Expert Systems With Applications 186 (2021) 115749

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Extended Hapicare: A telecare system with probabilistic diagnosis and self-adaptive treatment

Hossain Kordestani [a,c,*], Roghayeh Mojarad [b], Abdelghani Chibani [b], Kamel Barkaoui [c], Yacine Amirat [b], Wagdy Zahran [a]

[a] Department of Research and Innovation, Maidis SAS, Chatou, France
[b] Laboratory of Image, Signal & Intelligent Systems (LISSI), University of Paris Est-Paris Créteil, Vitry sur Seine, France
[c] Centre d'études et de recherche en informatique et communications, Conservatoire National des Arts et Métiers, Paris, France

ARTICLE INFO

Keywords:
Medical telemonitoring
Bayesian reasoning
Answer set programming
Self-adaptive systems

ABSTRACT

The massive growth of the population with chronic diseases calls for a telecare system to enhance their quality of life and reduce their treatment costs. Most of the current solutions depend on reliable data, deterministic rules, or the similarity of patients, while studies have shown otherwise. To this end, in this paper, we have extended our previous work on the Hapicare framework to integrate probabilistic diagnosis and self-adaptive treatment. Our new framework enables sensors' datastream analysis and online decision-making. Its ontology-based reasoning uses Systematized Nomenclature of Medicine - Clinical Terms (SNOMED-CT) ontology to add contextual information to the collected data. Moreover, probabilistic reasoning is applied for diagnosis and screening to manage the uncertainty and unreliability of data as well as the indeterministic medical rules. The treatment system is designed to be modifiable by the experts and automatically adaptable to patients' needs. The probabilistic diagnosis performance has been evaluated based on two public datasets regarding symptoms and risk factors of two chronic diseases: chronic kidney disease and dermatologic disease. The results show that our solution outperforms a classical classifier specifically when more than 40% of the data are missing. The proposed framework is also validated using four scenarios. The evaluation results demonstrate the ability of the proposed framework to help patients and doctors diagnose and treat medical conditions and episodes.

1. Introduction

Chronic diseases have created one of the biggest challenges in public health, as they are the leading cause of morbidity and mortality (Harris, 2019), in particular, during the pandemic of COVID-19 that increases the mortality of patients with chronic diseases (Paget et al., 2019). However, the quality of life and life expectancy of patients with chronic diseases can be improved using the existing knowledge (WHO, 2005). Nevertheless, given a high number of patients with chronic diseases, managing them would require a lot of medical efforts. To this end, the use of telecare has been favored. Telecare is a general term of combining the Greek prefix *tele*, meaning at a distance, and the word *care*, which Miller and O'Toole (2003b) have defined as the use of information and telecommunications technologies to monitor patients and deliver health care to them remotely. Telecare can be used for managing medical conditions, including the important acute events during medical conditions, which are called *episodes* (Miller & O'Toole, 2003a). For instance, a cardiac arrest is an acute health event that may occur during a cardiovascular medical condition; i.e., cardiovascular disease is a medical condition, while cardiac arrest is an episode. Moreover, hypoglycemic episode commonly occurs in patients suffering from Diabetes Mellitus (DM) and Chronic Kidney Disease (CKD).

Studies have shown that telecare enables patients to feel safe and reassured. It also provides an opportunity for better treatment to the physicians (de Bruin et al., 2018; Bujnowska-Fedak & Grata-Borkowska, 2015). Predominantly, the technology is used to remove the physical barrier between the medical team and patients to enable treating patients at their homes. For example, Bhatti et al. (2018) focus on treating patients in a remote area. However, the research is trending to facilitate the treatment by providing useful suggestions to patients and comprehensive information to their doctors. Parati et al. (2018)

Abbreviations: IoT, Internet of Things; ASP, Answer Set Programming; SNOMED-CT, Systematized Nomenclature of Medicine - Clinical Terms; DS, Dempster–Shafer; BN, Bayesian Network; DM, Diabetes Mellitus; CKD, Chronic Kidney Disease; CPT, Conditional Probability Table

* Corresponding author at: Centre d'études et de recherche en informatique et communications, Conservatoire National des Arts et Métiers, Paris, France.
E-mail addresses: hossain.kordestani@maidis.fr (H. Kordestani), roghayeh.mojarad@u-pec.fr (R. Mojarad), abdelghani.chibani@u-pec.fr (A. Chibani), kamel.barkaoui@cnam.fr (K. Barkaoui), yacine.amirat@u-pec.fr (Y. Amirat), wagdy.zahran@maidis.fr (W. Zahran).

https://doi.org/10.1016/j.eswa.2021.115749
Received 17 February 2020; Received in revised form 17 February 2021; Accepted 8 August 2021
Available online 28 August 2021
0957-4174/

have shown that telecare can better control the patient's condition and support doctors to optimize the treatment and, consequently, decrease healthcare expenditure.

Using telecare with different sensors would result in the holistic monitoring of patients. Nevertheless, there are various pitfalls to avoid (Bujnowska-Fedak & Grata-Borkowska, 2015). The essential principle in designing such systems is customizability; because each patient has a unique requirement, and a simplistic design might not be useful for all patients (Sultan et al., 2018; Vitacca et al., 2018). Another overlooked problem of such systems is the uncertainty of collected data, which would hugely affect the use of that data (Knowles et al., 2018). Monitoring vital signs, activities, and any other aspects of human life and health patterns must consider the heterogeneity of the different source types producing observations and the uncertainty of the observations.

We have proposed Hapicare in our previous work, a reactive telemonitoring system with probabilistic reasoning using Bayesian Network (BN) (Kordestani et al., 2019). BN is a probabilistic graphical model representing the initial knowledge and relationships among variables of complex systems using a directed acyclic graph. It also considers expert knowledge, empirical data, and their uncertainties to enable probabilistic reasoning at the same time. Regarding the categories of rules provided by Object Management Group, a computer industry standards consortium (OMG, 2009); in Hapicare, two types of rules are embedded in BN, namely (1) production rules that are used for diagnosis of episodes and (2) reactive rules that are used for selection of treatments based on the episode. The reactive rules are needed to be dynamic; however, the rules embedded in BN are hardly modifiable. To overcome this limitation, we have extended Hapicare in this paper to handle the uncertainty of data and rules for diagnosing the medical conditions/episodes of patients, and to automatically self-customize the treatment procedure based on patients' experiences, which is called self-adaptive treatment in the rest of this paper. Although the objective of telecare is holistic monitoring of patients, nonetheless, the proposed system can overcome missing data. The main contributions of the proposed Extended Hapicare are the following:

- An IoT-based telemonitoring system that considers the medical files for holistic monitoring of patients with chronic diseases
- Ontology-based reasoning for obtaining contextual information
- A probabilistic diagnosis: a computer-aided diagnosis to handle missing data, uncertain data, and probabilistic rules
- A self-adaptive treatment: A treatment service based on Answer Set Programming (ASP) with easily modifiable rules and automatically customizing itself for each patient

In Extended Hapicare, we have integrated a computer-aided diagnosis for screening the health of patients. The use of computers and smart systems for medical diagnosis has started roughly at the same time as the rest of computer-based systems in the 1950s. Nowadays, there are two types of approaches, either using *expert systems* or *machine learning methods* (Yanase & Triantaphyllou, 2019). On the other hand, according to Pramanik et al. (2017), smart health systems can be *pervasive*, *reactive*, or *hospital-based*. In Extended Hapicare, we have implemented hