the treatment actions to be considered in order to make continuous care decisions. In the action stub we can observe all the treatment actions identified in the CPGs. These can be discrete actions (i.e., antihistamine, epinephrine bolus, hydrocortisone, and resuscitation using intravenous fluids), continued actions (i.e., epinephrine infusion and norepinephrine infusion), and discharging actions (i.e., final emergency treatment).

This table captures the following treatment: If a patient with an anaphylactic shock arrives to the ICU, first we have to administer a bolus of epinephrine, and complement



Anaphylactoid shock																								
Antihistamine	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Arterial blood pressure		<65	65-70 <10	65-70	65-70 <10	<=70	<=70	<=70	<=70	<=70	<=70	<=70	<=70	<=70	<=70	<=70	>70	>70	>70	>70	>70	>70	>70	>70
Central venous pressure						NO	LOW	MEDIUM	HIGH	LOW	MEDIUM	HIGH	LOW	MEDIUM	HIGH	NO	NO	YES	NO	YES	NO	YES	NO	YES
Epinephrine	NO	1 DOSE	1 DOSE	1 DOSE	2 DOSES	2 DOSES	2 DOSES	2 DOSES	2 DOSES	2 DOSES	2 DOSES	2 DOSES	2 DOSES	2 DOSES	2 DOSES	2 DOSES	X Same	X Same	X Same	X Same	X Same	X Same	X Same	X Same
Epinephrine bolus																	Opt Same	X Opt Same	Opt Same	X Opt Same	Opt Same	X Opt Same	Opt Same	X Opt Same
Hydrocortisone																								
Norepinephrine																								
Epinephrine																								
Final emergency treatment																								
Hydrocortisone																								
Norepinephrine																								
Resuscitation using intravenous fluid	Opt	Opt	Same X	X	Same X																			
Antihistamine																								
Epinephrine						LOW	MEDIUM	HIGH	HIGH															
Epinephrine bolus										LOW	MEDIUM	HIGH	HIGH											
Final emergency treatment	X	X		X	Same									LOW	MEDIUM	HIGH	HIGH							
Hydrocortisone																								
Norepinephrine			Same X																					
Resuscitation using intravenous fluid	Opt																							

FIGURE 6.2: Anaphylactoid Treatment Shock Table

it with fluids if it is necessary (rule 1). After this first bolus, if the patient presents arterial blood pressure (ABP) under 70 bpm, then a second bolus of epinephrine is recommended (rule 2 and 4) or, alternatively, use fluids if the ABP is near 70 bpm and the central venous pressure (CVP) is not high (rule 3). Once a second bolus is taken, it is dangerous to supply new boluses. If the ABP continues under 70 bpm, it is possible

to provide fluids if the ABP is near 70 bpm and the CVP is not high (rule 5), or to use infusions of epinephrine (rules 6 to 9) or norepinephrine (rules 10 to 13). The infusions start with low dosage that increases until the patient is stabilized. When the patient is stabilized with ABP over 70 bpm, it is necessary to give antihistamine, and possibly hydrocortisone, before the patient is discharged (rules 14 to 17).

The number of possible ways to combine treatment actions are higher than the number of combinations of diagnostic test in chapter 5. This means that more rules, bigger tables, and less legibility are expected for TDTs than we had for TSDTs or EDTs.

For example, the treatment of APE RV Cardiogenic shock (see appendix B) it is not more complex than the treatment of the Anaphylactoid shock in number of steps, but the APE RV Cardiogenic shock has more treatment alternatives. This makes decision table of APE RV Cardiogenic shock to contains 85 rules, in front of 17 of the anaphylactic shock. The diagnostic clinical algorithms, same as the anaphylactoid shock treatment, are direct and have few alternatives. The number of rules of APE RV Cardiogenic shock could be higher without the mechanisms to reduce the number of combinations used in Treatment Tables, i.e. the use in TDTs of more complex types of value generalizations (e.g. ranges or dosage values) in the condition entries that in diagnostic tables.

6.3 Making Prognosis Decision Tables

Knowledge about prognosis is very difficult to obtain. Clinical practice guidelines do not provide this kind of knowledge, but it is possible to calculate it from clinical records about representative patients, or obtain it from medical experts.

In spite that we could not obtain data about clinical records from the health care centers which we collaborated with, namely the Hospital Clínic de Barcelona and SAGESSA, we could count with the appreciated collaboration of senior medical experts from one of these centers.

Our proposed methodology to capture prognostic knowledge to be represented in Prognostic Decision Tables (PDT) is based on the use of clinical surveys. Clinical surveys or questionnaires are frequent in medical practice as we stated in section 2.1.4. These questionnaires are addressed to medical experts whose medical knowledge about prognosis we want to elicit. The basic unit of these surveys is the question, and the questions that we defined for prognosis are of the form:

> If a patient with a disease d has a sign s with value v_s^0 and a clinical action a is performed, what is the expected new value v_s^1 for that sign?

For example, in the UCI consideration of shocks, one of the questions in the questionnaire was: If a patient with anaphylactic shock has a systolic blood pressure of *60 mm Hg* and she takes *a dose of antihistamine*, which is the expected new value for systolic blood pressure?

For each question in the survey we could capture units of prognostic information of the form (d, s, v_s^0, a, v_s^1) that could be stored in Prognostic Decision Tables after a processing.

Figure 6.3 shows a description of the phases to make the surveys and to use their results to build prognosis decision tables.

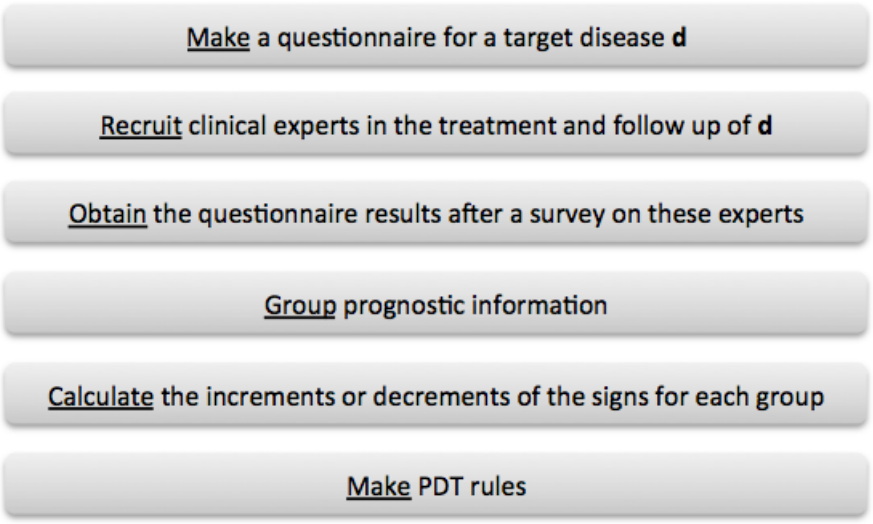


FIGURE 6.3: Methodology to make Prognosis Decision Tables

First, we select a target disease **d** from the set **D** of all the diseases studied. We obtain the set **S** composed by all the relevant signs and symptoms of at least one of the diseases of **D**. We also obtain the set **A** composed by all the therapeutic actions recommended by the CPGs for one or more diseases in **D**. For each sign $\mathbf{s}\in\mathbf{S}$, we select different initial values $\mathbf{v_s^0}$. These sign values can be classified into ranges (i.e. low, normal and high). For each range, a random number of values between 3 and 6 are selected. These values are taken at random from the ones possible for the respective range For each initial value, we construct a question in the survey in the form previously described.

For example, for the disease Anaphilactoid shock (**d**) with the treatment action antihistamine (**a**) , we want to assess the systolic blood pressure (**s**) . The range values for systolic blood pressure are: low (from 40 mmHg to 85 mmHg), normal (from 85 mmHg to 140 mmHg) and high (from 140 mmHg to 280 mmHg). Values under 40 mmHg or over 280 mmHg are dangerous. We select different initial values ($\mathbf{v_s^0}$) for systolic blood pressure: {45, 55, 70, 80, 90, 100, 115, 130, 145, 160, 190, 220, 260} mmHg , where values {45, 55, 70, 80} mmHg correspond to the low range, values {90, 100, 115, 130} mmHg

correspond to the normal range, and values {145, 160, 190, 220, 260} mmHg. For the first initial value 45 mmHg we obtain the question: *(anaphilactoid shock, antihistamine, systolic blood pressure, 45 mmHg)* that corresponds to the range low of systolic blood pressure.

The methodology continues with a survey that is conducted for a number of representative medical experts that are senior in the treatment and the follow up of the target disease. With the results, we group the questions about the same sign **s**, clinical action **a**, and initial value $\mathbf{v_s^0}$ in the same sign range : (**d**, **s**, $\mathbf{v_s^0}$, **a**, v_s^1). For each group of range questions, we calculate the mean value of the differences ($\mathbf{v_s^1} - \mathbf{v_s^0}$); i.e., the increment or the decrease that the sign has when the action is applied. With this information, we obtain the rules in the form:

disease + action + { sign = range } → { average of ($v_s^1 - v_s^0$) in range}

For example, if the answers of a survey are:

```
(Anaphilactoid, antihistamine, SBP, 45 mmHg) ->  40 mmHg
(Anaphilactoid, antihistamine, SBP, 55 mmHg) ->  48 mmHg
(Anaphilactoid, antihistamine, SBP, 70 mmHg) ->  62 mmHg
(Anaphilactoid, antihistamine, SBP, 80 mmHg) ->  72 mmHg
```

then we obtain the rule:

Anaphilactoid shock + Antihistamine + SBP = low → -7.0

In out study for the prognosis of evolution of patients with shock, we provided more than one thousand questions to emergency experts to cover all the possibilities of combinations of diseases, signs, and actions. Figure 6.4 shows an extract of the Anaphylactoid prognosis decision table built with this methodology that originally counts with 242 rules. Every PDT is focused in one of the seven different shocks studied, and contains between 200 and 300 prognostic rules.

In this table we can see how the actions have a different effect depending on the initial range of values. For example, antihistamine (in anaphylactic shock) provokes a bigger change in high values than in low values of systolic blood pressure, as rules 1 to 3 show. On the contrary, the same action has a larger effect in heart rate if the patient has medium values, as rules 7 to 9 indicate.

Anaphylactoid shock	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Antihistamine	x	x	x	x	x	x	x	x	x	x	x	x			
Atropine															
Central venous pressure										<3	3-8	>8			
Diastolic blood pressure				<50	50-90	>90									
Diuretic															
Dobutamine															
Dopamine															
Epinephrine															
Epinephrine bolus													x	x	x
Finding of hematocrit															
Heart rate							<60	60-100	>100						
Hydrocortisone															
Insertion of intra-aortic balloon counterpulsation															
Norepinephrine															
Pericardiocentesis															
Reperfusion															
Resuscitation using intravenous fluid															
Superior vena cava oxygen saturation															
Systolic blood pressure	<85	85-140	>140										<85	85-140	>140
Thrombolytic															
Transfusion of plasma															
Transfusion of red blood cells															
Vasodilators															
Central venous pressure										-0.1	-0.9	-1.8			
Diastolic blood pressure				-10.0	-10.0	-10.0									
Finding of hematocrit															
Heart rate							+6.0	+10.0	+2.0						
Superior vena cava oxygen saturation															
Systolic blood pressure	-7.0	-10.0	-23.3										+25.0	+25.0	+23.3

FIGURE 6.4: Extract of Anaphylactoid Prognosis Decision Table

6.4 Simulating Emergency Shock Patients

PDTs can also be used for patient simulation. However, this simulation requires a patient model (PM) in order to implement realistic simulations, as it was explained in section 4.3. Broadly speaking, this PM defines the initial state of the patient at the admission time, and her global parameters such as age, gender, weight and height, and

primary and secondary diseases. But it also describes the standard ranges of signs for that patient, and her sensitivities to clinical actions. PMs allow us to have different sorts of patients that can react differently in front of the same clinical actions contained in the PDTs. This causes prognosis to depend not only on the current condition of the patient and the clinical actions performed, but also on the patient herself, as it is observed in real clinical practice.

Defining a new PM consists on filling all the fields of the PM with the information about the case we want to simulate (see 4.3). The result is a virtual patient ready to be subject to the treatment recommendations suggested by PDTs.

In our test domain of seven different sorts of shocks, we developed a case-base with 51 PMs, representing patients with different shocks. These PMs were based on real patients observed at the ICU of the Hospital Clínic de Barcelona. The Emergency Department of this hospital represented the PM clinical description, history, initial values for signs and symptoms, personalized normality ranges of signs, and possible drug resistances or sensitivities.

Table 6.2 shows an example of a simulated patient with AMI LV Cardiogenic Shock. This patient represents a 67-year woman with obesity as a risk factor.

In section `VITAL SIGN RANGES`, her parameters of normality of the shock vital signs are defined normal, with exception of systolic blood pressure which is slightly high due to obesity (i.e., 50-95-160-290 mmHg values in comparison to normal SBP values 40-85-140-280 mmHg).

In section `ACTION SENSITIVITIES`, this patient is declared to have a high resistance to thrombolytic drugs, since they only have a 10% effect in this patient, when their normal effect should be 100%.

Section `INITIAL SIGNS`, describes the patient condition at the time she was admitted. We observe that all her initial sign values were normal, except systolic blood pressure, that was 76.0 mmHg, low.

6.5 Simulating Long-Term Treatment of Emergency Shock

The combined action of PDTs and PMs allows the simulation of a case personalized for a concrete patient. This simulation can be short-term (i.e., punctual in time), or long-term. In this work, short-term treatment simulation refers to the application of the treatment knowledge contained in the PDT to a patient who comes modeled by the PM. The result is a recommended personalized treatment described by a set of clinical

```
Patient ID : 33
AGE: 67            WEIGHT: 82 Kg
SEX: female        HEIGHT: 1.66 m

PRIMARY DIAGNOSIS: AMI LV Cardiogenic Shock
SECONDARY DIAGNOSES: Obesity

DESCRIPTION:
After informing the patient that her husband has a serious
disease with fateful prognosis, the patient begins oppressive
central chest pain radiating to the back and both arms.
She presents pallor and profuse sweating.
She refers feeling of dizziness.
Starts oxygen therapy with nasal specs at 2 L / min.
```

VITAL SIGN RANGES:	MIN	LOW	HIGH	MAX	UNITS
Diastolic blood pressure	20.0	50.0	90.0	130.0	mmHg
Systolic blood pressure	50.0	95.0	160.0	290.0	mmHg
Arterial blood pressure	26.6	61.6	106.7	180.0	mmHg
Heart rate	25.0	60.0	100.0	153.0	beats/min
Central venous pressure	0.0	3.0	8.0	20.0	cmH2O
Sup. vena cava ox. satur.	50.0	65.0	85.0	88.0	%
Finding of hematocrit	15.0	35.0	45.0	60.0	%

```
ACTION SENSITIVITIES:
Thrombolytic  ==>  10 %

INITIAL SIGNS:
  <Central venous pressure = 4.0>
  <Systolic blood pressure = 76.0>
  <Diastolic blood pressure = 50.0>
  <Finding of hematocrit = 36.0>
  <Heart rate = 96.0>
  <Superior vena cava oxygen saturation = 62.0>
```

TABLE 6.2: Example of patient for AMI LV Cariogenic Shock

actions, and an eventual forecast of the evolution of the patient, once these clinical actions are applied.

On the other hand, long-term treatment simulation in this thesis alludes to the automated combined use of PDTs and PMs for simulating treatment sequences that can last for several treatment steps or even the whole virtual patient treatment. In order to achieve automated long-term simulations, we propose an iteration system implementing the long-term treatment decision system designed in section 4.4.

Both sort of simulations are represented by figure 6.5.

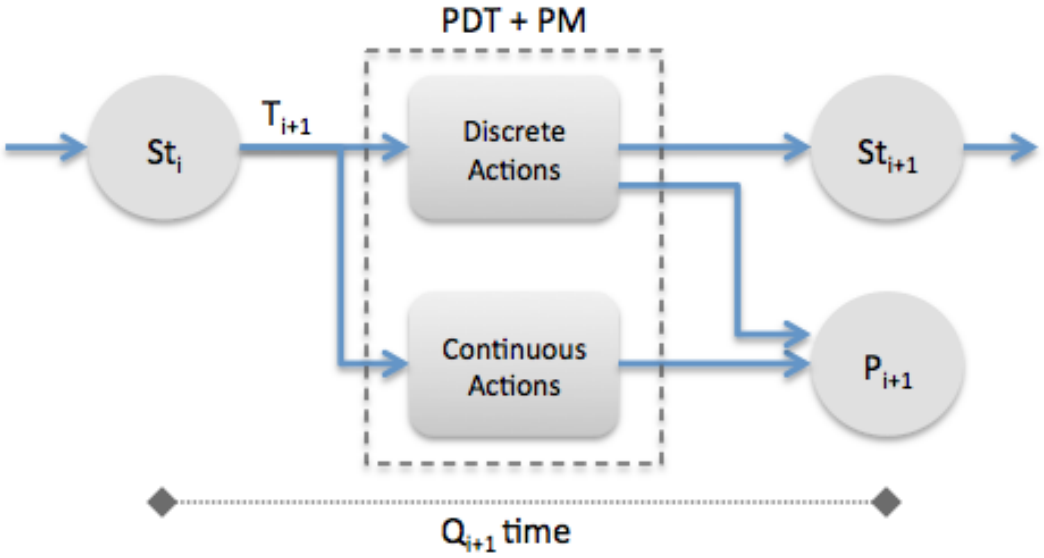


FIGURE 6.5: Iterations with PDT and PM

There, the current state of the patient (i.e., St_i) is treated with the set of clinical actions T_{i+1} provided by the TDT. The effects of these actions in a concrete patient are then calculated according to the PDT and modified with the patient sensitivities and normality ranges contained in the PM. In a short-term treatment simulation the patient will evolve to state $St_{i+1} \cup P_{i+1}$. Contrarily, in a long-term treatment simulation the initial state of the patient is taken from the section `INITIAL SIGNS` of the PM, and introduced as St_0. The patient is then treated with the set of clinical actions T_1 provided by the PDT for this patient. This set may contain both, continuous actions and discrete actions. Recall, that continuous actions are those which have an effect in the immediate next state of the patient, and discrete actions those whose effect persists for the whole treatment.

The application of both continuous and discrete actions in T_{i+1} to the patient transforms (some of) her vital signs obtaining a new patient state P_{i+1} which can be observed by the users of the simulator. This process is the result of applying PDT and PM to the pair (St_i, T_{i+1}). However, another internal state St_{i+1} of the patient is calculated by the simulator. This new state describes the vital signs of the patient from a global perspective required by the simulator to continue with a new iteration of the simulation process, after a time Q_{i+1}. This internal state is the result of applying PDT and PM to the pair (St_i , D_{i+1}), with $D_k = \{A_j^k \in T_k : A_j^k$ is a discrete action$\}$. St_{i+1} is hidden to the external users of the simulator.

After this treatment step and once a Q_{i+1} time has passed, the new state St_{i+1} can be considered the current state of the patient and the process can be repeated again. In order to calculate Q_{i+1} we take the largest response time of the clinical actions in T_{i+1}.

In our simulation of shocks, continuous clinical actions are dopamine infusion, dobutamine infusion, norepinephrine infusion, and epinephrine infusion. The rest of actions (antihistamine, hydrocortisone, epinephrine bolus, atropine, diuretic, fluid infusion, plasma transfusion, red blood cell packed, vasodilators, thrombolytic therapy, reperfusion (KT), pericardiocentesis, and insertion of intra-aortic balloon counterpulsation) are discrete, being the last four procedures.

6.6 Conclusions

Treatment is a complex decision task that involves plenty of considerations. Clinical practice guidelines provide different possibilities where, usually, the best one depends on the patient and her evolution.

In chapter 4, we presented the Treatment Decision Tables (TDT) as a way to describe the treatment options and the PDT as a possibility to explain the patients evolution. In this chapter we summarized step by step how to fill the tables using the emergency shock as a reference.

To obtain TDTs, medical experts construct global algorithms to represent the knowledge from the CPGs. By means of a new provided methodology, these algorithms are translated in general rules that are the base to create the TDT.

The PDT are constructed with information provided by surveys. For each set of diseases, a questionnaire was designed and provided to medical experts. The objective of this survey is to abstract the medical expertise to foresee the patient changes when different therapeutic actions are applied. This chapter explains how to structure and conduct the surveys and how to use the results to do the decision tables.

To personalize the knowledge of the PDT in the shock, a case base of 51 patients was created. This case base contains information of types of patient in the PM format.

Finally, we described our implemented process to simulate short- and long-time treatments by combining TDTs, PDTs and PMs.

UNIVERSITAT ROVIRA I VIRGILI
USE OF DECISION TABLES TO MODEL ASSISTANCE KNOWLEDGE TO TRAIN MEDICAL RESIDENTS
Francis Real Vázquez

Chapter 7

Using Decision Tables to Train Residents

In chapter 3 we introduced several new structures of decision tables to host the knowledge required to support differential diagnosis. Later on, in chapter 5 we showed how to implement these tables with the knowledge extracted from CPGs and clinical algorithms and how to apply a new knowledge acquisition methodology to obtain the rules that will be incorporated to the diagnostic decision tables in order to allow the diagnosis of secondary causes of hypertension.

Continuing with this work, the current chapter introduces an experiment in which diagnostic decision tables were tested as part of a learning application tool that was used to train residents at the Hospital Clínic de Barcelona. To do this, an on-line incremental knowledge-based training tool was created using the decision tables obtained in chapter 5. The training tool was oriented to medical residents with different specialties and experience. The objective of the application was to improve the adherence of health care professionals to official CPGs in DDx processes.

This chapter details the design and the construction of this tool, the development of the experiment, and the analysis of the results.

7.1 The DDx Training Tool

In section 2.2.5, the general architecture of a training tool was presented. Now, following the last model in figure 2.5(b), we design an Internet tool for the training of hospital residents in the DDx process.

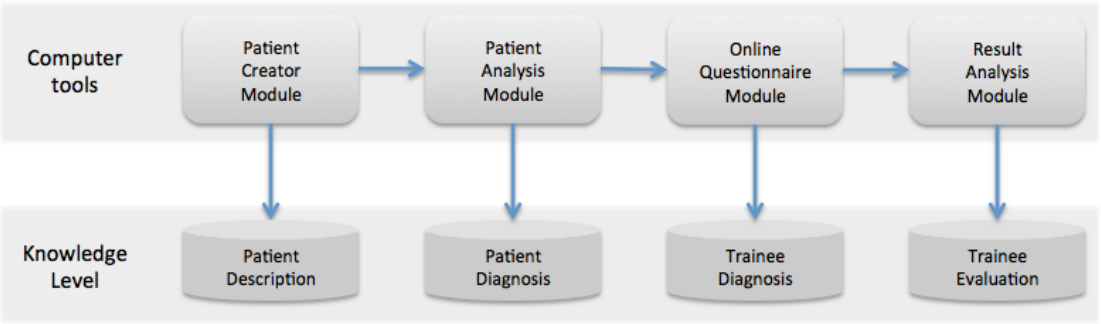


FIGURE 7.1: Diagnostic Computer Tools

The objective of the application is to train clinicians to manage, assimilate, and correctly apply the knowledge contained in the guidelines of the most frequent causes of AH.

To this end, we designed a set of computer tools that composed our training system. These computer tools are: a patient creator module, a patient analysis module, an online questionnaire module, and a result analysis module. Figure 7.1 shows a schematic representation of the architecture of the training system showing these computer tools and the sort of knowledge that they generate.

The Patient Creator Module (PCM) develops a case base containing clinical cases representing patients with none, one or more secondary causes of hypertension. The cases include the patient diseases, her sign and symptoms, and the patient response to all the possible diagnosis tests, in accordance to the sort of patient and health condition. This information about cases can be subject to controlled noise so that realistic situations could be simulated.

The Patient Analysis Module (PAM) simulates the process detailed in section 3.2. This module analyzes each patient case in the case base with the diagnostic decision tables for secondary causes of hypertension and it obtains the initial diagnostic hypotheses, the steps suggested by the clinical guidelines, and the final diagnosis.

The Online Questionnaire Module (OQM) manages the interaction with the medical residents that will be trained with the system. It proposes cases from the patient description case base and obtains the clinical responses from the residents. When a user of the tool diagnoses a case, the OQM provides a feedback with the diagnostic process followed as compared to the diagnostic process expected, in order to permit adherence analysis to standards.

The Result Analysis Module (RAM) analyzes the answers of the user with the help of the diagnostic decision tables and evaluates the responses. This evaluation is the basis to analyze the quality of the training system.

These modules are explained with more detail in next subsections.

7.1.1 The Patient Creator Module

The first step in the training of residents is to create a case base with simulated patients. In this work a base with 30 different patients was created.

The patients are obtained in a semi-random way. First, the target diseases to be part of the differential diagnosis process are introduced manually. This allows us to control the types of patient that will be created and to balance the amount of patients according to these diseases. With the PCM it is possible to define patients with none, one, or more than one diseases among the ones defined as secondary causes of hypertension. Cases with hypertension but with none secondary cause of hypertension are supposed to suffer from essential hypertension. Patients with one secondary cause of hypertension have to be diagnosed of this disease that is causing an hypertensive status in a patient. Essential hypertension should not be diagnosed. Multimorbid patients can have several secondary causes of hypertension simultaneously.

Although the secondary causes of hypertension do not have a high prevalence and the coexistence of more that one simultaneous secondary cause is an uncommon situation, there are clinical evidences of cases with more that one secondary causes in the same patient, for example adrenal Cushing's syndrome and glomerulonephritis [215], pheochromocytoma and hyperparathyroidism [216, 217], or acromegaly, hyperparothiroidism and pheocromocitoma [218]. Moreover, sometimes the presence of one secondary cause favors the presence of other secondary cause. For example, the sleep apnea syndrome is considered a common disorder associated to patients who suffer from acromegaly [219]. In response to these situations, the learning tool included some multimorbid patients.

Table 7.1 shows the distribution of diseases in the 30 patients of the AH case base. This table also shows the number of co-morbid cases.

Once the set of diseases of a case are defined, the process of identifying her signs and symptoms starts. This process associates each patient a set of signs and symptoms according to the patient's diseases.

To do this, the PCM applies a three-step process that begins with a selection of all the signs related to the patient diseases. This information is contained in the Diagnostic Hypothesis Decision Table (DHDT). In a second step the system deletes some of these signs at random, leaving at least one for each disease so that all the existing diseases can be suspected. This information is contained in the condition stub of the DHDT with **yes** or $\mathbf{Y\#}\boldsymbol{n}$ in the corresponding condition entry of the rules about the disease. The

Diagnosis	No of Cases
Acromegaly	4
Adrenal Cushing's syndrome	5
Coarctation of aorta	4
Essential AH	5
Glomerulonephritis	4
Hyperparathyroidism	4
Pheochromocytoma	5
Renovascular hypertension	4
Sleep apnea	4
Co-Morbidities	
none (essential)	5
1	18
2	5
3	2

TABLE 7.1: Distribution of diseases in the patient description case base.

deleted signs represent missing signs (i.e., signs that are related to the disease but that may not be developed in a concrete patient), or hidden signs (i.e., signs developed by the patient but not detected at that moment).

The final step is to randomly add none, one, or more new signs and symptoms that are not related to the diseases of the case. These added signs represent sings related to other conditions different to the patient diseases whose purpose is to add complexity to the diagnostic process. They are called noise signs. Noise signs can be found in the DHDT condition stub of the rules that are not describing the diseases of the case.

Table 7.2 shows the distribution of the signs in the patient description case base after apply the three-step process previously described to the 30 selected patients of secondary causes of hypertension.

When the signs and symptoms of all the cases have been determined, the PCM calculates the answer that each diagnostic test should provide to each case. This is obtained from the Evaluation Decision Table (EDT). All the relevant tests related to the patient diseases are selected and used to identify the value necessary to accept that disease. All the other tests receive values to discard the rest of diseases. The rationale to anticipate the response of each case to all the tests is to homogenize the answers of the training tool to the trainees. In other words, the cases act in a same way for all the users.

Diagnostic tests are not infallible. Sometimes they fail. In order to simulate this fact, all tests are given a confidence level among the values `Very Low`, `Low`, `Average`, `High`, or `Very High`. These values describe the possibility of the tests not to fail and produce a correct answer, the first time they are performed on a patient. In our studies we consider

Signs & Symptoms	Cases	Signs & Symptoms	Cases
Abdominal pain	6	Heart murmur	3
Acquired skeletal deformity	8	Hirsutism	4
Acute necrosis	5	Hypertensive disorder	30
Age more than 50 years	1	Insomnia	2
Amenorrhea	4	Muscle weakness	9
Anxiety	6	Nausea and vomiting	9
Apnea	3	Oliguria	5
Bone pain	6	Polyuria	3
Constipation	4	Progress satisfactory	1
Edema	6	Skin striae	4
Epistaxis	6	Snoring	5
Excessive sleepiness	6	Tachycardia	5
Fever	2	Unequal pulse	5
Headache	4		

TABLE 7.2: Distribution of signs in the patient description case base.

that diagnostic tests cannot fail more than once. For this reason, if a test is repeated, the result must be correct.

Table 7.3 shows the confidence level of each diagnosis test involved in the differential diagnosis of secondary causes of hypertension, according to the bibliography looked up by medical experts from the Hospital Clínic de Barcelona.

Figure 7.2 summarizes the PCM stages to construct a case base with simulated patients with diseases, signs and symptoms and their response to diagnosis tests.

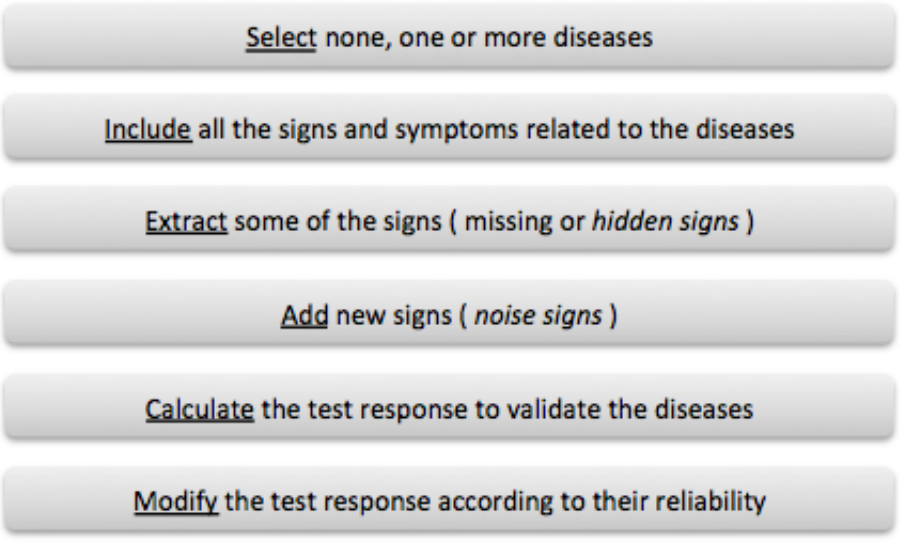


FIGURE 7.2: Summary of steps of the Patient Creator Module

Test	Confidence
Blood calcium level	Very High
Phosphorus measurement	Very High
Plasma growth hormone measurement	Average
Urine blood test	Low
Plasma IGF 1 measurement	Average
Urine protein test	Low
Plasma parathyroid hormone level	Average
Kidney biopsy	Very High
Measurement of hydrocortisone in saliva	Very Low
Plasma cortisol measurement	Average
Doppler studies	Average
Metanephrines	Low
Polysomnography	High
Clinical immunological test	Average
Magnetic resonance imaging	High
Magnetic resonance imaging of head	High
Kidney imaging, SPECT technique	High
Computed tomography of chest and abdomen	High
Dexamethasone	Average
Immunological treatment	Very Low

TABLE 7.3: Confidence level of diagnostic tests provided by medical experts

Table 7.4 shows an example of patient obtained after the application of the PCM. First, we choose the patient to be monomorbid with disease *hyperparathyroidism*.

Diseases:		
Hyperparathyroidism		
Signs:		
Hypertensive disorder		
Excessive sleepiness		
Polyuria		
Snoring		
Procedures:	Test-1	Test-n
Plasma parathyroid hormone level	True	True
Blood calcium level	True	True
Polysomnography	False	False
Metanephrines	True	False
...		

TABLE 7.4: Example of patient obtained after the application of PCM

The system obtains the signs for *hyperparathyroidism*. The only mandatory sign in table 5.3 is *hypertensive disorder*, and the alternative signs are *abdominal pain*, *muscle*

weakness, *nausea and vomiting*, *polyuria*, and *excessive sleepiness*. Among them *abdominal pain*, *muscle weakness* and *nausea and vomiting* are considered hidden signs by the PCM, and removed. The PCM also selects *snoring* as a noise sign, and it is included in the patient description. Consequently, our case will contain one mandatory sign (*hypertensive disorder*), two out of the five alternative signs (*excessive sleepiness* and *polyuria*), and one noise sign (*snoring*).

The tests related to *hyperparathyroidism* (see appendix A) are *plasma parathyroid hormone level* and *blood calcium level*. Both need to be true in order to confirm the disease. The rest of tests will obtain a negative result for hyperparathyroidism. As table 7.3 shows, *plasma parathyroid hormone level* has a confidence level `Average` and *blood calcium level* `Very High` which compute correct results the first time they are applied (i.e., `True` value in column Test-1) and the subsequent times (Test-n). `Metanephrines` however should provide a negative value, but the PCM determines that the first time this test is ordered for this patient a positive results will be obtained.

7.1.2 The Patient Analysis Module

The objective of the PAM is to simulate a gold standard physician that works as the clinical guidelines suggest. This tool analyzes each case in the patient description case base with the help of the diagnostic decision tables.

For each case, PAM selects the suspected diseases using the DHDT obtained in section 5.2. The result is a set of diagnostic hypotheses for the case. For example, the case in table 7.4 will produce the hypotheses *hyperparathyroidism* (due to signs *excessive sleepiness* and *polyuria*), and *sleep apnea* (due to signs *excessive sleepiness* and *snoring*).

With the set of hypotheses, the tool applies the methodology explained in section 3.2, using the tables obtained for each disease in chapter 5, to combine tests and evaluations until a final diagnosis is reached.

In the previous example, the TSDT of *hyperparathyroidism* determines that the test recommended is the *blood calcium level*, and the test recommended for *sleep apnea* is the *polysomnography*.

In the example, the results obtained the first time that the tests are performed on the patient are <*blood calcium level* , TRUE> and <*polysomnography* , FALSE>. These results are a new knowledge that can be added to the case description and the system can use this knowledge in future iterations.

On one hand, the EDT of *hyperparathyroidism* determines that the system has to maintain the hypothesis, and the TSDT suggests a new test: plasma parathyroid hormone level. On the other hand, the EDT of *sleep apnea* rejects the hypothesis and sleep apnea is considered never more.

In the next iteration, PAM generates new knowledge as a result of the test <*Plasma parathyroid hormone level* , TRUE>. With this new information, the EDT of *hyperparathyroidism* accepts the hypothesis.

As all the hypotheses have been accepted or rejected, the system can conclude a final diagnosis: the case is considered to have *hyperparathyroidism*.

Table 7.5 summarizes the above process of application of PAM to the patient example in table 7.4.

```
Initial suspects:
     Hyperparathyroidism
     Sleep apnea
Iterations:
     Iteration 1:
          Suspects:
               Hyperparathyroidism
               Sleep apnea
          Tests:
               Blood calcium level
               Polysomnography
     Iteration 2:
          New signs:
               Blood calcium level - TRUE
               Polysomnography - FALSE
          Suspects:
               Hyperparathyroidism
          Tests:
               Plasma parathyroid hormone level
     Iteration 3:
          New signs:
               Plasma parathyroid hormone level - TRUE
          Suspects:
               Hyperparathyroidism
Final diagnosis:
     Hyperparathyroidism
```

TABLE 7.5: Example of patient diagnosis after the application of PAM

For each case in the case base, this PAM process is followed. All the information generated along the process is used as feedback to be shown to the users of the training system.

7.1.3 The Online Questionnaire Module

The objective of the OQM is to allow the interaction of the trainees with the training system through the Internet. This interaction is possible because OQM works with an Internet browser interface that allows the users to have full-time remote access to the training system. It provides each physician with a different login and password for individual secure access.

Figure 7.3 shows the user interface to interact with the training system.

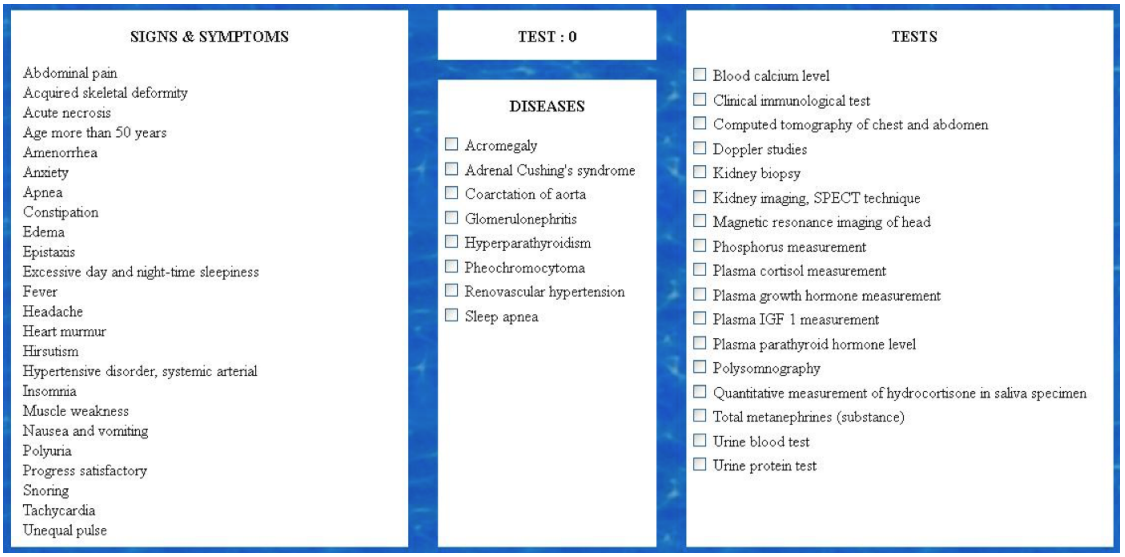


FIGURE 7.3: Hypertension Test Interface

The left column shows the signs and symptoms of the case, and also the additional information obtained by the tests in each iteration. In this figure, all the possible signs considered in the experiment are deployed. The central column contains a form to suggest possible diseases and the right column contains a form to indicate the possible diagnostic tests.

This module is designed to repeat a learning loop for each case in the training process. During the learning loop (see algorithm in figure 7.4) the user is informed of the signs and symptoms of the case.

After the observation of the signs and symptoms of the case, the trainee is allowed either to provide a final diagnosis or a hypothesis for the causes of AH (decision D1). Hypotheses need to be confirmed with additional diagnostic tests that the user is asked to order (decision D2). The system recovers the results of the tests from the available information about the case in the patient description case base, and shows this information to the user, before a new learning loop is started.

```
for each case in the patient description case base
   show the signs and symptoms of the case
   loop until there are no test to perform or final diagnosis confirmed
      wait for the user to choose a hypothesis (D1)
      wait for the user to order some diagnostic tests (D2)
      recover the results of the ordered tests from the case base
   end loop
   show user's diagnostic process confronted with guidelines' process
end for
```

FIGURE 7.4: Training Algorithm

At the end of the process, either if a final diagnosis is indicated or the case is dismissed, the system shows the whole differential diagnosis process followed by the user, confronted to the differential diagnosis process suggested by the decision tables obtained in the previous module.

7.1.4 The Result Analysis Module

The results obtained after the use of the OQM can be deeply analyzed with the help of the result analysis module (RAM). This module has two main objectives. On the one hand, it assesses the learning progress of the students and, on the other hand, it evaluates the performance of the system.

The differential diagnostic abilities that the RAM checks for the trainees are:

- The *initial hypotheses*, or the capacity of suspecting the correct diseases during the first encounter with the patient. Only the symptoms at the patient's arrival were known, before any test could be requested.
- The *final diagnosis*, or the capacity to reach a correct diagnosis after the DDx process. The commitment is not to affect the capability of the health care professionals to reach a correct diagnosis but to reeducate them to follow the DDx procedures contained in the CPGs.
- The *adherence to the diagnostic suspicions* in the CPGs, or the capacity of suspecting the diseases for which the CPGs had some evidence, along the DDx process (decision D1 in the algorithm in the training algorithm). Upon the observation of the available findings for the patient, the physician is expected to suspect the diseases whose CPGs contain enough evidence supporting these findings.
- The *adherence to tests requirements* in the CPGs, or the physicians' capacity to request, at the right time, the diagnostic tests that the CPGs recommend (decision D2 in the algorithm in the training algorithm).

The initial hypotheses and the final diagnosis can be evaluated by direct comparison of the trainee results with the patient description database. The diagnostic adherence and the test adherence are assessed with the evaluation of each case, step by step, with the decision tables.

Five different studies were decided: a global study with all the physicians and all the patients, separate studies according to the profile of the trained physician (depending on their experience and depending on whether they are GPs or specialists), and according to the complexity of the case (depending either on the number of morbidities, or on the number of signs and symptoms).

7.2 The Experiment

To evaluate the system, an experiment was carried out with the medical residents in the Hospital Clínic de Barcelona. For the study we counted with twenty-three physicians. The profile of the residents was diverse in terms of speciality (specialists versus general practitioners), and the number of years of residency. Table 7.6 shows the characteristics of the residents in the experiment.

Specialty	**Residency year**	**Quantity**
generic	1	2
generic	2	4
generic	3	1
generic	4	2
generic	junior associate	1
specialized	1	9
specialized	2	3
specialized	3	1

TABLE 7.6: Characteristics of the residents

The 30 cases in the patient description case base were used as baseline for the physicians. These cases were shuffled and presented in a different order to the different residents to avoid case discussions between the users of the training system. For each resident, ten additional cases were taken at random from the patient description database with the purpose of evaluating the achievements of the learning system when the users were asked to diagnose repeated different cases.

The cases were exposed one by one to the residents by means of a web server. For three weeks the web page remained open so that the physicians could work with the cases. The system stored the information after each case was closed, so the residents could interrupt their training at any time and continue with the following training cases whenever they

wanted, later. If a case was left unsolved at the end of a work session, the DDx decisions made for that case were lost and the case was the first one to diagnose from scratch, in the following session. Physicians were able to look up the CPGs [191, 195–202] at any time (direct links to the CPGs were provided in the web page).

After three weeks, the web access to the system was blocked. The stored information was used to compare the training of the residents in terms of improvement and stabilization. For the analysis of *training improvement* we compared the mean user's adherence to CPGs in the first 5 cases of the training (cases 1 to 5), with the user's adherence to CPGs in the last 5 training cases (cases 26 to 30). For the analisis of training stabilization we compared the mean adherence to CPGs in the last 5 training cases of the training (cases 26 to 30), with the user's adherence to CPGs in the repeated additional 10 cases (cases 31 to 40).

Figure 7.5 details the sequence of cases in the experiment.

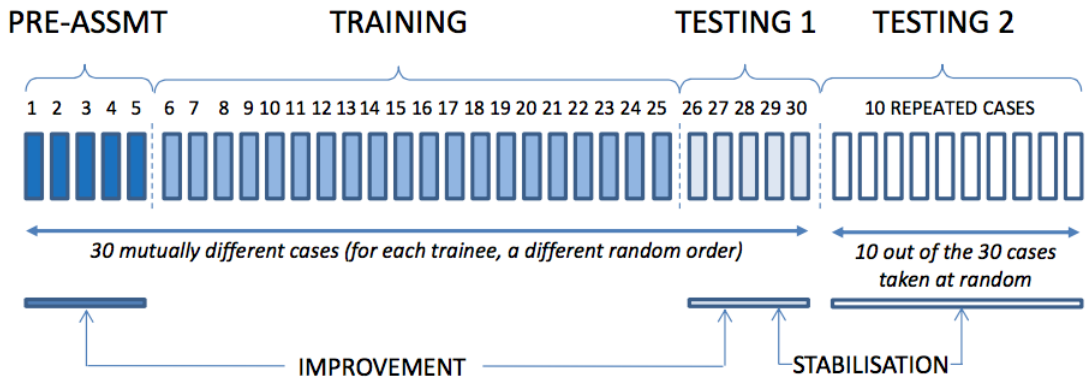


FIGURE 7.5: Sequencying of the DDx pre-assessment, training, and testing stages

7.3 Results

After the experiment, the abilities as described in section 7.1.4 were analyzed. In all the analyses we calculated the accuracy, sensitivity, specificity, and positive and negative predictive values of residents' decisions, and performed t-Student's t-tests to obtain the p-values in table 7.7

In the next subsections we present the results obtained in terms of the average improvement, after training, and the average stabilization of the adherence to the CPGs recommendations. Results are presented globally for all the residents and cases together, but also relative to the resident's experience and specialty, and the patient's complexity.

Variable	**Cases 1-5**	**Cases 26-30**	**Cases 31-40**	**Improv. P-Value**	**Stabil. P-Value**
Initial Hypothesis					
Accuracy	0.64	0.71	0.72	**0.026**	**0.98**
Sensitivity	0.43	0.57	0.56	**0.001**	**0.64**
Specificity	0.83	0.82	0.86	0.795	0.42
Positive Predictive Value	0.65	0.67	0.72	0.544	0.43
Negative Predictive Value	0.67	0.77	0.74	**0.009**	**0.28**
Final Diagnosis					
Accuracy	0.90	0.92	0.90	0.90	0.36
Sensitivity	0.58	0.60	0.62	0.87	0.85
Specificity	0.95	0.96	0.94	0.49	0.38
Positive Predictive Value	0.67	0.68	0.66	0.94	0.71
Negative Predictive Value	0.94	0.95	0.95	0.96	0.48
Mean Adherence (*thypothesis selection* - decision D1)					
Accuracy	0.90	0.93	0.90	**0.01**	0.008
Sensitivity	0.40	0.57	0.51	0.58	0.77
Specificity	0.91	0.93	0.91	0.02	0.04
Positive Predictive Value	0.40	0.56	0.48	0.65	0.68
Negative Predictive Value	0.98	0.99	0.98	**0.003**	0.004
Mean Adherence (*test selection* - decision D2)					
Accuracy	0.89	0.94	0.93	**0.03**	**0.14**
Sensitivity	0.69	0.79	0.79	**0.03**	**0.57**
Specificity	0.90	0.95	0.94	**0.04**	**0.22**
Positive Predictive Value	0.53	0.71	0.67	**0.01**	**0.22**
Negative Predictive Value	0.97	0.98	0.97	0.12	0.20

TABLE 7.7: Estimates of mean improvement after the use of the training system

7.3.1 Global Results

Evidence was found that the training system improved the sensitivity, negative predictive value and the accuracy of residents at the time of suspecting initial hypotheses (p-values 0.001, 0.009, and 0.026, respectively). Moreover, this improved ability could remain after the training, as column stabilization does not provide evidence of differences ($P>0.28$).

After training, there was not a significant change in the diagnostic capacities of the residents that remained high ($P\approx0.8$).

An improvement in the accuracy, specificity, and negative predictive value of residents' in hypothesis selection along the differential diagnosis process was also detected with $P=0.01$, $P=0.02$, and $P=0.003$, respectively. That is to say, with the use of the system, residents learned to disregard unfounded hypotheses. However this acquired ability did not last ($P<0.04$).

Finally, we observed a lasting improvement of residents accuracy, sensitivity, specificity, and positive predictive value at the time of selecting diagnostic tests along differential diagnosis (P=0.03, P=0.03, P=0.04, and P=0.01, respectively). In other words, residents improved and stabilized their ability to order the proper tests, and only the proper tests, in accordance to the indications in the CPGs.

7.3.2 Resident's Experience and Specialty

Residents' experience ranges between 1 and 4 years, after graduation. When we analyzed the data depending on this experience, the following results were obtained: residents in the first year were the most influenced physicians by the training tool. This tool improved 10% their accuracy at the time of suspecting the right diseases (P=0.045) and 21.4% their ability not to rule out diagnoses that are possible (P=0.007), during the first encounter with the case. Moreover, first-year residents benefited from the tool by improving 10% their adherence to the CPGs recommendations (P=0.046), in average. For the choice of tests, the benefits were 12% improvement of the mean accuracy (P=0.03), 20% risk reduction of forgetting relevant tests (P=0.008) and 12.3% risk reduction of asking irrelevant tests (P=0.025).

For second-, third- and forth-year residents (i.e., experienced residents), improvements were moderate and non conclusive. This could be attributable to the low number of trainees with these levels of experience (n=7, 2, and 2, respectively), but also to their reluctance to change their DDx schemes when they were already having 88% mean accuracy in their final diagnoses (in the pre-assessment stage). This percentage rose to 90% after the training.

We also observed that first-year residents failed more than experienced residents in the selection of the correct tests (P=0.042). Their failure ratios being 77.7% and 87.4%, respectively. This left novice residents more room to improve than senior residents, as the results after training showed, with novices reaching 89.6% of accuracy (P=0.03) and seniors showing a non significant improvement (P=0.723).

Interestingly, the accuracy of first-year residents to correctly diagnose patients was increased to the level of the accuracy of forth-year residents (91%), after the training. However, the DDx process followed by first-year residents to reach these correct diagnoses improved its adherence to CPGs from 68.5% before the training to 78% after the training (P=0.046), but remained unchanged and below 70% for forth-year residents (P=0.453). For first-year residents we could not find evidence that the improvements were lost after the training (0.32<P<98).

Resident's specialty was used to compare the influence of the training tool between GPs and specialists. For GPs the accuracy and the sensitivity of the initial suspected diseases improved 14% (P=0.005) and 13.7% (P=0.073), respectively. Conversely, the specialists improved their sensitivity to identify initial suspected diseases in 15% (P=0.02), but also their accuracy in 10% (P=0.019), sensitivity in 13.3% (P=0.033) and specificity in 10% (P=0.015) at the time of ordering diagnostic tests, as CPGs recommend. In all these results the benefits lasted after the training (P>0.7).

7.3.3 Patient's Complexity

In this study, the patient complexity is measured in terms of either the number of multimorbidities of the case, or the number of signs and symptoms describing the case. As the number of morbidities (or signs and symptoms) increase, the patient is considered to show a higher complexity which may affect the DDx process.

For patients with essential hypertension (n=5) (i.e., not a single secondary cause of hypertension is found) the improvement of resident's DDx decisions was only conclusive for the determination of the suspected diseases along the diagnostic process: accuracy improved 10% (P=0.018), and sensitivity 26.2% (P=0.035).

For the diagnosis of patients with one secondary cause of hypertension, some improvements were observed on the accuracy 11.6% (P=0.003), and the sensitivity 18.2% (P=0.005) of the initial suspected hypotheses. For the full DDx process, the most outstanding improvements were: 12% in accuracy (P=0.001), and 10% in specificity (P=0.001). For the adherence to DDx tests, improvements were observed with regard to average accuracy 13.2% (P<0.001), sensitivity 15.6% (P=0.005), and specificity 11.8% (P<0.001). Unfortunately for most of these improvements, we did not find significant evidence that they lasted after the training.

For patients with multimorbidity, residents tend to suspect of only one disease. This is observed when we compare the sensitivity of the suspected diseases along the DDx process for single-disease (35.3%) and for multimorbid (19%) patients. That is to say, the observed mean probability of failing to consider an existing disease is 24.3% lower for single-disease cases than for multimorbid cases (P=0.002).

As far as the number of signs and symptoms, four patient complexities were studied: those with 3 or less initial signs and symptoms, with 4 or 5, with 6 or 7, and those with 8 or more.

The cases in the range 4 to 5 was the group that our tool provided a better training for residents. For these patients, the accuracy of trainees to provide good diagnostic

hypotheses during the first encounter increased 11.7% (P=0.016), and for specificity 14.9% (P=0.01). Also, the average capacity of residents to discard irrelevant diagnoses along the DDx process increased 10% (P=0.018). As far as the capacity of residents to propose the diagnostic test pointed out in the guidelines, this was improved in terms of accuracy 13.8% (P=0.005), sensitivity 22% (P=0.005), and specificity 13.8% (P=0.005).

7.4 Conclusions

In this chapter we ran a test about the training of residents at the Hospital Clínic de Barcelona for the diagnosis of secondary causes of hypertension affecting single-disease and multimorbid patients. The DDx Training tool was described in terms of its four components: the patient creator module, the patient analysis module, the online questionnaire module, and the result analysis module. A DDx training experiment was designed with the help of physicians from that hospital. In the experiment, we studied the accuracy, the sensitivity and the specificity of the DDx processes and observed a significant improvement (between 10% and 20%) of the residents' mean sensitivity in all the sort of decisions of the DDx process: determination of diseases along the whole process (including the first encounter), and selection of the correct tests. These results are even better for first-year residents who enhanced more than 20% their capacity not to forget possible diseases during the first encounter, and not to rule out required tests along the DDx process.

The system allows users to train with multimorbid cases and promotes their change of mentality from traditional single-disease thinking to modern multimorbid consideration in a natural progressive way.

Apart of the good results of the tool in the training of residents to identify essential hypertension following the standard DDx procedures, the residents that initially suspected a correct essential hypertension were also instructed to consider possible secondary causes of hypertension.

Our study states that the lack of adherence to CPGs is better corrected in first-year residents than in residents in subsequent years, which drives us to consider that DDx training tools are more suitable to improve the expertise of novice physicians who are more given to consult and apply the DDx processes contained in the CPGs.

Chapter 8

Use of Treatment Decision Tables to Train Residents in Emergency Shock

In chapter 4, the structure of the treatment and prognosis decision tables were presented. In chapter 6, the structure of treatment tables were used to model the treatment of clinical practice guidelines about eight emergency shocks, and the structure of prognosis tables were used to simulate the consequences of the treatment actions during the treatment of these sorts of shock.

In a same way that in chapter 7, this chapter introduces a new experiment where the knowledge contained in the treatment decision tables is tested in a training program of the Hospital Clínic de Barcelona by means of an incremental knowledge-based on-line training tool. The objective of this tool is to improve the adherence of the hospital residents to CPGs in the treatment of shock at the Emergency Unit.

This chapter presents the training tool and how it was used to train medical residents. At the end of the chapter, the results of the experiment are evaluated.

8.1 The training tool

An online learning training tool was implemented to train medical residents in the adherence to treatment. Figure 8.1 details the scheme of the application.

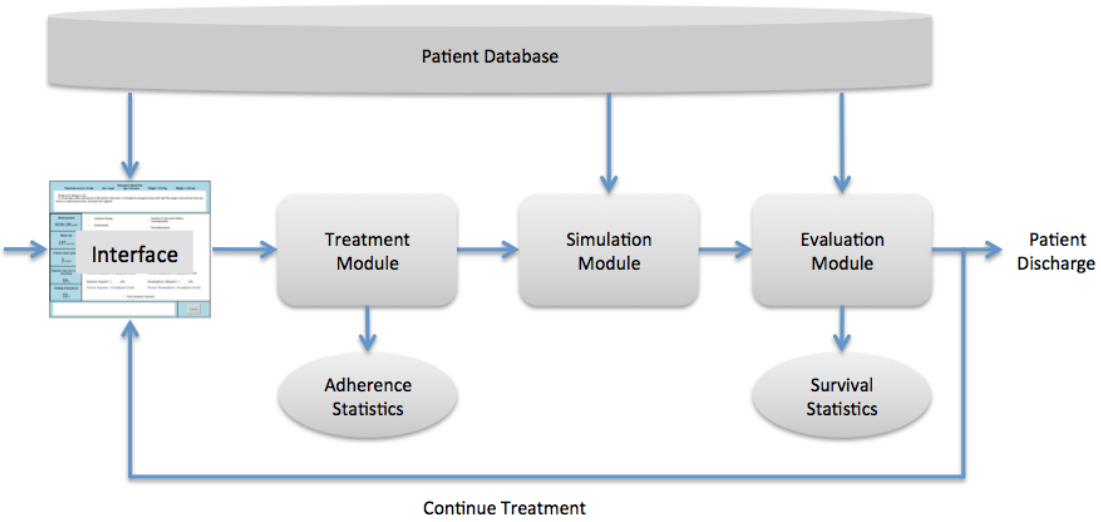


FIGURE 8.1: Emergency Shock Treatment Modules

At the beginning of each case, the interface module obtains the information of the patient from the patient database. This information is displayed and provided to the trainee from which some treatment recommendation is expected.

When the student provides a recommendation, the treatment module compares this recommendation with the actions suggested by treatment decision tables. Let us recall that the actions in the treatment decision tables represent the knowledge obtained from specialized clinical practice guidelines. This comparison generates some adherence statistics for a later analysis.

The next step of the learning process is to modify the health condition of the patient accordingly to the prescribed actions. The simulation module uses the prognosis decision tables and the patient information (i.e., current patient condition, and her resistances and sensitivities to clinical actions) to evolve the patient condition. In the experiment about the treatment of shock only the hemodynamic parameters are taken into account.

Finally, the evaluation module decides whether the patient is discharged or if the training process continues. The patient may be discharged if the trainee decides to discharge the patient, or if some dangerous action has been prescribed that provokes the patient's signs to be out of range or if the training time is over. This module also generates some statistics about the patient's survival. The patient survival is a measure that provides information about the status of the patient. This measure takes into account the vital signs and their values, penalizing the ones that are out of the normality range.

The next subsections detail each one of the components of the learning tool in the context of the training of hospital residents of the treatment of emergency shock.

8.1.1 The Patient Database

This experiment uses the patient database described in section 6.4. All the patients in the database are adults, because children require pediatric CPGs which are different to the ones used in our study.

The treatment of shock involves the combination of ventilation support, fluid resuscitation, and vasoactive agents. In our database, all patients have received ventilator support before the process started, consequently the experiment is focused on the administration of fluids and vasoactive agents.

Table 8.1 shows the distribution of cases in the patient database. Each patient is assigned a simulator depending on the type of shock. Usually there is one patient simulator per shock, but it is possible that one shock might be due to different causes that can drive to different responses in front of a same treatment. In such a case, several patient simulators are needed. This is the case of Septic Shock which requires two different simulators (high and low) as the table shows.

Shock	Simulator	Cases	Modif.	Resist.	Sensit.
Anaphylactoid shock	Anaphylactoid	7	4	4	1
APE RV Cardiogenic Shock	APE RV	7	0	0	0
AMI LV Cardiogenic Shock	AMI LV	8	8	8	0
Hemorrhagic shock	Hemorrhagic	7	6	6	2
Cardiac tamponade	Tamponade	7	2	2	0
Neurogenic shock	Neurogenic	7	5	5	2
Septic shock	Septic High	6	5	1	5
	Septic Low	2	2	1	2
		51	32	27	12

TABLE 8.1: Typologies of patients in the shock database

The column Cases shows the total number of cases in the database, for each simulator. The column Modif in the table shows the number of patient that have some resistance or sensitivity to one or more clinical actions. Patients can have one or both resistance or sensitivity to several clinical actions. For example, a patient can be resistant to dobutamine, but sensitive to dopamine. The last two columns show the number of patients with some resistance and the number of patients with some sensitivity to actions, respectively.

8.1.2 The Interface

The Interface Module (IM) is the training tool component that interacts with the trainee. It is based on the form presented in figure 8.2.

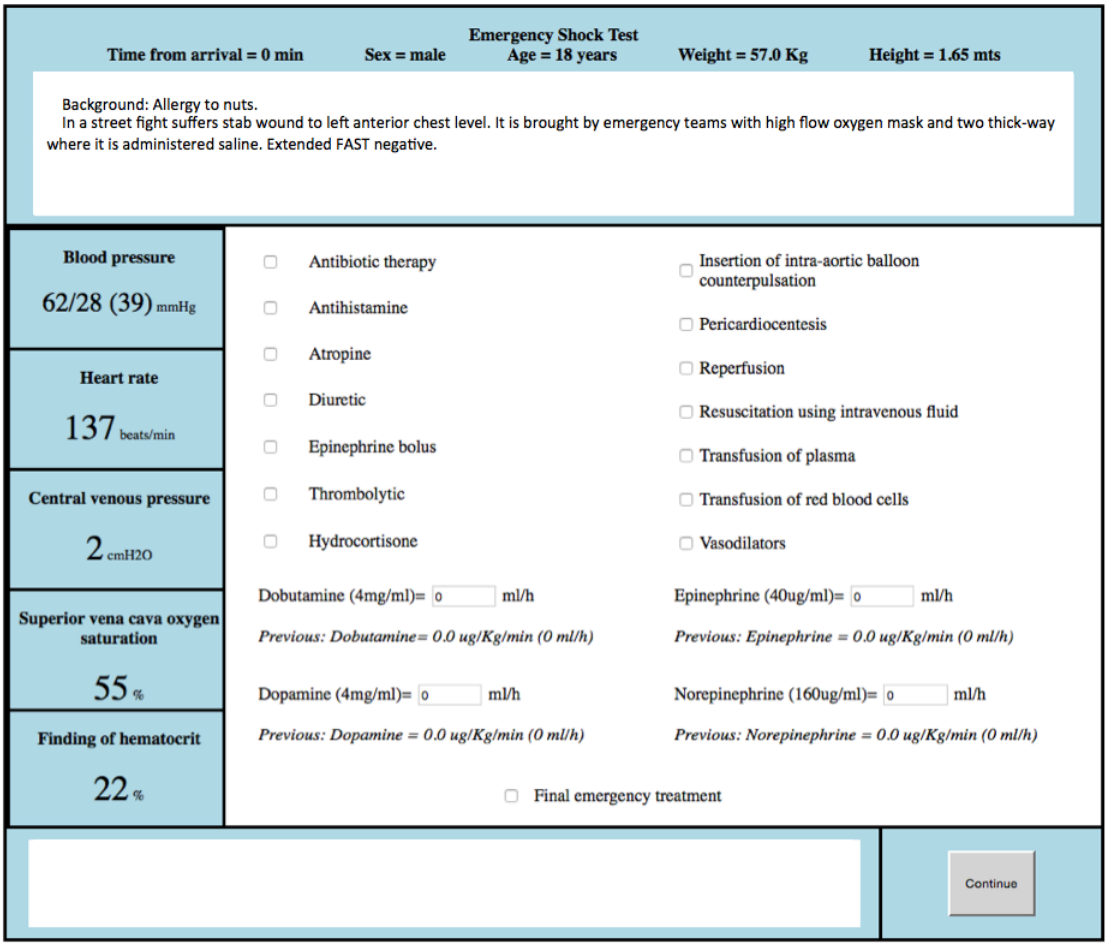


FIGURE 8.2: Emergency Shock Interface

In the top, the patient information is shown. This includes the patient's clinical description, the history, and the time from arrival to the ICU. This information is intended to provide evidence to the trainee about the type of shock, the possible actions to do, and the modulation of the treatment. For example, the trainee has to be careful with actions that may produce tachycardia in older patients, or to prescribe larger dosages in patients with obesity.

The information in the history should provide enough evidence for a trained practitioner to diagnose the type shock. However, the objective of this tool is to train the treatment, not the diagnosis. For example, the case displayed in figure 8.2 describes a 18 year young man, allergic to nuts, who suffers from a stab wound and blood loss. The background may suggest a anaphylactic shock caused by the allergy, but the wound and the lost of blood determined by the low value in finding of hematocrit (22%), provide enough evidence to diagnose a hemorrhagic shock.

The middle left column shows the patient signs related to the hemodynamic parameters (i.e., blood pressure, heart rate, central venous pressure, superior vena cava oxygen

saturation, and finding of hematrocrits). These signs are updated accordingly every time that a treatment action is carried out.

The middle right area of the screen presents a form to be filled by the trainee with the decided treatment. This form includes all the discrete treatment actions (see the fourteen upper actions), the continuous actions or infusions (see the four numerical fields), and the discharge action (see the action at the bottom). Discrete and discharge actions only need to be selected, but the infusions require a selection of the corresponding dosages.

Finally, the button at the right bottom corner is to conclude the treatment step once the trainee has decided the actions to perform. The rectangle at the left bottom area is used to provide feedback information to the trainee.

8.1.3 The Treatment Module

The Treatment Module (TM) is responsible of evaluating the adherence of the actions picked by the trainee to the actions suggested by the CPGs. This is calculated step by step after each iteration.

To make the comparison between the answers of the user and the guideline indications, the system introduces the patient data in the TDTs. As a result, the system obtains different treatment possibilities. Then, the system compares the user answer with each one of the treatment suggestions, obtaining the number of true positives (i.e., actions suggested by the CPGs that the user has indicated), false negatives (i.e., actions suggested by the CPGs that the user has not indicated), and false positives (i.e. actions indicated by the user that the CPGs do not suggest).

Among the treatment possibilities, the one that maximizes the accuracy is considered the most nearby, and the one used for comparison.

There are two levels of comparison, the pharmacological level and the dosage level. The pharmacological level assesses whether the treatment actions chosen by the students are the same actions that the guidelines suggest. The dosage level assesses not only the action, but also the correct dosage. For example, if the decision table recommends epinephrine with low dosage, and the trainee select epinephrine with high dosage, the answer is correct at pharmacological level, but is wrong at dosage level.

The adherence is measured with the metrics accuracy, distance, sensitivity, specificity, positive and negative predictive values, and t-measure, which are based on the number of true positives, false negatives and false positives. These measures explain to what extent the answers of the user match the suggestions of the CPGs.

8.1.4 The Simulation Module

The Simulation Module (SM) uses the prognosis decision tables as a patient simulator to calculate the evolution of the cases under study. This module considers the iterations presented in section 6.5 (see figure 6.5) as it follows.

In each iteration, for the initial state St_i, the SM uses the prognosis decision tables (PDT) to calculate the patient state P_{i+1} and the internal state St_{i+1}. The patient state P_{i+1} takes into consideration both discrete and continuous actions of the treatment whilst the internal state St_{i+1} only the discrete actions. The users receive only the information of P_{i+1}.

The changes produced for discrete actions are obtained from the PDT and modified accordingly the resistances and sensitivities of the patient (i.e., the patient model or PM). The trainees do not know the patient's resistances and sensitivities, but they can suspect them if the patient's response is not the expected one.

Table 8.2 shows the range boundaries for infusions. The concentrations in the table are expressed in μg/Kg/min as they appear in the CPGs, but residents use to work in ml/h units.

	Low	Med	High	Max	Conversion factor
Dopamine	5.0	12.0	30.0	70.0	66.7 / weight
Dobutamine	7.0	15.0	30.0	60.0	66.7 / weight
Epinephrine	0.08	0.2	0.4	1.0	1.3 / weight
Norepinephrine	0.12	0.6	1.8	3.0	0.7 / weight

TABLE 8.2: Infusion ranges and conversions in μg/Kg/min

The IM allows the students to work in ml/h, but the dosage prescribed to the patient has to be converted before the knowledge in the decision tables are applied. The last column in the table shows the conversion factor used based in the drug infusion concentrations.

The effect of the drug depends on the dosage. The result calculated by the prognosis decision table is adjusted according to a modification factor represented by functions as the one in figure 8.3, representing the adjustment factors of dobutamine.

The procedure can be understood with an example. A resident recommends 18 ml/h of dobutamine to a patient of 80 kg, and the prognosis decision table shows that it has a modification of +10 in the heart rate. Under such circumstances, the simulator converts the value from ml/h to μg/Kg/min with the conversion factor in table 8.2 (i.e., 66.7/weight). It obtains a dosage of 15 μg/Kg/min (i.e., $18ml \times 66.7/80Kg$). According to the function in figure 8.3, for a dosage of 15 μg/Kg/min a conversion factor of 1.75 is obtained, and the heart rate is modified with +17.5 (i.e., (+10) $\times$ 1.75).

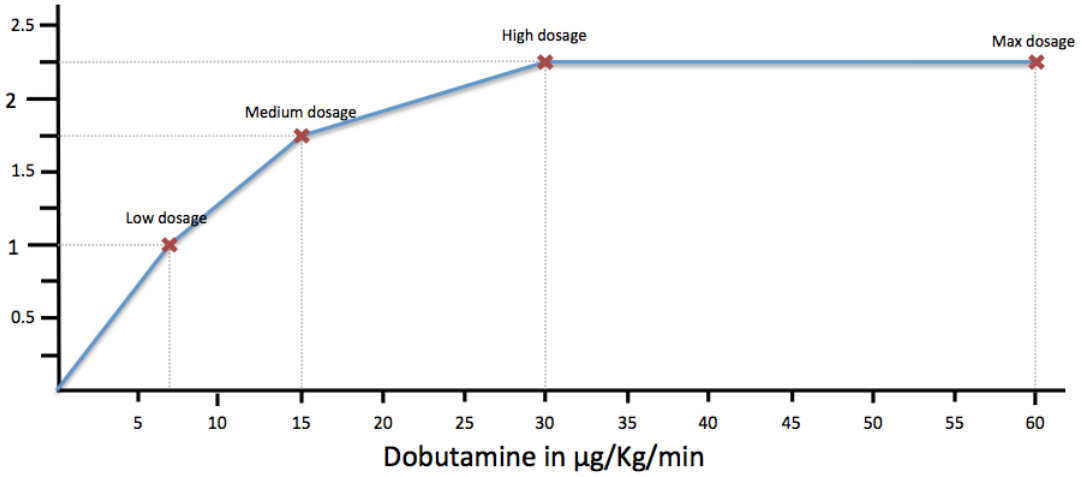


FIGURE 8.3: Example of dobutamine effects

The infusions can also be affected by resistances or sensitivities of the patients. For example, if the patient above was 110% sensitive to the effects of dobutamine, the heart rate would have a final increment of +19.25 (i.e., (+17.5) × 110%).

In order to simulate other uncontrolled factors that may affect the effects of the treatment actions, and also to provide more realism along the training, when the new values of all the parameters are calculated as previously explained, the system introduces a variation of ±10% in the parameter modifications. For example, if decision tables conclude that the heart rate suffers a modification of +20, the simulator provides a value between +18 and +22.

8.1.5 The Evaluation Module

The Evaluation Module (EM) decides whether the case remains active, or if the patient is discharged, if the student has exposed the patient to a dangerous action, if the time for the treatment is over.

The patient is discharged if the resident uses the discharge action *final emergency treatment*. In this case, the system evaluates the vital signs of the patient and provide a survival score based on the patient's shock stability.

When the patient is discharged, the final survival score is calculated. This score compares each vital sign with the normality range. If all the sign values are in the normality range, the score is 100%. If one or more signs are out of the normality range, the relative distance between the normality range is computed for all the abnormal signs, and take the minimum score of the signs.

The case is forced to finish if the patient is exposed to dangerous situations. Dangerous situations happen when risk measures are applied to wrong shocks, infusion doses are prescribed over the maximum limits, or a vital sign goes out of the security limits. In these cases the survival value is set to 0%.

The case is also finished if the security time to treat a shock is over. In this case, the survival value are calculated, but the case is closed.

Although the objective of the training tool is to improve the adherence to the treatment proposed by the CPGs, the survival values provide additional information about the quality of the treatment.

8.2 The Long-Term Treatment Experiment

This new training tool has been tested with the support of residents from the Hospital Clínic de Barcelona. These residents were involved in a training program to analyze the consequences of introducing artificial intelligence tools in the improvement of the adherence of medical students (trainees) to the clinical practice guidelines for the treatment of seven different sorts of shocks in an emergency unit.

The experiment divided the trainees into two groups, the intervention group and the control group. Figure 8.4 shows a summary of the sequence of the cases in the study as they were exposed for treatment to the intervention and control groups.

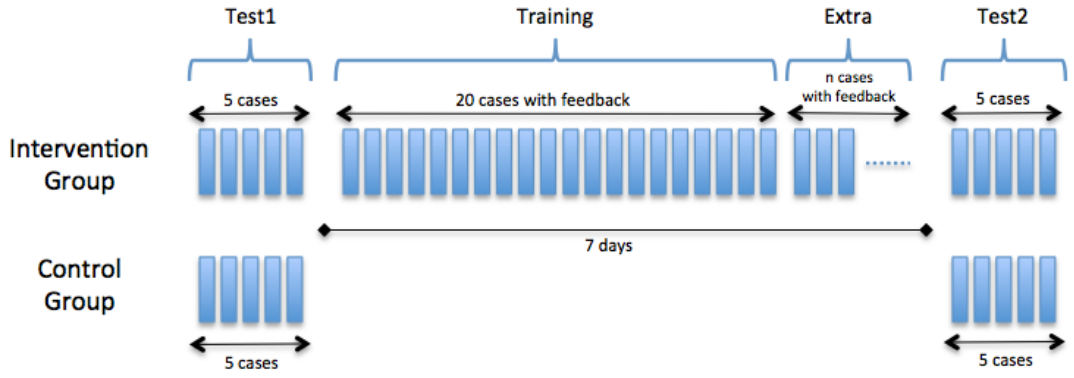


FIGURE 8.4: Sequencying of cases in intervention and control groups

During test 1, both groups are assessed about their initial ability to treat shocks. Five different cases of different shocks took at random were exposed to the participants. Then, during seven days, the members of the intervention group were allowed to train with 20 cases, providing them with an online feedback of their correct and incorrect decisions (*Training*). These cases were taken at random from the case base, promoting

cases with different sorts of shock. Trainees in the intervention group were given the possibility to solve additional cases, if they were solved during the seven training days (*Extra*). After the training, a new test of five cases taken at random without feedback was proposed to assess the new skills of the students (*Test2*). If the training with the 20 cases did not finish in seven days, the system delayed the second test until all the training cases were finished.

After the first test, the members of the control group had to wait for seven days before they started the second test (*Test2*).

Table 8.3 shows the distribution of the participants in the experiment in the intervention and control groups and their category.

Group	**Category**	**Participants**
Intervention	R1	11
Intervention	R2	3
Intervention	R3	2
Intervention	R4	1
Intervention	Jr. Associate	1
Control	R1	8
Control	R2	3
Control	Jr. Associate	4

TABLE 8.3: Participants in shock experiment

8.3 Results

The results obtained in the experiment have been analyzed to measure the adherence to the clinical practice guidelines and the patient survival.

To measure the adherence we calculated the accuracy, distance, sensitivity, specificity, positive and negative predictive values, and t-measure for the intervention and the control groups. These values were obtained for each test (i.e., tests 1 and 2), and level (i.e., pharmacological and dosage levels). For survival values only the arithmetic means were considered.

Both, adherence and survival statistics, are analyzed with the application of t-Student's tests to obtain the corresponding p-values.

In the next subsections we present the results obtained These are separated in global results and results conditioned to the resident's experience and their specialty.

8.3.1 Global Results

Table 8.4 shows the mean values obtained for each quality parameter analyzed in the shock treatment experiment.

	Intervention				**Control**			
	Test1		**Test2**		**Test1**		**Test2**	
	Pha.	**Dos.**	**Pha.**	**Dos.**	**Pha.**	**Dos.**	**Pha.**	**Dos.**
Accuracy	0.915	0.912	0.934	0.930	0.906	0.901	0.920	0.920
Distance	0.477	0.457	0.546	0.520	0.463	0.429	0.463	0.443
Sensitivity	0.655	0.626	0.699	0.661	0.696	0.652	0.604	0.584
Specificity	0.937	0.936	0.952	0.950	0.923	0.921	0.945	0.943
Pos. Pred. Val.	0.542	0.515	0.595	0.566	0.515	0.482	0.525	0.502
Neg. Pred. Val.	0.969	0.968	0.976	0.974	0.974	0.971	0.968	0.967
F-Measure	0.563	0.537	0.620	0.590	0.554	0.519	0.530	0.510
Survival	0.309		0.450		0.358		0.464	

TABLE 8.4: Means for shock experiment

The results show that there are no evidences of a representative difference between the pharmacological and dosage levels. No p-values under 0.05 have been found between the different groups and tests, for any quality measure, to support significant differences. From this analysis, we could not distinguish between pharmacological and dosage levels, so, for the sake of brevity, the next conclusions are presented only at pharmacological level.

Table 8.5 shows the p-values found for test 1, test 2 (comparing the intervention and the control groups results), and the intervention and control groups (comparing the results between test 1 before training and test 2 after training).

	Test1	Test2	Intervention	Control
	Inter. vs Control	Inter. vs Control	Test1 vs Test2	Test1 vs Test2
Accuracy	0.359	0.078	**0.011**	0.173
Distance	0.718	**0.036**	0.064	0.990
Sensitivity	0.393	0.052	0.334	0.074
Specificity	0.135	0.315	**0.023**	**0.034**
Pos. Pred. Val.	0.532	0.100	0.184	0.826
Neg. Pred. Val.	0.294	0.055	0.083	0.200
F-Measure	0.834	**0.029**	0.142	0.559
Survival	0.320	0.784	**0.004**	**0.041**

TABLE 8.5: P-Values between groups in the shock treatment study

Before the training, the treatment skills of the members in the intervention and control groups were tested with the test 1. The results do not provide evidence about differences between the levels of the groups (all p-values are above 0.05). Consequently, there are

not reasons to suspect that the quality of the members in the intervention group is higher or lower than the one of the members in the control group.

However, the comparison of the means of the quality parameters between the two groups before and after the training (colum Test2) shows that the adherence to the guidelines of the intervention group is always better that the results of the control group. After the experiment, the intervention group presented less wrong answers (distance p-value 0.036, with respective means 0.546 and 0.463) and a general better performance (F-measure p-value 0.029, with means 0.620 and 0.530).

After the training, the physicians in the intervention group improved all the quality measures (see column intervention in table 8.4. That is to say, all the arithmetic means were found better in test 2 than in test 1. Moreover, in column intervention of table 8.5, several significant improvements were found in their accuracy (p-value 0.11, with means 0.915 for test 1 and 0.934 for test 2), their specificity (p-value 0.023, with means 0.937 for test 1 and 0.952 for test 2), and the patient stabilization (p-value 0.004, with means 0.309 vs. 0.450).

Comparatively, after the delay time, the members of the control group (without a real training) had a better score avoiding false positives. This is represented in column control of table 8.4 by an improvement of the specificity and the positive predictive values, but results provide that the number of false negatives also increase. This comes represented by a worse sensitivity and negative predictive values. However, they present a great improvement in the survival score (p-value 0.041, with means 0.358 and 0.464) in column control of table 8.5.

In general, we observe that both groups present improvements between test 1 and test 2. The intervention group betters both, their adherence to guides and the patient stabilization, but the control group only the patient stabilization.

8.3.2 Resident's Experience

In the groups of intervention and control, the first year residents (R1) before the training did not present evidences to consider relevant different skills treating shock (all p-values are greater than 0.05).

After the training, all the statistical averages of the intervention group was better that the same averages of the control group. The improvement was relevant in distance (p-value 0.016, averages 0.542 vs. 0.419), sensibility (p-value 0.032, averages 0.704 vs 0.561), positive predictive value (p-value 0.021, averages 0.596 vs 0.468), and F-measure (p-value 0.009, averages 0.622 vs 0.481).

After the training, the intervention group improved all the quality measures in average. The most relevant improvements were in the accuracy (p-value 0.010, averages 0.91 vs 0.934) and the specificity (p-value 0.011, averages 0.93 vs 0.952). The control group, after the delay of 7 days, improved the averages of accuracy, specificity, negative predictive value, and survival, but get worse averages for distance, sensibility, positive predictive value, and F-measure. There is not a relevant change in the adherence to CPGs, but the changes in survival are significant (p-value 0.042, averages 0.311 vs 0.460).

The residents of second year (R2) do not exhibit neither before nor after a training difference between groups. There are no evidences that the training improves their abilities. The only significant difference appears in the intervention group, after the training, with an improvement of the survival score (p-value 0.001, averages 0.318 vs 0.662). In fact, the results of the R2 intervention group after the training for the survival score beat the other residents (p-value$<$0.001, averages 0.662 vs 0.421).

This shows that novel physicians obtain more benefits from the tool in order of acquire treatment skills from guidelines. Experienced physicians seem to use the tool to increase their abilities for patient stabilization.

8.3.3 Resident's Specialty

Depending on the specialty, we split physicians into two groups, ERM and NO-ERM. ERM gathers the users with a specialty related to ICU. Their specialties are close to the diseases and treatments used in the study. In fact, before training, differences are detected in the adherence of the trainees of both groups in terms of sensibility (p-value 0.047, averages 0.671 vs 0.510). Contrarily, NO-EMR comprises all the residents participant in the study who are specialists in medical domains that are not close to ICU.

After the training, the ERM users in the intervention group improved their mean survival score (p-value 0.013, averages 0.296 vs 0.441). However, the NO-ERM users in the intervention group improved their adherence to clinical practice guidelines, specially in terms of distance (p-value 0.040, averages 0.421 vs 0.566) and negative predictive values (p-value 0.038, averages 0.963 vs 0.977). Interestingly, after the training, the differences in adherence to CPGs are not detected between these groups (no p-values $<$0.05).

Once again, the group with less experience takes a greater benefit from the learning tool in order to learn the clinical procedures while the experienced users train their competency to stabilize patients.

8.4 Conclusions

This chapter presents a learning tool that combines treatment decision tables with prognosis decision tables. The knowledge in the treatment decision tables is used to teach the treatment recommendations in clinical practice guidelines while the prognosis decision tables are used as patient simulators to allow long term treatment training.

Prognosis decision tables (PDT) are combined with a patient model (PM) to personalize the behavior of each case. The effect of the actions may be obtained directly from the combination of PDT and PM or, in the case of infusions, obtained by consideration of more factors, such as the dosage.

This tool has been adapted and used in a training program of residents in the Hospital Clínic de Barcelona to teach the clinical procedures contained in clinical practice guidelines for the treatment of seven types of emergency shocks. After the training, the inexperienced physicians (first-year residents or R1s and residents with a specialty distant to ICU or NO-ERMs) improved their adherence to clinical practice guidelines. The experienced physicians benefited from the simulator in order to better stabilize the patients.

UNIVERSITAT ROVIRA I VIRGILI
USE OF DECISION TABLES TO MODEL ASSISTANCE KNOWLEDGE TO TRAIN MEDICAL RESIDENTS
Francis Real Vázquez

Chapter 9

Conclusions

This thesis is about the application of decision tables to support medical practice. Medical practice is represented as a continuous process that includes the cognitive tasks of diagnostic, treatment, and prognosis. The ultimate aim is to design decision tables for these tasks and to find a methodology to represent the knowledge of clinical practice guidelines in these tables that are easy to manage by practitioners, and that can be used as base in computer tools as decision support systems or training tools.

We began studying the clinical process and we focused on the three cognitive clinical tasks previously mentioned. For diagnosis we identified three main subtasks related to differential diagnosis (DDx): making diagnostic hypothesis, selecting appropriate diagnostic tests, and discarding negligible hypotheses. We put special emphasis on the representation of multimorbidity. Multimorbidity is the simultaneous coexistence of several diseases in the same individual none of them considered an index disease [220]. The increasing prevalence of chronic diseases in the last decades and the expected evolution of chronic pathologies entail a growth of multimorbid cases nowadays and in the years to come [221–223]. As a result to these expectations, the US Department of Health and Human Services (HHS) activated a plan to improve the response to multimorbidity [224]. One of the goals of this plan was to encourage the search for scientific evidence for multimorbid conditions to ameliorate the utility of clinical practice guidelines [131, 225, 226].

The second step in this work was to formalize decision tables structures for each one of the clinical tasks. To do this, we proposed a new formalism of decision tables, the grouping decision tables model (GDT), as an upgrade of semantic decision tables. We used the GDT to represent the three subtasks of differential diagnosis. We designed the diagnostic hypothesis decision tables (DHDT) to make diagnostic hypothesis, the test selection decision tables (TSDT) to select the appropriate diagnostic tests, and the evaluation decision tables (EDT) to discard negligible hypotheses. In order to model

DDx with decision tables, a differential diagnostic system was presented that integrates all the diagnostic decision tables: DHDT, TSDT, and EDT.

For clinical treatment, we designed the treatment decision tables (TDT) that summarize the therapeutic knowledge in clinical practice guidelines. These tables needed variable types to solve the problems of multiple choice and ambiguity in the treatment.

To solve the prognosis of short-term evolution, we proposed the prognosis decision tables (PDT). These tables represent the evolution of standard patients. In order to approach real patient prognosis and to personalize the evolution of concrete patients, a patient model (PM) has been introduced. This was useful to use the prognosis tables as simulators.

In addition to the new structures of decision tables designed, we proposed a methodology to combine the treatment and prognosis decision tables to plan long-term treatments. Moreover, we employed the patient model combined with the prognosis decision tables as the basis to create a patient simulator.

After designing decision tables for each clinical task, the third step was the use of the these table structures to host the knowledge of CPGs.

For the diagnosis tables, we modeled the diagnosis knowledge of the eight more frequent causes of secondary hypertension. To obtain the tables we proposed a methodology based in an intermediate representation as clinical algorithm. DHDT were obtained after a process that started with the identification of the signs and symptoms obtained from the clinical practice guidelines involved and their later classification by experts in mandatory, alternative, irrelevant, and impossible signs. The TSDT and EDT were built automatically from clinical algorithms.

The CPGs of seven emergency shocks were taken to test the treatment and prognosis decision tables.

Treatment decision tables were made after a semi-automatic process in which medical experts analyzed the therapeutic knowledge in CPGs and provided short clinical algorithms for the different treatment options of each shock. The different possibilities of treatment and the possibility of a treatment to change along the management of patients in shock, implied additional difficulties that produced bigger tables and the lose of simplicity and compehension. This was partially solved with the incorporation of new variable types.

The prognosis decision tables of emergency shock were acquired from the experience of medical experts by surveys. These decision tables had big dimensions, but the information was stored structured and easy to understand and validate.

Prognostic tables were used for patient simulation. A database with 51 sorts of patient were created. All these cases were represented with the PM structure. The patients in the database were synthetic, but based on real cases observed at the Emergency Department of the Hospital Clínic de Barcelona.

In addition, we worked with the simulation of long-term treatment the concept of action effect duration. For emergency shocks we distinguished between discrete actions, with long duration, and continuous actions, with short duration. We presented the algorithm used to solve the different time effects.

Finally, to test the decision tables for the diagnosis of secondary causes of hypertension and treatment of emergency shocks, two studies were carried out with residents in Hospital Clínic de Barcelona. In both experiments we obtained similar conclusions:

- Decision tables are suitable to represent medical knowledge and to build training tools.
- The training tools improve the abilities of physicians, ease and accelerate the acquisition and update of procedural knowledge of the health care professionals.
- The efficiency of these tools is higher with inexperienced and untrained physicians. These tools make residents to improve their abilities in a faster and more efficacious way than traditional learning.

The main contributions of this thesis, in relation to the objectives enumerated in chapter 1, are:

1. **Design decision tables to expose the medical process:** In this work we propose a representation of the medical process based on differential diagnosis, treatment and personalized prognosis. All these stages are represented with decision tables.
2. **Use decision tables to represent the medical knowledge:** The decision tables designed have been used successfully to represent the knowledge of eight CPGs of secondary causes of hypertension for diagnosis, the knowledge of seven CPGs of emergency shocks for treatment, and the medical experience of evolution of shock patients for prognosis.
3. **Build training tools with decision tables:** Two learning tools have been built using the decision tables as a base of knowledge. These tools use simulated situations to test the knowledge of the users and provide feedback to correct the mistakes.

4. **Evaluate the performance of the tools with medical students:** The training tools have been tested with residents in Hospital Clínic de Barcelona. The results of the tests support the capability of the decision-table based tools to improve the abilities of the users, especially the inexperienced users.

This thesis opens new working lines as future work:

1. **Reduction of decision tables rules:** One of the problems of decision tables is the rise of dimension as new variables are added. This problem is not exclusive of this work. In this thesis, our approach to solve this problem was based on providing more expressiveness to the variables. Complex treatments with a lot of possibilities of therapeutic actions to perform may be translated in huge tables difficult to treat.

2. **Comorbidity in the treatment:** In this thesis we present how to manage the comorbidity and multimorbidity in the diagnosis. For clinical treatment, the control of comorbidity is highly relevant but still difficult to obtain from clinical practice guidelines because these use to target only one disease.

3. **Patient simulation:** One of the main contributions of this work is the creation of a patient simulator. The simulators allow residents to train their clinical skills and gain experience. The approach in this work is to create general simulators (PDT) from clinical surveys and personalize different patient behaviors with a set of simulated patients. As a future work, we should be able to automate the process of constructing PDTs with real patient data, and obtain different simulators for each typology of patient. The simulators obtained in this way would not only be used by training systems, but also by decision support systems.

4. **Extend the prognosis concept:** In this work, we attacked the problem of prognosis only as a short-term evolution of the patients, and we provided tools to estimate the long-term evolution, but this cognitive task comprise more subtasks, for example the prognosis of times in the evolution of the diseases for different patients and under different treatments. Our experience along this thesis suggests that prognosis is a hard cognitive task with limited medical technologies contributing. Most of them are based on the analysis of retrospective clinical data, but this approach has some limitations as for example the representativeness, completeness and soundness of the data analyzed. Alternatively, this knowledge can also be obtained from the personal experience of clinical experts by means of surveys. This is the approach that we followed in this thesis, but it still has similar drawbacks mainly related to the quality and quantity of experts involved in the elicitation of

prognostic knowledge. If decision tables have been shown in this thesis as valid structures to host prognostic knowledge, the process to obtain this knowledge deserves further attention.

UNIVERSITAT ROVIRA I VIRGILI
USE OF DECISION TABLES TO MODEL ASSISTANCE KNOWLEDGE TO TRAIN MEDICAL RESIDENTS
Francis Real Vázquez

Appendix A

Hypertension Secondary Causes Algorithms and Tables

A.1 Acromegaly

Acromegaly is a hormonal disorder that results from too much growth hormone (GH) in the body. The pituitary, a small gland in the brain, makes GH. In acromegaly, the pituitary produces excessive amounts of GH. Usually the excess GH comes from benign, or noncancerous, tumors on the pituitary.

Acromegaly is most often diagnosed in middle-aged adults, although symptoms can appear at any age. If not treated, acromegaly can result in serious illness and premature death. Acromegaly is treatable in most patients, but because of its slow and often "sneaky" onset, it often is not diagnosed early or correctly. The most serious health consequences of acromegaly are type 2 diabetes, high blood pressure, increased risk of cardiovascular disease, and arthritis. Patients with acromegaly are also at increased risk for colon polyps, which may develop into colon cancer if not removed.

Figure A.1 shows the clinical algorithm obtained from the CPG [195].

Figure A.2 shows the test selection decision table for Acromegaly.

Figure A.3 shows the evaluation decision table for Acromegaly.

A.2 Adrenal Cushing's Syndrome

Adrenal Cushing's syndrome is a metabolic disorder caused by excessive production of cortisol by a tumor of an adrenal gland. Cortisol is a hormone made by the adrenal

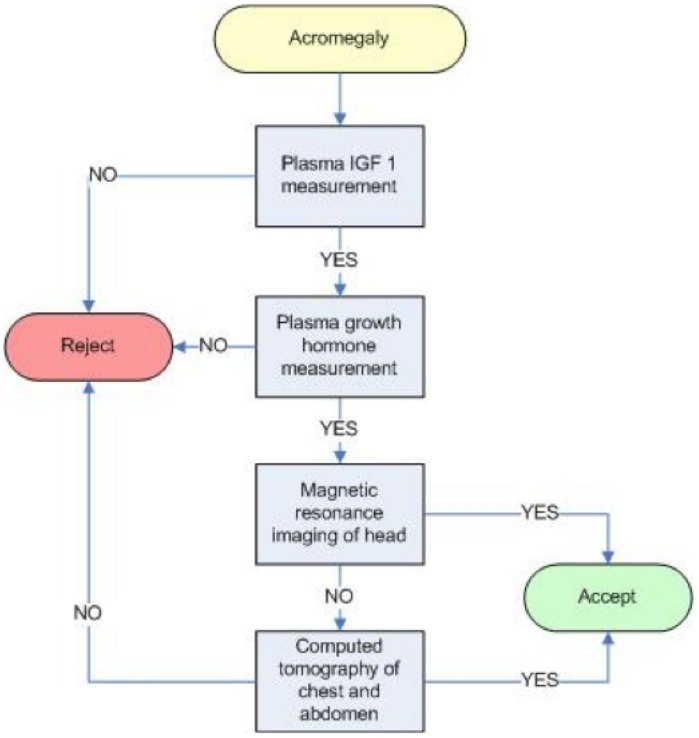


FIGURE A.1: Acromegaly Algorithm

Acromegaly	Yes	Yes	Yes	Yes
Computed tomography of chest and abdomen				?
Magnetic resonance imaging of head			?	No
Plasma IGF 1 measurement	?	Yes	Yes	Yes
Plasma growth hormone measurement		?	Yes	Yes
Computed tomography of chest and abdomen				X
Magnetic resonance imaging of head			X	
Plasma IGF 1 measurement	X			
Plasma growth hormone measurement		X		

FIGURE A.2: Acromegaly Test Selection Decision Table

Acromegaly	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Computed tomography of chest and abdomen						?	Yes	No
Magnetic resonance imaging of head				?	Yes	No	No	No
Plasma IGF 1 measurement	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Plasma growth hormone measurement		?	No	Yes	Yes	Yes	Yes	Yes
Acromegaly		?		?	X	?	X	

FIGURE A.3: Acromegaly Evaluation Decision Table

glands that plays an essential role in the stress response. Normal cortisol levels are necessary to sustain life, to maintain normal sleep-wake cycles, and to enable the body to respond to stressful events. Though limited bursts of cortisol are normal, long-term elevations of the cortisol level are harmful to many organ systems.

Figure A.4 shows the clinical algorithm obtained from the CPG [196].

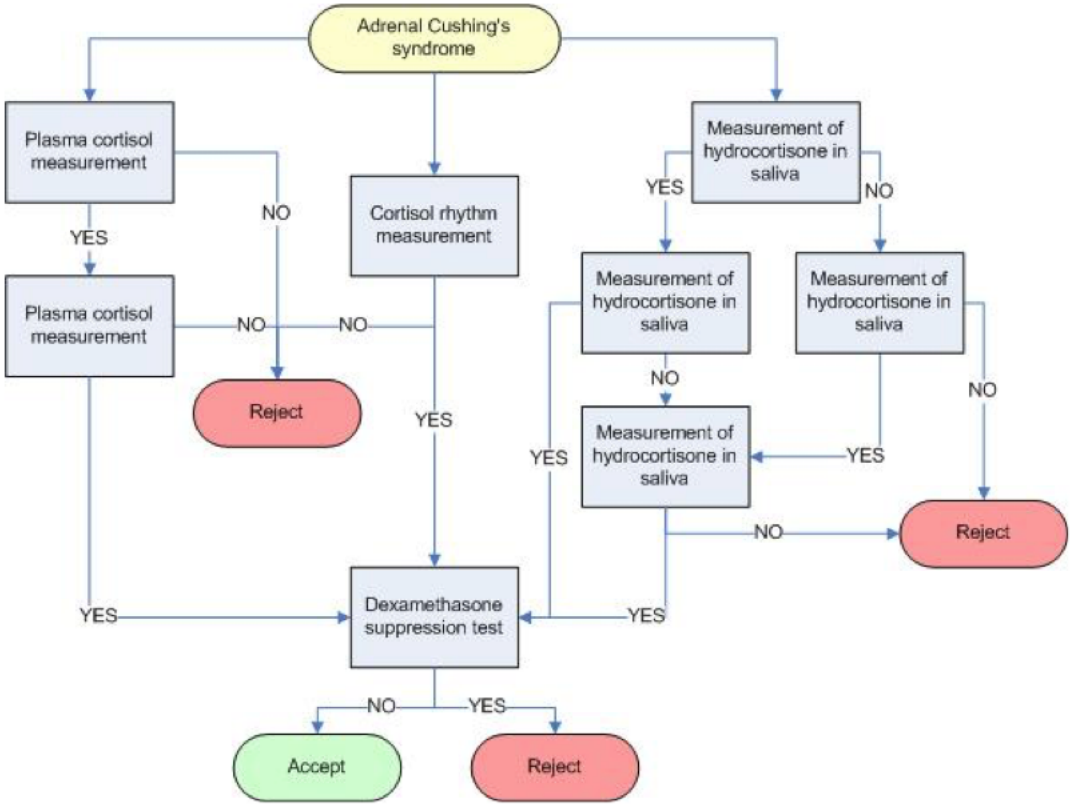


FIGURE A.4: Adrenal Cushing's syndrome Algorithm

Adrenal Cushing's syndrome	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cortisol rhythm measurement	?	?	?						Yes
Dexamethasone suppression test					?			?	?
Measurement of hydrocortisone in saliva{1}	?	?	?			No	Yes		
Measurement of hydrocortisone in saliva{2}						?	?	Yes	
Plasma cortisol measurement{1}	?	?	?	Yes	Yes				
Plasma cortisol measurement{2}				?	Yes				
Cortisol rhythm measurement			X						
Dexamethasone suppression test					X			X	X
Measurement of hydrocortisone in saliva		X				X	X		
Plasma cortisol measurement	X			X					

FIGURE A.5: Adrenal Cushing's syndrome Test Selection Decision Table

A.3 Coarctation of Aorta

Coarctation of the aorta (or aortic coarctation) is a narrowing of the aorta, the large blood vessel that branches off your heart and delivers oxygen-rich blood to your body. When this occurs, your heart must pump harder to force blood through the narrow part of your aorta. It is a type of birth defect. Aortic coarctation is more common in persons with certain genetic disorders, such as Turner syndrome.

Figure A.7 shows the clinical algorithm obtained from the CPG [197].

Adrenal Cushing's syndrome	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cortisol rhythm measurement											No	Yes		
Dexamethasone suppression test				?				?		?		?	No	Yes
Measurement of hydrocortisone in saliva{1}					No	Yes	No	No	Yes	Yes				
Measurement of hydrocortisone in saliva{2}					?	?	No	Yes	No	Yes				
Plasma cortisol measurement{1}	No	Yes	Yes	Yes										
Plasma cortisol measurement{2}		?	No	Yes										
Adrenal Cushing's syndrome		?		?	?	?		?		?		?		X

FIGURE A.6: Adrenal Cushing's syndrome Evaluation Decision Table

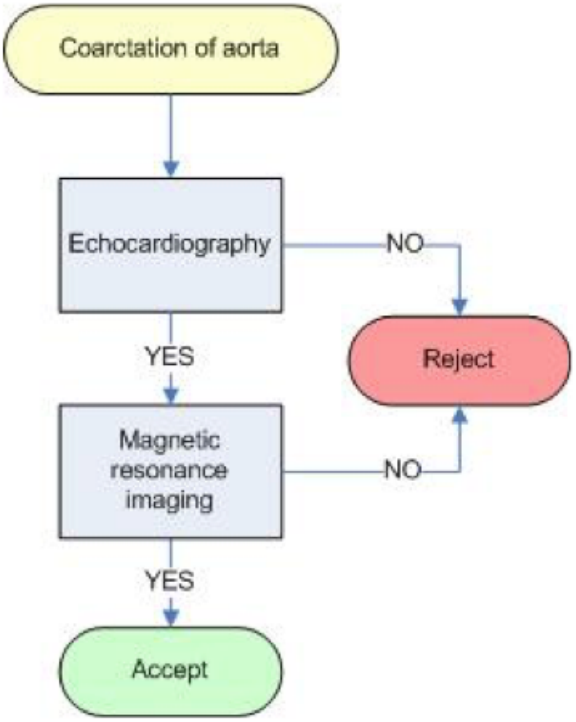


FIGURE A.7: Coarctation of aorta Algorithm

Coarctation of aorta	Yes	Yes
Echocardiography	?	Yes
Magnetic resonance imaging		?
Echocardiography	X	
Magnetic resonance imaging		X

FIGURE A.8: Coarctation of aorta Test Selection Decision Table

A.4 Glomerulonephritis

Glomerulonephritis is a group of diseases that injure the part of the kidney that filters blood (called glomeruli). Glomeruli remove excess fluid, electrolytes and waste from your bloodstream and pass them into your urine. When the kidney is injured, it cannot

Coarctation of aorta	Yes	Yes	Yes	Yes
Echocardiography	No	Yes		
Magnetic resonance imaging			Yes	No
Coarctation of aorta		?	X	

FIGURE A.9: Coarctation of aorta Evaluation Decision Table

get rid of wastes and extra fluid in the body. If the illness continues, the kidneys may stop working completely, resulting in kidney failure.

Glomerulonephritis can be acute or chronic. Acute glomerulonephritis can be a response to an infection such as strep throat or an abscessed tooth. It may be due to problems with your immune system overreacting to the infection. The chronic form of glomerulonephritis can develop over several years with no or very few symptoms. This can cause irreversible damage to your kidneys and ultimately lead to complete kidney failure.

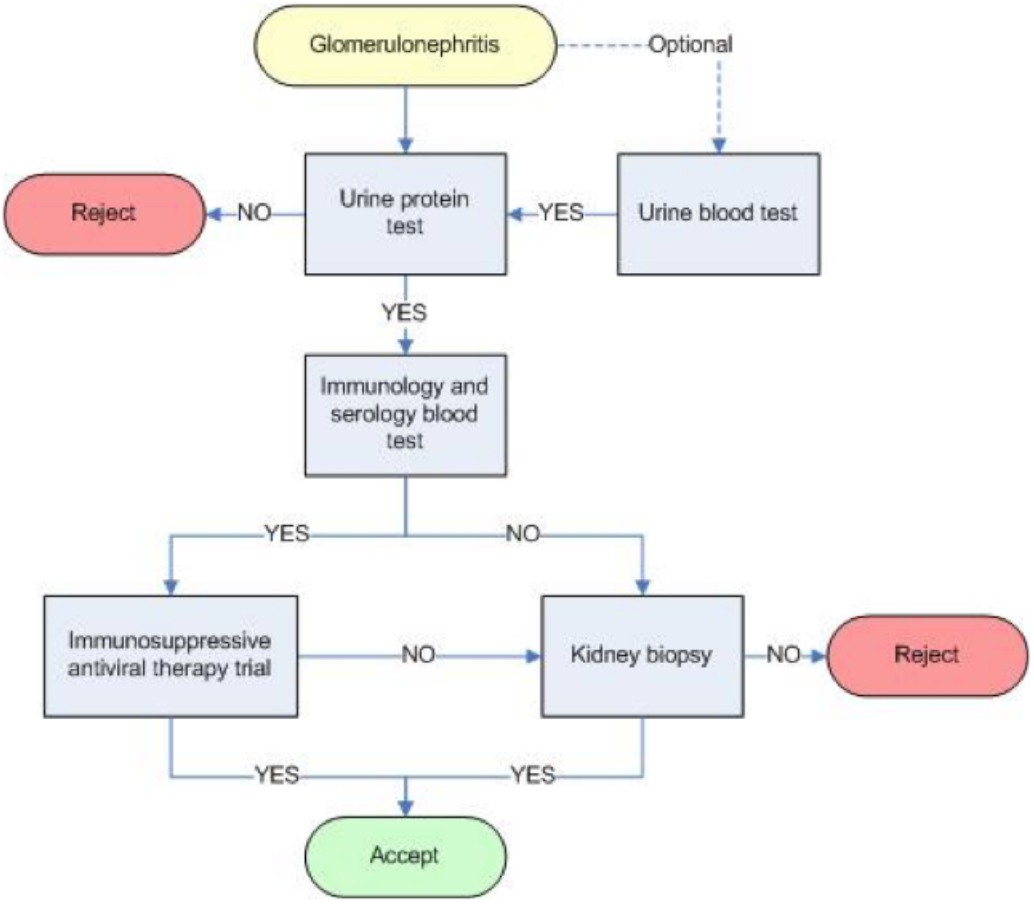


FIGURE A.10: Glomerulonephritis Algorithm

Figure A.10 shows the clinical algorithm obtained from the CPG [198].

Glomerulonephritis	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Immunology and serology blood test					?	Yes	Yes	No
Immunosuppressive antiviral therapy trial						?	No	
Kidney biopsy							?	?
Urine blood test	?	?	Yes	No				
Urine protein test	?	?	?	?	Yes	Yes	Yes	Yes
Immunology and serology blood test					X			
Immunosuppressive antiviral therapy trial						X		
Kidney biopsy							X	X
Urine blood test		X						
Urine protein test	X		X	X				

FIGURE A.11: Glomerulonephritis Test Selection Decision Table

Glomerulonephritis	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Immunology and serology blood test				?	Yes	Yes	Yes	Yes	Yes	No	No	No
Immunosuppressive antiviral therapy trial					?	Yes	No	No	No			
Kidney biopsy							?	Yes	No	?	Yes	No
Urine blood test	No	Yes										
Urine protein test	?	?	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Glomerulonephritis	?	?		?	?	X	?	X		?	X	

FIGURE A.12: Glomerulonephritis Evaluation Decision Table

A.5 Hyperparathyroidism

Hyperparathyroidism is an excess of parathyroid hormone in the bloodstream due to overactivity of one or more of the body's four parathyroid glands, located in the neck, near or attached to the back of the thyroid. The parathyroid glands produce parathyroid hormone, which helps maintain an appropriate balance of calcium in the bloodstream and in tissues that depend on calcium for proper functioning.

Hyperparathyroidism could be caused by a tumor, gland enlargement, or other structural problems of the parathyroid glands. This causes kidneys and intestines to absorb a larger amount of calcium. It also results in more calcium being removed from bones.

Figure A.13 shows the clinical algorithm obtained from the CPG [199].

A.6 Pheochromocytoma

A pheochromocytoma is a rare, usually noncancerous (benign) tumor that develops in cells in the center of an adrenal gland, which is called the adrenal medulla. The adrenal medulla is responsible for the normal production of adrenaline, which our body requires to help maintain blood pressure and to help cope with stressful situations. A tumor that arises from the adrenal medulla and overproduces adrenaline can be a deadly tumor because of the severe elevation in blood pressure it causes.

Figure A.16 shows the clinical algorithm obtained from the CPG [200].

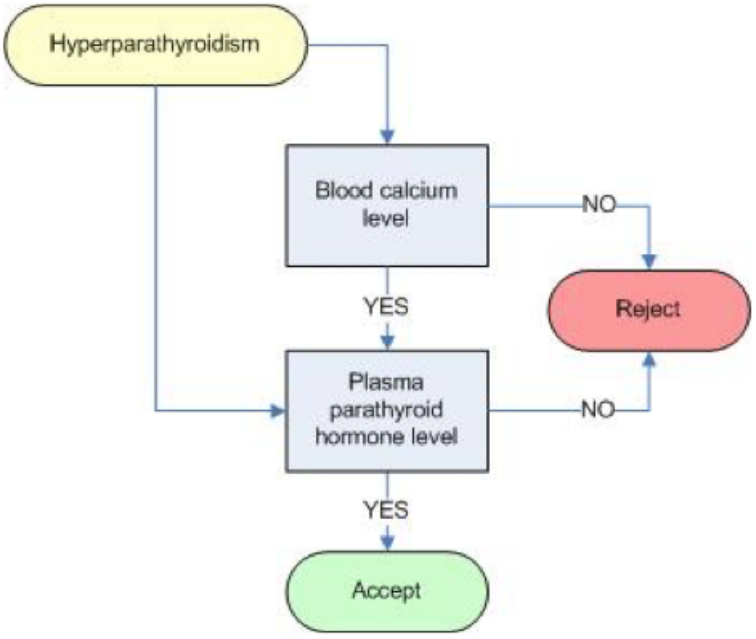


FIGURE A.13: Hyperparathyroidism Algorithm

Blood calcium level	?	?	Yes
Hyperparathyroidism	Yes	Yes	Yes
Plasma parathyroid hormone level	?	?	?
Blood calcium level	X		
Plasma parathyroid hormone level		X	X

FIGURE A.14: Hyperparathyroidism Test Selection Decision Table

Blood calcium level	Yes	No		
Hyperparathyroidism	Yes	Yes	Yes	Yes
Plasma parathyroid hormone level	?	?	Yes	No
Hyperparathyroidism	?		X	

FIGURE A.15: Hyperparathyroidism Evaluation Decision Table

A.7 Renovascular Hypertension

Renovascular hypertension is blood pressure elevation due to partial or complete occlusion of one or more renal arteries or their branches. When the kidneys receive low blood flow, they act as if the low flow is due to dehydration. So they respond by releasing hormones that stimulate the body to retain sodium and water. Blood vessels fill with additional fluid, and blood pressure goes up.

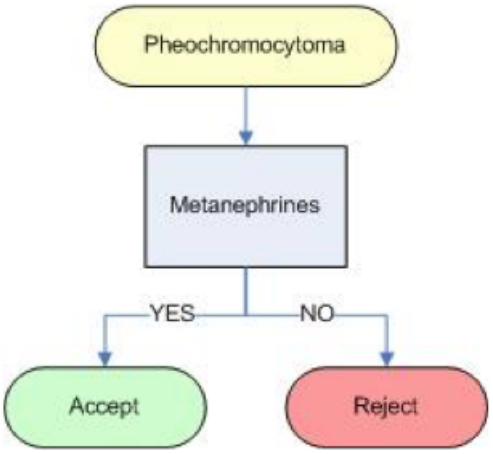


FIGURE A.16: Pheochromocytoma Algorithm

Metanephrines	?
Pheochromocytoma	Yes
Metanephrines	X

FIGURE A.17: Pheochromocytoma Test Selection Decision Table

Metanephrines	Yes	No
Pheochromocytoma	Yes	Yes
Pheochromocytoma	X	

FIGURE A.18: Pheochromocytoma Evaluation Decision Table

The narrowing in one or both renal arteries is most often caused by atherosclerosis, or hardening of the arteries. This is the same process that leads to many heart attacks and strokes.

Renovascular hypertension is the most frequent form of secondary hypertension. It is most often diagnosed among elderly patients and has significant effects on prognosis and patient outcomes.

Figure A.19 shows the clinical algorithm obtained from the CPG [201].

A.8 Sleep Apnea

Sleep apnea is a common disorder in which you have one or more pauses in breathing or shallow breaths while you sleep. Breathing pauses can last from a few seconds to minutes. This means the brain, and the rest of the body, may not get enough oxygen.

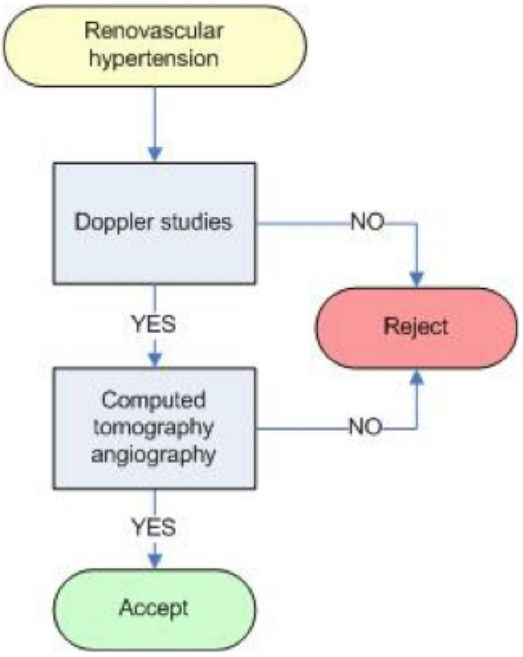


FIGURE A.19: Renovascular hypertension Algorithm

Computed tomography angiography		?
Doppler studies	?	Yes
Renovascular hypertension	Yes	Yes
Computed tomography angiography		**X**
Doppler studies	**X**	

FIGURE A.20: Renovascular hypertension Test Selection Decision Table

Computed tomography angiography		?	No	Yes
Doppler studies	No	Yes	Yes	Yes
Renovascular hypertension	Yes	Yes	Yes	Yes
Renovascular hypertension		**?**		**X**

FIGURE A.21: Renovascular hypertension Evaluation Decision Table

Typically, normal breathing then starts again, sometimes with a loud snort or choking sound.

Sleep apnea usually is a chronic condition that disrupts the sleep. When breathing pauses or becomes shallow, patients will often move out of deep sleep and into light sleep. As a result, the quality of their sleep is poor, which makes you tired during the day. Sleep apnea is a leading cause of excessive daytime sleepiness.

Figure A.22 shows the clinical algorithm obtained from the CPG [202].

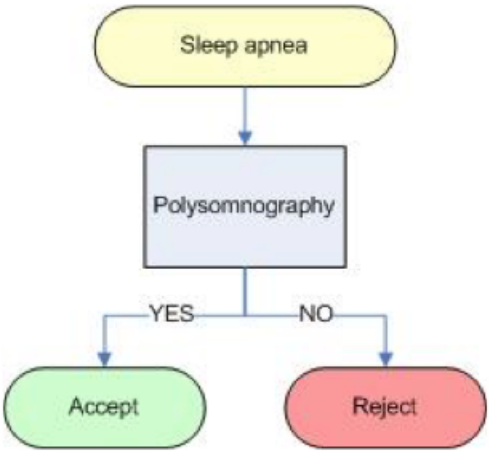


FIGURE A.22: Sleep apnea Algorithm

Polysomnography	?
Sleep apnea	Yes
Polysomnography	**X**

FIGURE A.23: Sleep apnea Test Selection Decision Table

Polysomnography	Yes	No
Sleep apnea	Yes	Yes
Sleep apnea	**X**	

FIGURE A.24: Sleep apnea Evaluation Decision Table

Appendix B

Emergency Shock Tables

B.1 Anaphylactoid Shock

The Anaphylactoid shock [227], also called anaphylaxis, is a severe, potentially life-threatening allergic reaction. It can occur within seconds or minutes of exposure to something you are allergic to, such as peanuts or the venom from a bee sting.

The flood of chemicals released by your immune system during anaphylaxis can cause you to go into shock; your blood pressure drops suddenly and your airways narrow, blocking normal breathing. Signs and symptoms of anaphylaxis include a rapid, weak pulse, a skin rash, and nausea and vomiting. Common triggers of anaphylaxis include certain foods, some medications, insect venom, and latex.

Anaphylaxis requires an immediate trip to the emergency department and an injection of epinephrine. If anaphylaxis is not treated right away, it can lead to unconsciousness or even death.

Figure B.1 shows the Anaphylactoid Shock Treatment Decision Table obtained from the CPG [206].

B.2 Cardiogenic Shock

Cardiogenic shock occurs if the heart suddenly cannot pump enough oxygen-rich blood to the body. The most common cause of cardiogenic shock is a damage of the heart muscle from a severe heart attack.

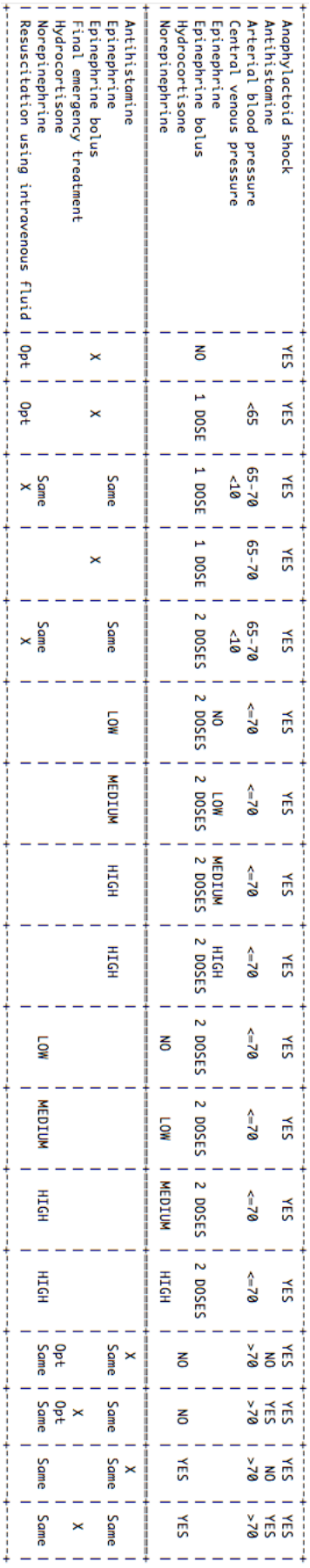


FIGURE B.1: Anaphylactoid Shock Table

Without enough oxygen-rich blood reaching the body's major organs, many problems can occur, organs can stop working well, their cells can die, and the organs may never recover again.

Cardiogenic shock is rare, but it is often fatal if not treated immediately. If it is treated immediately, about half of the people who develop the condition survive.

We consider two types of cardiogenic shock in this work, the AMI LV Cardiogenic Shock and the APE RV Cardiogenic Shock, and we obtain a different treatment decision table for each shock.

B.2.1 AMI LV Cardiogenic Shock

The damage in the heart muscle prevents the heart's main pumping chamber, the left ventricle (LV), from working well. As a result, the heart cannot pump enough oxygen-rich blood to the rest of the body.

A heart attack, also called a myocardial infarction, occurs when the flow of blood to the heart is blocked. The interrupted blood flow can damage or destroy part of the heart muscle. A heart attack can be fatal, but treatment has improved dramatically over the years.

Myocardial infarction with left ventricular (LV) failure remains the most common cause of cardiogenic shock [207].

Figure B.2 shows the AMI LV Cardiogenic Shock Treatment Decision Table obtained from the CPG [207].

B.2.2 APE RV Cardiogenic Shock

In about three percent of cardiogenic shock cases, the heart's lower right chamber, the right ventricle, does not work well. This means the heart cannot properly pump blood to the lungs, where it picks up oxygen to bring back to the heart and the rest of the body [228].

The Acute Pulmonary Embolism (APE) is a blockage in one of the pulmonary arteries in the lungs. In most cases, pulmonary embolism is caused by blood clots that travel to the lungs from the legs or, rarely, other parts of the body (deep vein thrombosis). APE can damage your heart and other organs in your body.

Figure B.3 shows the APE RV Cardiogenic Shock Treatment Decision Table obtained from the CPG [208].

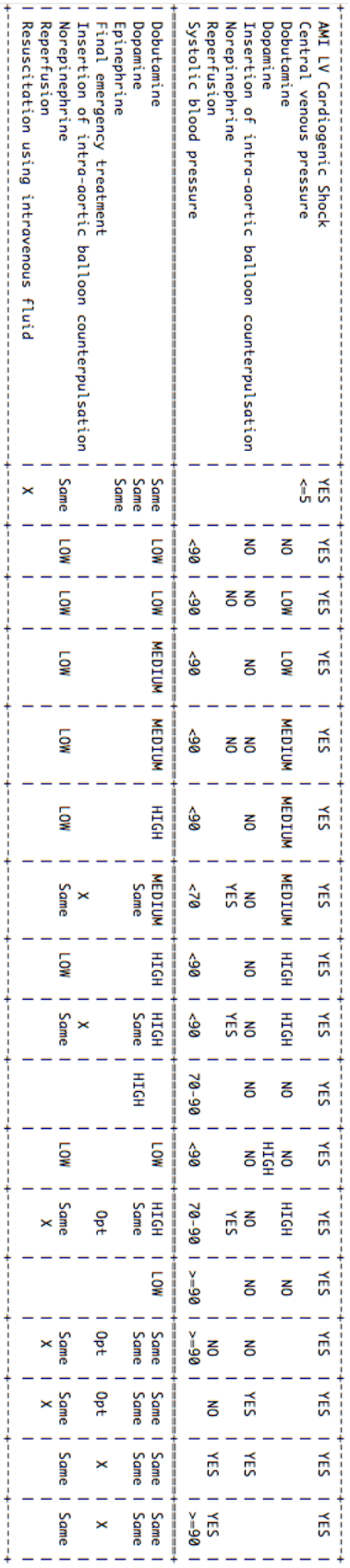


FIGURE B.2: AMI LV Cardiogenic Shock Table

Emergency Shock Tables

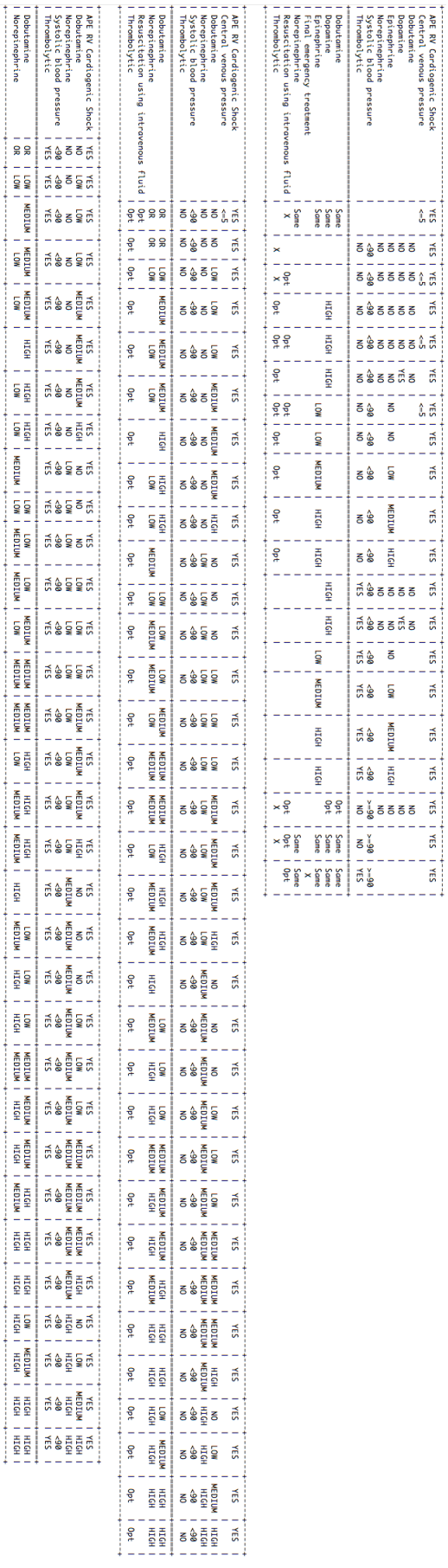


FIGURE B.3: APE RV Cardiogenic Shock Table

B.3 Cardiac Tamponade

Cardiac tamponade is an unusual pressure on the heart that occurs when blood or fluid builds up in the space between the heart muscle (myocardium) and the outer covering sac of the heart (pericardium).

In this condition, blood or fluid collects in the pericardium, the sac surrounding the heart. This prevents the heart ventricles from expanding fully. The excess pressure from the fluid prevents the heart from working properly. As a result, the body does not get enough blood.

Figure B.4 shows the Cardiac Tamponade Treatment Decision Table obtained from the CPG [209].

Cardiac tamponade	YES	YES	YES	YES	YES	YES	YES	YES
Central venous pressure		<10						
Dobutamine			NO	LOW	MEDIUM	HIGH		
Pericardiocentesis	NO	NO	NO	NO	NO	NO	YES	YES
Resuscitation using intravenous fluid		NO						
Systolic blood pressure		<100						
Dobutamine	**Same**	**Same**	**LOW**	**MEDIUM**	**HIGH**	**HIGH**	**Same**	
Final emergency treatment							X	X
Pericardiocentesis	X	Opt	Opt	Opt	Opt	Opt		
Resuscitation using intravenous fluid		X						

FIGURE B.4: Cardiac Tamponade Table

B.4 Hemorrhagic Shock

Hemorrhagic shock, also called hypovolemic shock, occurs when the body begins to shut down due to heavy blood loss. People suffering injuries that cause heavy bleeding may go into hemorrhagic shock if the bleeding is not stopped immediately. Common causes of hemorrhagic shock are: severe burns, deep cuts, gunshot wounds, trauma, and amputations.

According to the Mayo Clinic, hemorrhagic shock is the leading cause of death in people with traumatic injuries.

When heavy bleeding occurs, there is not enough blood flow to the organs in the body. Blood carries oxygen and other essential substances to your organs and tissues. When these substances are lost more quickly than they can be replaced, organs in the body begin to shut down.

Figure B.5 shows the Hemorrhagic Shock Treatment Decision Table obtained from the CPG [210].

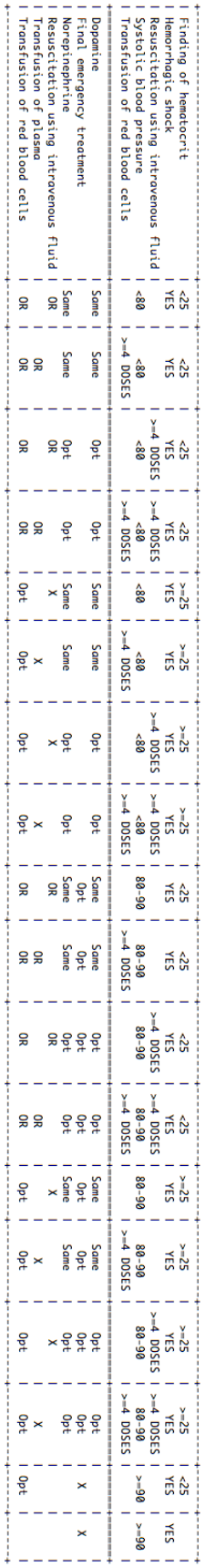


FIGURE B.5: Hemorrhagic Shock Table

B.5 Neurogenic Shock

Neurogenic shock may occurs after injuries in the spinal cord and when there is disruption in the blood circulation throughout the body due to injury or illness. It affects important nerves that make up the autonomic nervous system.

Injury to these nerves causes the walls of the blood vessels to relax resulting in slowing of the heart rate or bradycardia which can be fatal.

It is a serious and life-threatening condition, which requires prompt medical attention without any delay. If the treatment is delayed, then it causes irreversible tissue damage and even death. Out of the different types of the shocks, neurogenic shock is the most difficult to manage, mainly because of the irreversible damage to the tissues.

Figure B.6 shows the Neurogenic Shock Treatment Decision Table obtained from the CPGs [211, 212].

B.6 Septic Shock

Septic shock is what happens as a complication of an infection where toxins can initiate a full-body inflammatory response. It often occurs in people who are elderly or have a weakened immune system.

It is thought that the inflammation resulting from sepsis causes tiny blood clots to form, which can block oxygen and nutrients from reaching vital organs. As a result, the organs fail, causing a profound septic shock. This may cause a drop in blood pressure and may result in death. In fact, septic shock is the most common cause of death in intensive care units in the United States [229]

Figure B.7 shows the Septic Shock Treatment Decision Table obtained from the CPGs [213, 214].

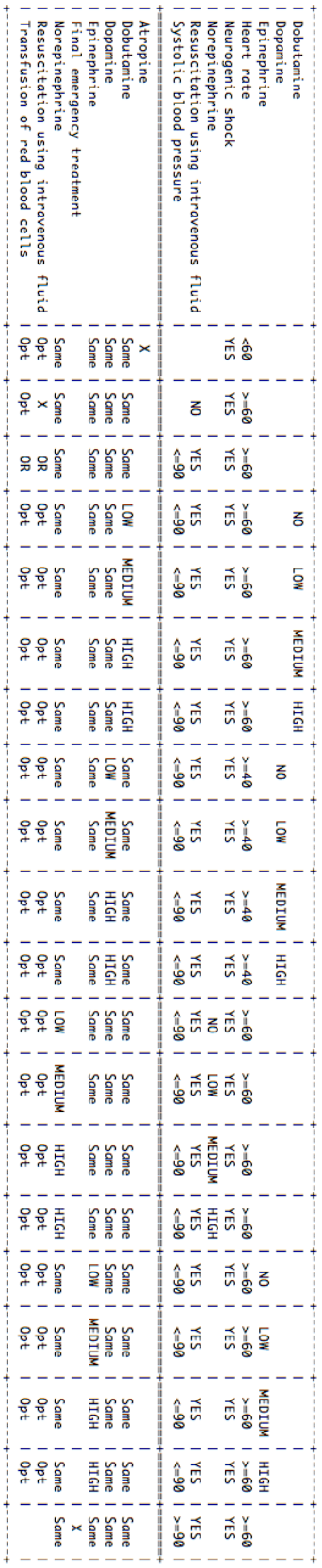


FIGURE B.6: Neurogenic Shock Table

Antibiotic therapy											
Arterial blood pressure			<65	<65	<65	<65	<65	<65	<65	<65	<65
Central venous pressure	<8	>12	8-12	8-12	8-12	8-12	8-12	8-12	8-12	8-12	8-12
Dobutamine											
Dopamine			NO	NO	YES						
Epinephrine											LOW
Finding of hematocrit											
Norepinephrine			NO	NO	NO	LOW	MEDIUM	MEDIUM	HIGH	HIGH	
Septic shock	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Superior vena cava oxygen saturation											
Antibiotic therapy	Opt	Opt	Opt	Opt	Opt	Opt	Opt	Opt	Opt	Opt	Opt
Diuretic		X									
Dobutamine	Same	Same									
Dopamine	Same	Same	HIGH								
Epinephrine	Same	Same								LOW	MEDIUM
Final emergency treatment											
Norepinephrine	Same	Same		LOW	LOW	MEDIUM	MEDIUM	HIGH	HIGH	HIGH	HIGH
Resuscitation using intravenous fluid	X										
Transfusion of red blood cells											
Vasodilators											

Antibiotic therapy										NO	YES
Arterial blood pressure	<65	<65	>90	65-90	65-90	65-90	65-90	65-90	65-90	65-90	65-90
Central venous pressure	8-12	8-12	8-12	8-12	8-12	8-12	8-12	8-12	8-12	8-12	8-12
Dobutamine					NO	LOW	MEDIUM	MEDIUM	HIGH		
Dopamine											
Epinephrine	MEDIUM	HIGH									
Finding of hematocrit				<30	>=30	>=30	>=30	>=30	>=30		
Norepinephrine											
Septic shock	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Superior vena cava oxygen saturation				<70	<70	<70	<70	<70	<70	>=70	>=70
Antibiotic therapy	Opt	Opt	Opt	Opt	Opt	Opt	Opt	Opt	Opt	X	Opt
Diuretic											
Dobutamine				Same	LOW	MEDIUM	MEDIUM	HIGH	HIGH	Same	Same
Dopamine				Same	Same	Same	Same	Same	Same	Same	Same
Epinephrine	HIGH	HIGH		Same	Same	Same	Same	Same	Same	Same	Same
Final emergency treatment										Opt	X
Norepinephrine	HIGH	HIGH		Same	Same	Same	Same	Same	Same	Same	Same
Resuscitation using intravenous fluid											
Transfusion of red blood cells				X							
Vasodilators			X								

FIGURE B.7: Septic Shock Table

Bibliography

[1] M.F. Collen. Origins of medical informatics. *Medical informatics [Special Issue] West J Med*, 145:778–785, December 1986.

[2] P. Braveman and L. Gottlieb. The social determinants of health: it's time to consider the causes of the causes. *Public Health Rep*, 129(Suppl 2):19–31, Jan-Feb 2014.

[3] R.H. Fletcher and S.W. Fletcher. Clinical practice guidelines. *Ann Intern Med*, 113(9):645–646, November 1990.

[4] R.A. Miller. Medical diagnostic decision support systems–past, present, and future: a threaded bibliography and brief commentary. *J Am Med Inform Assoc*, 1(1):8–27, January 1994.

[5] E.H. Shortliffe. Computer programs to support clinical decision making. *JAMA*, 258(1):61–66, July 1987.

[6] R. Klar and U. Bayer. Computer-assisted teaching and learning in medicine. *Int J Biomed Comput*, 26(1-2):7–27, 1990.

[7] D. Riaño, J.A. Bohada, A. Collado, and J.A. Lopez-Vallverdu. Mpm: A knowledge-based functional model of medical practice. *J Biom Inf*, 46(3): 379–387, 2013.

[8] M.A. Jones and D.A. Rivett. *Clinical reasoning for manual therapists (introduction to clinical reasoning).* Elsevier, 2004.

[9] N.B. Peek. *Decision-Theoretic Planning of Clinical Patient Management.* PhD thesis, Universiteit Utrecht, 2000.

[10] F. Rivas Echevarria and C. Rivas Echevarria. Application of expert systems in medicine. In *Proc 2006 Conference on Artificial Intelligence Research and Development*, pages 3–4, Amsterdam, The Netherlands, 2006.

[11] P.J. Lucas, L.C. van der Gaag, and A. Abu-Hanna. Bayesian networks in biomedicine and health-care. *Artif Intell Med*, 30(3):201–214, March 2004.

[12] M. Peleg. Computer-interpretable clinical guidelines: A methodological review. *J Biomed Inform*, 46(4):744–63, August 2013.

[13] Y.J. Son, H.G. Kim, E.H. Kim, S. Choi, and S.K. Lee. Application of support vector machine for prediction of medication adherence in heart failure patients. *Healthc Inform Res*, 16(4):253–259, 2010.

[14] M. Critchley. Medical dictionary. *Butterworths*, 1986.

[15] P.J.F. Lucas and A. Abu-Hanna. Prognostic methods in medicine. *Artif Intell Med*, 15(2):105–9, February 1999.

[16] J.A. Bohada, D. Riaño, and F. Real. Induction of partial orders to predict patient evolutions in medicine. *AIME 2007, LNAI 4594*, pages 489–499, 2007.

[17] P.G. Shekelle, S.C. Morton, and E.B. Keeler. Costs and benefits of health information technology. *Evid Rep Technol Assess (Full Rep)*, 132:1–71, 2006.

[18] B. Chaudhry, J. Wang, S. Wu, M. Maglione, W. Mojica, E. Roth, S.C. Morton, and P.G. Shekelle. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. *Ann Intern Med*, 144(10): 742–752, May 2006.

[19] C.M. DesRoches, E.G. Campbell, S.R. Rao, K. Donelan, T.G. Ferris, A. Jha, R. Kaushal, D.E. Levy, S. Rosenbaum, A.E. Shields, and D. Blumenthal. Electronic health records in ambulatory care - a national survey of physicians. *N Engl J Med*, 359(1):50–60, July 2008.

[20] P.C. Tang, J.S. Ash, D.W. Bates, J.M. Overhage, and D.Z. Sands. Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. *J Am Med Inform Assoc*, 13(2):121–126, Mar-Apr 2006.

[21] S.G. Peters and M.A. Khan. Electronic health records: current and future use. *J Comp Eff Res*, 3(5):515–522, September 2014.

[22] S. Bowman. Impact of electronic health record systems on information integrity: quality and safety implications. *Perspect Health Inf Manag*, 10:1c, October 2013.

[23] T.D. Gunter and N.P. Terry. The emergence of national electronic health record architectures in the united states and australia: models, costs, and questions. *J Med Internet Res*, 7(1):e3, March 2005.

[24] K. Häyrinen, K. Saranto, and P. Nykänen. Definition, structure, content, use and impacts of electronic health records: a review of the research literature. *Int J Med Inform*, 77(5):291–304, May 2008.

[25] S. Clamp and J. Keen. Electronic health records: is the evidence base any use? *Med Inform Internet Med*, 32(1):5–10, 2007.

[26] D.L. Hunt, R. Brian Haynes ans S.E. Hanna, and K. Smith. Effects of computer-based clinical decision support systems on physician performance and patient outcomes. a systematic review. *JAMA*, 280(15):1339–46, October 1998.

[27] A.A. Montgomery, T. Fahey, T.J. Peters, C. Macintosh, and D.J. Sharp. Evaluation of computer based clinical decision support system and risk chart for management of hypertension in primary care: randomised controlled trial. *BJM*, 320(7236):686–90, 2000.

[28] M.J. Romano and R.S. Stafford. Electronic health records and clinical decision support systems: impact on national ambulatory care quality. *Arch Intern Med*, 171(10):897–903, May 2011.

[29] D.F. Sittig and H. Singh. Rights and responsibilities of users of electronic health records. *CMAJ*, 184(13):1479–83, September 2012.

[30] J. Fox and R. Thomson. Clinical decision support systems: a discussion of quality, safety and legal liability issues. In *Proc AMIA Symp*, pages 265–269, 2002.

[31] R.N. Freed. Legal aspects of computer use in medicine. *Law and Contemporary Problems*, 32:674–706, 1967.

[32] E. Sezgin and S.O. Yıldırım. A literature review on attitudes of health professionals towards health information systems: From e-health to m-health. *Procedia Technology*, pages 1317–1326, 2014.

[33] C.L. Chi. *Medical decision support systems based on machine learning.* PhD thesis, University of Iowa, 2009.

[34] Z. Walter and M.S. Lopez. Physician acceptance of information technologies: Role of perceived threat to professional autonomy. *Decision Support Systems*, 46 (1):206–215, 2008.

[35] P.J.H. King. Decision tables. *The Comput J*, 10:4345–4353, 1967.

[36] M. Schaafsma, W.V.D. Deijl, J.M. Smits, A.O. Rahmel, P.F.D.V. Robbe, and A.J.Hoitsma. Decision tables and rule engines in organ allocation systems for optimal transparency and flexibility. *European Society for Organ Transplantation*, 24:433–440, 2011.

[37] R.N. Shiffman. Representation of clinical practice guidelines in conventional and augmented decision tables. *J Am Med Inform Assoc*, 4(5):382–393, 1997.

[38] J. Wartak, D. Fenna, E.T. Gelfand, and J.C. Callaghan. Diagnostic evaluation of chest pain using decision tables. *Computers in Cardiology*, pages 235–237, 1987.

[39] R.N. Shiffman and R.A. Greenes. Use of augmented decision tables to convert probabilistic data into clinical algorithms for the diagnosis of appendicitis. *Proc Ann Symp Comp Appl in Medical Care*, 1991.

[40] W.R. Booden. Adaptation of the decision table in teaching oral diagnosis. *J Dent Educ*, 41(10):626–629, 1977.

[41] D. Riaño. A systematic analysis of medical decisions. how to store knowledge and experience in decision tables. *Proc Int workshop KR4HC*, Springer LNAI 6924:23–36, 2011.

[42] F. Real, D. Riaño, and J.R. Alonso. Training residents in the application of clinical guidelines for differential diagnosis of the most frequent causes of arterial hypertension with decision tables. *KR4HC*, LNAI 8903:147–159, 2014.

[43] A. Baerheim. The diagnostic process in general practice: has it a two-phase structure? *Family Practice*, 18:243–245, 2001.

[44] D. Sackett and G. Haynes. *Clinical Epidemiology*. Lippincott Williams and Wilkins, 1991.

[45] P.M. Healey and E.J. Jacobson. *Common Medical Diagnoses: An Algorithmic Approach*. Saunders (Ed.), 2006.

[46] D.C. Hadorn. Use of algorithms in clinical guideline development in clinical practice guideline development: methodology perspectives. Agency HC Policy Res, 1995.

[47] Society for Medical Decision Making Committee on Standardization of Clinical Algorithms. Proposal for clinical algorithm standards. *Med Decis Making*, 12(2):149–154, Apr-Jun 1992.

[48] W. Siegenthaler. *Differential Diagnosis in Internal Medicine: From Symptom to Diagnosis*. Thieme Medical Publishers, New York, 2007.

[49] F.J. Ingelfinger, A.S. Relman, and M. Finland. *Controversy in internal medicine*. Saunders, 1966.

[50] M.C. Peterson, J.H. Holbrook, D. Von Hales, N.L. Smith, and L.V. Staker. Contributions of the history, physical examination, and laboratory investigation in making medical diagnoses. *West J Med.*, 156(2):163–165, February 1992.

[51] D. Kasper, J. Loscalzo, S. Hauser, A. Fauci, J. Jameson, and D. Longo. *Harrison's Principles of Internal Medicine*. McGraw-Hill, 19 edition, April 2015.

[52] P. Little, H. Everitt, I. Williamson, G. Warner, M. Moore, C. Gould, K. Ferrier, and S. Payne. Preferences of patients for patient centred approach to consultation in primary care: observational study. *BMJ*, 322(7284):468–472, February 2001.

[53] D.F. Cella. Quality of life: concepts and definition. *J Pain Symptom Manage*, 9 (3):186–192, 1994.

[54] D. Felce and J. Perry. Quality of life: its definition and measurement. *Res Dev Disabil*, 16(1):51–74, 1995.

[55] A. Leplège and S. Hunt. The problem of quality of life in medicine. *JAMA*, 278 (1):47–50, 1997.

[56] R.S. Morrison and D.E. Meier. Palliative care. *N Engl J Med*, 350:2582–2590, June 2004.

[57] D.L. Katz and A. Ali. Preventive medicine, integrative medicine, and the health of the public. Commissioned for the IOM Summit on Integrative Medicine and the Health of the Public, February 2009.

[58] J.M. Valderas, B. Starfield, B. Sibbald, C. Salisbury, and M. Roland. Defining comorbidity: Implications for understanding health and health services. *Ann Fam Med*, 7(4):357–363, 2009.

[59] J. Hilden and J.D.F. Habbema. Prognosis in medicine: An analysis of its meaning and rôles. *Theoretical Medicine*, 8(3):349–365, October 1987.

[60] A. Abu-Hanna and P.J.F. Lucas. Intelligent prog- nostic methods in medical diagnosis and treatment planning. In *Computational Engineering in Systems Applications (IMACS-IEEE, UCIS, Lille)*, 1998.

[61] J. Neuberger. Prognostic models - what is their future? *European Journal of Gastroenterology and Hepatology*, 9(12):1145–47, 1997.

[62] A. Abu-Hanna and P.J.F. Lucas. Prognostic models in medicine. ai and statistical approaches. *Methods Inf Med*, 40(1):1–5, 2001.

[63] M.J. Field and K.N. Lohr, editors. *Clinical Practice Guidelines: Directions for a New Program*. Committee to Advise the Public Health Service on Clinical Practice Guidelines, Institute of Medicine, 1990.

[64] R.M. Rosenfeld and R.N. Shiffman. Clinical practice guideline development manual: A quality-driven approach for translating evidence into action. *Otolaryngol Head Neck Surg*, 140(6 Suppl 1):S1–43, November 2009.

[65] D.C. Hadorn. Use of algorithms in clinical guideline development. In Agency for Health Care Policy and Research, editors, *Clinical Practice Guideline Development: Methodology Perspectives.*, volume 95-0009, pages 93–104, Rockville, MD, 2005.

[66] C.Z. Margolis. Uses of clinical algorithms. *JAMA*, 249(5):627–632, February 1983.

[67] G. Georg. Computerization of clinical guidelines: an application of medical document processing. In B.G. Silverman, A. Jain, A. Ichalkaranje, and L.C. Jain, editors, *Intelligent Paradigms for Healthcare Enterprises*, volume 184 of *Studies in Fuzziness and Soft Computing*, pages 1–30. Springer Berlin Heidelberg, 2005.

[68] F. Real, D. Riaño, and J.A. Bohada. Automatic generation of formal intervention plans based on the sda representation model. In *Proceedings of Computer Based Medical Systems*, pages 575–580, Maribor, Slovenia, 2007.

[69] P.B. Burns, R.J. Rohrich, and K.C. Chung. The levels of evidence and their role in evidence-based medicine. *Plastic and reconstructive surgery*, 128(1):305–310, July 2011.

[70] A.W. Chan and D.G. Altman. Epidemiology and reporting of randomised trials published in pubmed journals. *Lancet*, 365(9465):1159–1162, April 2005.

[71] S. Hopewell, S. Dutton, L.M. Yu, A.W. Chan, and D.G. Altman. The quality of reports of randomised trials in 2000 and 2006: comparative study of articles indexed in pubmed. *BMJ*, 340:c723, 2010.

[72] A.E. Bennett and K. Ritchie. *Questionnaires in medicine : a guide to their design and use*. London ; New York : Published for the Nuffield Provincial Hospitals Trust by the Oxford University Press, 1975.

[73] E. McColl, A. Jacoby, L. Thomas, J. Soutter, C. Bamford, N. Steen, R. Thomas, E. Harvey, A. Garratt, and J. Bond. Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients. *Health Technol Assess*, 5(31):1–256, 2001.

[74] P.M. Boynton and T. Greenhalgh. Hands-on guide to questionnaire research: Selecting, designing, and developing your questionnaire. *BMJ*, 328(7451): 1312–1315, 2004.

[75] P.M. Boynton. Hands-on guide to questionnaire research: Administering, analysing, and reporting your questionnaire. *BMJ*, 328(7452):1372–1375, 2004.

[76] J. Geddes. Answering clinical questions about prognosis. *Evid Based Mental Health*, 3:100–101, 2000.

[77] A.R. Artino Jr., J.S. La Rochelle, K.J. Dezee, and H. Gehlbach. Developing questionnaires for educational research: Amee guide no. 87. *Med Teach*, 36(6): 463–474, 2014.

[78] D.M. Gaba. The future vision of simulation in healthcare. *Simul Healthc*, 2(2): 126–135, 2007.

[79] J.C. Benneyan. An introduction to using computer simulation in healthcare: patient wait case study. *J Soc Health Syst*, 5(3):1–15, 1997.

[80] J. Barjis. Healthcare simulation and its potential areas and future trends. *SCS M&S Magazine*, 2(5):1–6, 2011.

[81] M.L. Good. Patient simulation for training basic and advanced clinical skills. *Med Educ*, Suppl 1:14–21, November 2003.

[82] D.A. Cook and M.M. Triola. Virtual patients: a critical literature review and proposed next steps. *Med Educ*, 43(4):303–311, 2009.

[83] R. Kneebone. Simulation in surgical training: educational issues and practical implications. *Med Educ*, 37(3):267–277, March 2003.

[84] Y. Okuda, E.O. Bryson, S. DeMaria Jr, L. Jacobson, J. Quinones, B. Shen, and A.I. Levine. The utility of simulation in medical education: What is the evidence? *Mt Sinai J Med*, 76(4):330–343, August 2009.

[85] W.C. McGaghie, S.B. Issenberg, E.R. Petrusa, and R.J. Scalese. A critical review of simulation-based medical education research: 2003-2009. *Med Educ*, 44 (1):50–63, January 2010.

[86] G.I. Mihalas, J. Zvarova, C. Kulikowski, M. Ball, J. van Bemmel, A. Hasman, I. Masic, D. Whitehouse, and B. Barber. History of medical informatics in europe - a short review by different approach. *Acta Inform Med*, 22(1):6–10, February 2014.

[87] G.I. Mihalas. Evolution of trends in european medical informatics. *Acta Inform Med*, 22(1):37–43, February 2014.

[88] T. Greenhalgh. Computer assisted learning in undergraduate medical education. *BMJ*, 322(7277):40–44, 2001.

[89] Shankaracharya, D. Odedra, S. Samanta, and A.S. Vidyarthi. Computational intelligence in early diabetes diagnosis: a review. *Rev Diabet Stud*, 7(4):252–262, February 2010.

[90] L.H. Eadie, P. Taylor, and A.P. Gibson. A systematic review of computer-assisted diagnosis in diagnostic cancer imaging. *Eur J Radiol*, 81(1): e70–76, February 2012.

[91] P. Hucklenbroich. Steps towards a theory of medical practice. *Theor Med Bioeth*, 19(3):215–228, 1998.

[92] G. Squires. Modelling medicine. *Med Educ*, 36(11):1077–82, 2002.

[93] A. Hasman and R. Haux. Modeling in biomedical informatics – an exploratory analysis (part 2). *Int J Med Inform*, 76:96–102, 2006.

[94] E. Maij, V.E. van Reijswoud, P.J. Toussaint, E.H. Harms, and J.H. Zwetsloot-Schonk. A process view of medical practice by modeling communicative acts. *Methods Inf Med*, 39(1):56–62, 2000.

[95] S. Malhotra, D. Jordan, E. Shortliffe, and V.L. Patel. Workflow modeling in critical care: piecing together your own puzzle. *J Biomed Inform*, 40(2):81–92, 2007.

[96] G.A. Gorry. Modelling the diagnostic process. *J Med Educ*, 45(5):293–302, May 1970.

[97] S.I. Herbert, C.J. Gordon, A. Jackson-Smale, and J.L.R. Salis. Protocols for clinical care. *Comput Methods Programs Biomed*, 48:21–26, 1995.

[98] G.T. Jun. Health care process modelling: which method when? *Int J Qual Health Care*, 21(3):214–224, 2009.

[99] C. Wiesemann. The significance of prognosis for a theory of medical practice. *Theor Med Bioeth*, 19(3):253–267, 1998.

[100] L. Ohno-Machado. Modeling medical prognosis: survival analysis techniques. *J Biomed Inform*, 34(6):428–39, 2001.

[101] A. Abu-Hanna and P.J.F. Lucas. Prognostic models in medicine (editorial). *Methods Inf Med*, 40:1–5, 2010.

[102] R. El-Kareh, O. Hasan, and G.D. Schiff. Use of health information technology to reduce diagnostic errors. *BMJ Qual Saf*, 0:1–12, 2013.

[103] J.M. Overhage, W.M. Tierney, X.H. Zhou, and C.J. McDonald. A randomized trial of "corollary orders to prevent errors of omission. *J Am Med Inform Assoc*, 4(5):364–375, 1997.

[104] S. Shea, W. DuMouchel, and L. Bahamonde. A meta-analysis of 16 randomized vontrolled trials to evaluate computer-based clinical reminder systems for preventative care in the ambulatory setting. *J Am Med Inform Assoc*, 3(6): 399–409, 1996.

[105] M. Peleg, S. Tu, J. Bury, P. Ciccarese, J. Fox, R.A. Greenes, R. Hall, P.D. Johnson, N. Jones, A. Kumar, S. Miksch, S. Quaglini, A. Seyfang, E.H. Shortliffe, and M. Stefanelli. Comparing computer-interpretable guideline models: A case-study approach. *Journal of the American Medical Informatics Association*, 10(1):52–68, January 2003.

[106] M. Peleg, A.A. Boxwala, O. Ogunyemi, Q. Zeng, S. Tu, R. Lacson, E. Bernstam, N. Ash, P. Mork, L. Ohno-Machado, E.H. Shortliffe, and R.A. Greenes. Glif3: the evolution of a guideline representation format. In *Proc AMIA Symp*, pages 645–649, 2000.

[107] P.D. Johnson, S.W. Tu, N. Booth, B. Sugden, and I.N. Purves. Using scenarios in chronic disease management guidelines for primary care. In *Proc AMIA Annu Fall Symp*, pages 389–393, 2000.

[108] S.W. Tu and M.A. Musen. A flexible approach to guideline modeling. In *Proc AMIA Symp*, pages 420–424, 1999.

[109] J. Fox and A. Rahmanzadeh. Disseminating medical knowledge: the proforma approach. *Artif Intell Med*, 14:157–181, 1998.

[110] Y. Shahar, S. Miksch, and P. Johnson. The asgaard project: a task-specific framework for the application and critiquing of time-oriented clinical guidelines. *Artif Intell Med*, 14:29–51, 1998.

[111] A. Seyfang, S. Miksch, and M. Marcos. Combining diagnosis and treatment using asbru. *International Journal of Medical Informatics*, 68(1):49–57, 2002.

[112] M.A. Musen, S.W. Tu, A.K. Das, and Y. Shahar. Eon: A component-based approach to automation of protocol-directed therapy. *Journal of the American Medical Informatics Association*, 3(6):367–388, 1996.

[113] A. Latoszek-Berendsen, P. de Clercq, J. van der Herik, and A. Hasman. Intention-based expressions in gastine. *Methods of Information in Medicine*, 48 (4):391–396, 2009.

[114] P.A. de Clercq, J.A. Blom, H.H.M. Korsten, and A. Hasman. Approaches for creating computer-interpretable guidelines that facilitate decision support. *Artificial Intelligence in Medicine*, 31(1):1–27, 2004.

[115] L. Anselma, P. Terenziani, S. Montani, and A. Bottrighi. Towards a comprehensive treatment of repetitions, periodicity and temporal constraints in clinical guidelines. *Artificial Intelligence in Medicine*, 38(2):171–195, 2006.

[116] P. Terenziani, G. Molino, and M. Torchio. A modular approach for representing and executing clinical guidelines. *Artificial Intelligence in Medicine*, 23(3): 249–276, 2001.

[117] A.A. Boxwala, S. Tu, M. Peleg, Q. Zeng, O. Ogunyemi, R.A. Greenes, E.H. Shortliffe, and V.L. Patel. Toward a representation format for sharable clinical guidelines. *Journal of Biomedical Informatics*, 34(3):157–169, 2001.

[118] A.A. Boxwala, M. Peleg, S. Tu, O. Ogunyemi, Q.T. Zeng, D. Wang, V.L. Patel, R.A. Greenes, and E.H. Shortliffe. Glif3: a representation format for sharable computer-interpretable clinical practice guidelines. *Journal of Biomedical Informatics*, 37(3):147–161, 2004.

[119] A. Skonetzki, H.J. Gausepohl, M. van der Haak, S. Knaebel, O. Linderkamp, and T. Wetter. Helen, a modular framework for representing and implementing clinical practice guidelines. *Methods Inf Med*, 43(4):413–426, 2004.

[120] S. Panzarasa, S. Madde, S. Quaglini, C. Pistarini, and M. Stefanelli. Evidence-based careflow management systems: the case of post-stroke rehabilitation. *Biomedical Informatics*, 35(2):123–139, 2002.

[121] S. Quaglini, M. Stefanelli, G. Lanzola, V. Caporusso, and S. Panzarasa. Flexible guideline-based patient careflow systems. *Artif Intell Med*, 22:65–80, 2001.

[122] P. Ciccarese, E. Caffi, S. Quaglini, and M. Stefanelli. Architectures and tools for innovative health information systems: The guide project. *International Journal of Medical Informatics*, 74(7):553–562, 2005.

[123] D.R. Sutton and J. Fox. The syntax and semantics of the proforma guideline modeling language. *Journal of the American Medical Informatics Association*, 10 (5):433–443, 2003.

[124] S.W. Tu, J.R. Campbell, J. Glasgow, M.A. Nyman, R. McClure, J. McClay, C. Parker, K.M. Hrabak, D. Berg, T. Weida, J.G. Mansfield, M.A. Musen, and R.M. Abarbanel. The sage guideline model: Achievements and overview. *J Am Med Inform Assoc*, 14(5):589–98, 2007.

[125] D. Riaño. The sda model: A set theory approach. In *Proc. IEEE Computer-Based Medical Systems*, pages 563–568, 2007.

[126] D. Wang, M. Peleg, S.W. Tu, E.H. Shortliffe, and R.A. Greenes. Representation of clinical practice guidelines for computer-based implementations. *Stud Health Technol Inform*, 84(Pt 1):285–289, 2001.

[127] D. Wang, M. Peleg, S.W. Tu, A.A. Boxwala, R.A. Greenes, V.L. Patel, and E.H. Shortliffe. Representation primitives, process models and patient data in computer-interpretable clinical practice guidelines: a literature review of guideline representation models. *Int J Med Inform*, 68(1-3):59–70, December 2002.

[128] N. Mulyar, W.M. van der Aalst, and M. Peleg. A pattern-based analysis of clinical computer-interpretable guideline modeling languages. *J Am Med Inform Assoc*, 14(6):781–787, Nov-Dec 2007.

[129] J. Houser and K.S. Oman. *Evidence-based Practice: An implementation guide for Healthcare Organizations*. Jones & Bartlett Learning, 2011.

[130] C. Bähler, C.A. Huber, B. Brüngger, and O. Reich. Multimorbidity, health care utilization and costs in an elderly community-dwelling population: a claims data based observational study. *BMC Health Services Research*, 15:23, January 2015.

[131] R.E. Upshur. Do clinical guidelines still make sense? no. *Ann Fam Med*, 12(4): 301, 2014.

[132] S.R. Abidi and S.S.R. Abidi. Towards the merging of multiple clinical protocols and guidelines via ontology-driven modeling. In *Artificial Intelligence in Medicine*, volume 5651 of *Lecture Notes in Computer Science*, pages 81–85. Springer Berlin Heidelberg, 2009.

[133] V. Zamborlini, R. Hoekstra, M. da Silveira, C. Pruski, A. ten Teije, and F. van Harmelen. A conceptual model for detecting interactions among medical recommendations in clinical guidelines: A case-study on multimorbidity. In

Knowledge Engineering and Knowledge Management, volume 8876 of *Lecture Notes in Computer Science*, pages 591–606. Springer International Publishing, 2014.

[134] J.E. Wilk, J.C. West, W.E. Narrow, S. Marcus, M. Rubio-Stipec, D.S. Rae, H.A. Pincus, and D.A. Regier. Comorbidity patterns in routine psychiatric practice: is there evidence of underdetection and underdiagnosis? *Compr Psychiatry*, 47(4): 258–264, Jul-Aug 2006.

[135] X. Hu, J. Huang, Y. Lv, G. Li, and X. Peng. Status of prevalence study on multimorbidity of chronic disease in china: systematic review. *Geriatr Gerontol Int*, 15(1):1–10, January 2015.

[136] D. Riaño, F. Real, J.A. López-Vallverdú, F. Campana, S. Ercolani, P. Mecocci, R. Annicchiarico, and C. Caltagirone. An ontology-based personalization of health-care knowledge to support clinical decisions for chronically ill patients. *Journal of Biomedical Informatics*, 45(3):429–446, 2012.

[137] R.A. Miller. Computer-assisted diagnostic decision support: history, challenges, and possible paths forward. *Adv Health Sci Educ Theory Pract*, 14(Suppl 1): 89–106, 2009.

[138] W.F. Bond, L.M. Schwartz, K.R. Weaver, D. Levick, M. Giuliano, and M.L. Graber. Differential diagnosis generators: an evaluation of currently available computer programs. *Journal of General Internal Medicine*, 27(2):213–219, 2012.

[139] E.P. Hoffer, M.J. Feldman, R.J. Kim, K.T. Famiglietti, and G.O. Barnett. Dxplain: Patterns of use of a mature expert system. In *AMIA Annu Symp Proc*, pages 321–324, 2005.

[140] P. Ramnarayan, G. Kulkarni, A. Tomlinson, and J. Britto. Isabel: a novel internet-delivered clinical decision support system. current perspectives in healthcare computing. Current perspectives in healthcare computing, 2004.

[141] J. Britto. Isabel at the helm. a web-based diagnosis system speeds clinical decisions for pediatric physicians. *Health Manag Technol*, 25(7):28–29, July 2004.

[142] L.R. Amy, S.M. Borowitz, P.A. Brown, M.J. Mendelsohn, and J.A. Lyman. Impact of a web-based diagnosis reminder system on errors of diagnosis. In *AMIA Annu Symp Proc*, page 843, 2006.

[143] P. Ramnarayan, G.C. Roberts, M. Coren, V. Nanduri, A. Tomlinson, P.M. Taylor, J.C. Wyatt, and J.F. Britto. Assessment of the potential impact of a

reminder system on the reduction of diagnostic errors: a quasi-experimental study. In *BMC Med Inform Decis Mak*, volume 6, page 22, 2006.

[144] B.A. Bauer, M. Lee, L. Bergstrom, D.L. Wahner-Roedler, J. Bundrick, S. Litin, E. Hoffer, R.J. Kim, K. Famiglietti, G.O. Barnett, and P. Elkin. Internal medicine resident satisfaction with a diagnostic decision support system (dxplain) introduced on a teaching hospital service. In *Proc AMIA Symp*, pages 31–35, 2002.

[145] P. Ramnarayan, R.R. Kapoor, M. Coren, V. Nanduri, A.L. Tomlinson, P.M. Taylor, J.C. Wyatt, and J.F. Britto. Measuring the impact of diagnostic decision support on the quality of clinical decision making: Development of a reliable and valid composite score. *J Am Med Inform Assoc*, 10(6):563–572, 2003.

[146] M.L. Graber, D. Tompkins, and J.J. Holland. Resources medical students use to derive a differential diagnosis. *Med Teach*, 31(6):522–527, 2009.

[147] S. London. Dxplain: a web-based diagnostic decision support system for medical students. *Med Ref Serv Q*, 17(2):17–28, 1998.

[148] P.L. Elkin, M. Liebow, B.A. Bauer, S. Chaliki, D. Wahner-Roedler, J. Bundrick, M. Lee, S.H. Brown, D. Froehling, K. Bailey, K. Famiglietti, R. Kim, E. Hoffer, M. Feldman, and G.O. Barnett. The introduction of a diagnostic decision support system (dxplain) into the workflow of a teaching hospital service can decrease the cost of service for diagnostically challenging diagnostic related groups (drgs). *Int J Med Inform*, 79(11):772–777, 2010.

[149] J.R. Hammersley and K. Cooney. Evaluating the utility of available differential diagnosis systems. In *Proc Annu Symp Comput Appl Med Care*, pages 229–231, 1988.

[150] E.S. Berner, G.D. Webster, A.A. Shugerman, J.R. Jackson, J. Algina, A.L. Baker, E.V. Ball, C.G. Cobbs, V.W. Dennis, E.P. Frenkel, L.D. Hudson, E.L. Mancall, C.E. Rackley, and O.D. Taunton. Performance of four computer-based diagnostic systems. *N Engl J Med*, 330:1792–1796, June 1994.

[151] E.S. Berner, J.R. Jackson, and J. Algina. Relationships among performance scores of four diagnostic decision support systems. *J Am Med Inform Assoc*, 3 (3):208–215, 1996.

[152] D. Riaño and F. Real. Performance of ddx generators when dealing with multimorbid cases. *Submitted to J Am Med Assoc*, 2016.

[153] S.L. Pollack, H.T. Hicks, and W.J. Harrison. *Decisión tables: theory and practice*. Wiley-Interscience, 1971.

[154] A.M. Moreno Garcia, M. Verhelle, and J. Vanthienen. An overview of decision table literature 1982 - 2000. Technical Report 0044, Katholieke Univ. Leuven, Dep. Toegepaste Economische Wetenschappen, 2000.

[155] J.D. Bezek. *Decision table language and its parallel execution architecture with applications in expert systems*. PhD thesis, State University of New York at Binghamton, Watson School of Engineering, Applied Science, and Technology, 1990.

[156] V. Chvalovsky. Decision tables. *Software: Practice and Experience*, 13(5): 423–429, 1983.

[157] A.I. Salah. *An Integration of Decision Tables and a Relational Database System into a Prolog Environment*. PhD thesis, University of Alabama, Huntsville, AL, USA, 1986.

[158] R. Canonne, J. Damret, J.P. Remezy, and D. Rollinat. Computerized inventory control with tables of decision. In *Proceedings of the IFIP WG 5.7 Working Conference on Advances in Production Management Systems*, pages 173–181. Elsevier Science, 1984.

[159] S. Biswas and V. Rajaraman. Algorithm to decide feasibility of linear integer constraints occurring in decision tables. *IEEE Transactions on Software Engineering*, 13(12):1340–1347, 1987.

[160] R. Kohavi and D. Sommerfield. Targeting business users with decision table classifiers, 1998.

[161] S. Lusignan. Codes, classifications, terminologies and nomenclatures: definition, development and application in practice. *Inform Prim Care*, 13(1):65–70, 2005.

[162] J.E. Mezzich. International surveys on the use of icd-10 and related diagnostic systems. *Psychopathology*, 35(2-3):72–75, 2002.

[163] B.G. Bentsen. International classification of primary care. *Scand J Prim Health Care*, 4(1):43–50, 1986.

[164] G.C. Miller and H. Britt. A new drug classification for computer systems: the atc extension code. *Int J Biomed Comput*, 40(2):121–124, October 1995.

[165] P. Ruch, J. Gobeill, C. Lovis, and A. Geissbühler. Automatic medical encoding with snomed categories. *BMC Med Inform Decis Mak*, 8(Suppl 1):S6, 2008.

[166] D. Riaño, F. Real, F. Campana, S. Ercolani, and R. Annicchiarico. An ontology for the care of the elder at home. In *Proceedings of the 12th Conference on Artificial Intelligence in Medicine: Artificial Intelligence in Medicine*, AIME '09, pages 235–239, Berlin, Heidelberg, 2009. Springer-Verlag.

[167] R. Kohavi. The power of decision tables. In *Proceedings of the 8th European Conference on Machine Learning*, ECML '95, pages 174–189, London, UK, 1995. Springer-Verlag.

[168] B.S.Yang, D.S. Lim, and A.C.C. Tan. Vibex: an expert system for vibration fault diagnosis of rotating machinery using decision tree and decision table. *Expert Systems with Applications*, 28(4):735–742, May 2005.

[169] G. Wets, F. Witlox, H. Timmermans, and J. Vanthienen. Locational choice modelling using fuzzy decision tables. In M. Smith, M. Lee, J. Keller, and J. Yen, editors, *Proceedings of the 1996 Biennial Conference of the North American Fuzzy Information Processing Society*, pages 80–84, 1996.

[170] G. Wets, F. Witlox, H. Timmermans, and J. Vanthienen. A fuzzy decision table approach for business site selection. *FUZZ-IEEE '96 - Proceedings of the fifth IEEE international conference on fuzzy systems*, 1-3:1605–1610, 1996.

[171] R. Hewett and J. Leuchner. The power of second-order decision tables. In *Proceedings of the 2002 SIAM International Conference on Data Mining*, 2002.

[172] Y. Tang and R. Meersman. Towards directly applied ontological constraints in a semantic decision table. In *Proceedings of the 5th International Conference on Rule-based Modeling and Computing on the Semantic Web*, RuleML'11, pages 193–207. Springer-Verlag, 2011.

[173] M.L. Schneider. Weighted decision tables - an alternative solution for ambiguity. *The Computer Journal*, 28(4):366–371, 1985.

[174] R.N. Shiffman and R.A. Greenes. Improving clinical guidelines with logic and decision-table techniques: application to hepatitis immunization recommendations. *Med Decis Making*, 14(3):245–254, 1994.

[175] R.N. Shiffman, L.L. Leape, and R.A. Greenes. Translation of appropriateness criteria into practice guidelines: application of decision table techniques to the rand criteria for coronary artery bypass graft. In *Proc Annu Symp Comput Appl Med Care*, pages 248–252, Washington DC., 1993.

[176] R.N. Shiffman and R.A. Greenes. Rule set reduction using augmented decision table and semantic subsumption techniques: application to cholesterol guidelines. *Proc Annu Symp Comput Appl Med Care*, pages 339–43, 1992.

[177] R.B. Mishra. A decision table and rule based interpretation system for epileptic discharges. *Int J Clin Monit Comput*, 9(3):165–178, 1992.

[178] R.R. Holland. Decision tables. their use for the presentation of clinical algorithms. *JAMA*, 233(5):455–457, August 1975.

[179] B. Williams. Case based learning—a review of the literature: is there scope for this educational paradigm in prehospital education? *Emerg Med J*, 22(8): 577–581, August 2005.

[180] J.G. Ruiz, M.J. Mintzer, and R.M. Leipzig. The impact of e-learning in medical education. *Acad Med*, 81(3):207–212, 2006.

[181] S. Kim. The future of e-learning in medical education: Current trend and future opportunity. *J Educ Eval Health Prof*, 3(3):3, September 2006.

[182] S.L. Clever, D.H. Novack, D.G. Cohen, and W. Levinson. Evaluating surgeons' informed decision making skills: pilot test using a videoconferenced standardised patient. *Med Educ*, 37(12):1094–1099, December 2003.

[183] D.K. Chan, T.H. Gallagher, R. Reznick, and W. Levinson. How surgeons disclose medical errors to patients: a study using standardized patients. *Surgery*, 138(5): 851–858, November 2005.

[184] D. Lee, R. Cornet, F. Lau, and N. de Keizer. A survey of snomed ct implementations. *Journal of Biomedical Informatics*, 46(1):87–96, February 2013.

[185] J.A. López-Vallverdú, D. Riaño, and J.A. Bohada. Improving medical decision trees by combining relevant health-care criteria. *Expert Systems with Applications*, 39(14):11782–11791, 2012.

[186] D.S. Bongala, R.M.F. Santos, M.M. Panaligan, F.E. Cercenia, A.A. Paulino, I.D.E. Ampil, J.D. Quebral, and A.L. Anastacio. Evidence-based clinical practice guidelines on the diagnosis and treatment of acute appendicitis. Philippine College of Surgeons. Committee on Surgical Infections., 2002.

[187] A.V. Chobanian, G.L. Bakris, H.R. Black, W.C. Cushman, L.A. Green, J.L. Izzo Jr, D.W. Jones, B.J. Materson, S. Oparil, J.T. Wright Jr, and E.J. Roccella. Seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure. *Hypertension*, 42(6): 1206–1252, December 2003.

[188] G. Mancia, R. Fagard, K. Narkiewicz, J. Redon, A. Zanchetti, M. Böhm, T. Christiaens, R. Cifkova, G. De Backer, A. Dominiczak, M. Galderisi, D.E.

Grobbee, T. Jaarsma, P. Kirchhof, S.E. Kjeldsen, S. Laurent, A.J. Manolis, P.M. Nilsson, L.M. Ruilope, R.E. Schmieder, P.A. Sirnes, P. Sleight, M. Viigimaa, B. Waeber, and F. Zannad. 2013 esh/esc guidelines for the management of arterial hypertension: the task force for the management of arterial hypertension of the european society of hypertension (esh) and of the european society of cardiology (esc). *J Hypertens*, 31(7):1281–1357, July 2013.

[189] D.D. Eisenhower. BrainyQuote.com. URL http://www.brainyquote.com/quotes/quotes/d/dwightdei164720.html.

[190] F. Real, D. Riaño, and J.R. Alonso. A patient simulation model based on decision tables for emergency shocks. in: Knowledge representation for health care. In D. Riaño, R. Lenz, S. Miksch, M. Peleg, M. Reichert, and A. ten Teije, editors, *Knowledge Representation for Health Care*, volume 9485 of *Lecture Notes in Computer Science*, pages 21–33. Springer International Publishing, 2015.

[191] P.A. James, S.Oparil, B.L. Carter, W.C. Cushman, C. Dennison-Himmelfarb, J. Handler, D.T. Lackland, M.L. LeFevre, T.D. MacKenzie, O. Ogedegbe, S.C. Smith Jr, L.P. Svetkey, S.J. Taler, R.R. Townsend, J.T. Wright Jr, A.S. Narva, and E. Ortiz. 2014 evidence-based guideline for the management of high blood pressure in adults: Report from the panel members appointed to the eighth joint national committee (jnc 8). *JAMA*, 311(5):507–20, February 2014.

[192] P.M. Kearney, M. Whelton, K. Reynolds, P.K. Whelton, and J. He. Worldwide prevalence of hypertension: a systematic review. *J Hypertens*, 22(1):11–19, 2004.

[193] M.S. Kallistratos, A. Giannakopoulos, and A.J. Manolis V. German. Diagnostic modalities of the most common forms of secondary hypertension. *Hellenic J Cardiol*, 51:518–529, 2010.

[194] S.J. Taler. Secondary causes of hypertension. *Prim Care*, 35(3):489–500, 2008.

[195] A. Duran Rodriguez-Hervada, J.A. Diaz Perez, P. Martin Rojas-Marcos, and A.L. Charro Salgado. Acromegalia. *Medicine*, 09:766–73, June 2004.

[196] J.M. Miralles Garcia. Hipercortisolismo de origen suprarrenal: síndrome de cushing. *Medicine*, 10:967–75, September 2008.

[197] J.A. Hurtado Martinez, F. Teruel Carrillo, A. Garcia Alberola, M. Valdes Mas, and M. Valdes Chavarri. Cardiopatías congénitas en el adulto. *Medicine*, 9(37): 2463–70, June 2005.

[198] J. Egido and J. Rojas-Rivera. Consideraciones generales y diagnóstico del síndrome nefrótico. *Medicine*, 10:5581–3, June 2011.

[199] J.A. Amado Señarís. Trastornos de las paratiroides. *Medicine*, 10:1057–62, September 2008.

[200] F.J. Tebar Masso and J.M. Rodriguez Gonzalez. Feocromocitoma. *Medicine*, 10: 997–1005, September 2008.

[201] M. Gago Fraile, G. Fernandez Fresnedo, and M. Arias Rodriguez. Protocolo diagnóstico de la patología vascular renal. *Medicine*, 10:5515–7, June 2011.

[202] E. Martinez-Ceron, I. Fernandez Navarro, and J. Fernandez Lahera. Síndrome de apneas hipopneas del sueño. *Medicine*, 10:4345–53, October 2010.

[203] F. Real and D. Riaño. An autonomous algorithm for generating and merging clinical algorithms. In D. Riaño, editor, *Knowledge Management for Health Care Procedures*, volume 5626 of *Lecture Notes in Computer Science*, pages 13–24. Springer Berlin Heidelberg, 2009.

[204] J.L. Vincent and D. Backer:. Circulatory shock. *N Engl J Med*, 369:1726–1734, 2013.

[205] S.A. Hill. Pharmacokinetics of drug infusions. *Contin Educ Anaesth Crit Care Pain*, 4(3):76–80, 2004.

[206] F.E. Simons, L.R. Ardusso, M.B. Bilo, V. Dimov, M. Ebisawa, Y.M. El-Gamal, D.K. Ledford, R.F. Lockey, J. Ring, M. Sanchez-Borges, G.E. Senna, A. Sheikh, B.Y. Thong, and M. Worm. 2012 update: World allergy organization guidelines for the assessment and management of anaphylaxis. *Curr Opin Allergy Clin Immunol*, 12(4):389–399, 2012.

[207] H.R. Reynolds and J.S. Hochman. Cardiogenic shock: current concepts and improving outcomes. *Circulation*, 117(5):686–697, February 2008.

[208] S. Konstantinides, A. Torbicki, G. Agnelli, N. Danchin, D. Fitzmaurice, N. Galiè, J.S. Gibbs, M. Huisman, M. Humbert, N. Kucher, I. Lang, M. Lankeit, J. Lekakis, C. Maack, E. Mayer, N. Meneveau, A. Perrier, P. Pruszczyk, L.H. Rasmussen, T.H. Schindler, P. Svítil, A. Vonk Noordegraaf, J.L. Zamorano, and M. Zompatori. 2014 esc guidelines on the diagnosis and management of acute pulmonary embolism. *European Heart Journal*, 35:3033–3080, 2014.

[209] B. Maisch, P.M. Seferović, A.D. Ristić, R. Erbel, R. Rienmüller, Y. Adler, W.Z. Tomkowski, G. Thiene, and M.H. Yacoub. Guidelines on the diagnosis and management of pericardial diseases executive summary; the task force on the diagnosis and management of pericardial diseases of the european society of cardiology. *Eur Heart J*, 25(7):587–610, 2004.

[210] S. Pascoe and J. Lynch. Management of hypovolaemic shock in the trauma patient. Adult Trauma Clinical Practice Guidelines, 2007.

[211] L. Levi, A. Wolf, and H. Belzberg. Hemodynamic parameters in patients with acute cervical cord trauma: description, intervention, and prediction of outcome. *Neurosurgery*, 33(6):1007–1016, 1993.

[212] Consortium for Spinal Cord Medicine. Early acute management in adults with spinal cord injury: a clinical practice guideline for health-care professionals. *J Spinal Cord Med*, 31(4):403–479, 2008.

[213] E. Rivers, B. Nguyen, S. Havstad, J. Ressler, A. Muzzin, B. Knoblich, E. Peterson, and M. Tomlanovich. Early goal-directed therapy in the treatment of severe sepsis and septic shock. *N Engl J Med*, 8(345):1368–1377, 2001.

[214] R.P. Dellinger, M.M. Levy, A. Rhodes, D. Annane, H. Gerlach, S.M. Opal, J.E. Sevransky, C.L. Sprung, I.S. Douglas, R. Jaeschke, T.M. Osborn, M.E. Nunnally, S.R. Townsend, K. Reinhart, R.M. Kleinpell, D.C. Angus, C.S. Deutschman, F.R. Machado, G.D. Rubenfeld, S. Webb, R.J. Beale, J.L. Vincent, and R. Moreno. Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock, 2012. *Intensive Care Med*, 39(2):165–228, 2013.

[215] Y. Hiraiwa, N. Kusunoki, I. Araki, T. Monji, T. Shinagawa, M. Nagamori, and R. Takeda. A case report of cushing syndrome associated with nephrotic syndrome. *Nihon Naika Gakkai Zasshi*, 77(7):1099–1100, July 1988.

[216] J. Baughan, C. de Gara, and D. Morrish. A rare association between acromegaly and pheochromocytoma. *Am J Surg*, 182(2):185–187, 2001.

[217] K.K. Behera, A. Nanaiah, A. Gupta, and S. Rajaratnam. Neurofibromatosis type 1, pheochromocytoma with primary hyperparathyroidism: A rare association. *Indian J Endocrinol Metab*, 17(2):349–351, March 2013.

[218] J.H. Myers and J.J. Eversman. Acromegaly, hyperparathyroidism, and pheochromocytoma in the same patient. a multiple endocrine disorder. *Arch Intern Med*, 141(11):1521–1522, October 1981.

[219] M.V. Davi, L. Dalle Carbonare, A. Giustina, M. Ferrari, A. Frigo, V. Lo Cascio, and G. Francia. Sleep apnoea syndrome is highly prevalent in acromegaly and only partially reversible after biochemical control of the disease. *Eur J Endocrinol*, 159(5):533–540, 2008.

[220] S.M. Smith, H. Soubhi, M. Fortin, C. Hudon, and T. O'Dowd. Managing patients with multimorbidity: systematic review of interventions in primary care and community settings. *BMJ*, 345, 2012.

[221] B.W. Ward and J.S. Schiller. Prevalence of multiple chronic conditions among us adults: estimates from the national health interview survey, 2010. *Prev Chronic Dis*, 10:E65, April 2013.

[222] B.W. Ward, J.S. Schiller, and R.A. Goodman. Multiple chronic conditions among us adults: a 2012 update. *Prev Chronic Dis*, 11:E62, April 2014.

[223] M. E. Tinetti, T.R. Fried, and C.M. Boyd. Designing health care for the most common chronic condition-multimorbidity. *JAMA*, 307(23):2493–2494, June 2012.

[224] A.K. Parekh, R. Kronick, and M. Tavenner. Optimizing health for persons with multiple chronic conditions. *JAMA*, 312(12):1199–1200, 2014.

[225] C.O. Weiss, R. Varadhan, M.A. Puhan, A. Vickers, K. Bandeen-Roche, C.M. Boyd, and D.M. Kent. Multimorbidity and evidence generation. *J Gen Intern Med*, 29(4):653–660, 2014.

[226] K. Uhlig, B. Leff, D. Kent, S. Dy, K. Brunnhuber, J.S. Burgers, S. Greenfield, G. Guyatt, K. High, R. Leipzig, C. Mulrow, K. Schmader, H. Schunemann, L.C. Walter, J. Woodcock, and C.M. Boyd. A framework for crafting clinical practice guidelines that are relevant to the care and management of people with multimorbidity. *J Gen Intern Med*, 29(4):670–679, 2014.

[227] Clinica mayo, 2015. URL http://www.mayoclinic.org.

[228] A.K. Jacobs, J.A. Leopold, E. Bates, L.A. Mendes, L.A. Sleeper, H. White, R. Davidoff, J. Boland, S. Modur, R. Forman, and J.S. Hochman. Cardiogenic shock caused by right ventricular infarction: a report from the shock registry. *Journal of the American College of Cardiology*, 41(8):1273–1279, 2003.

[229] S.J. Fitch and J.R. Gossage. Optimal management of septic shock. rapid recognition and institution of therapy are crucial. *Postgrad Med*, 111(3):53–56, 2002.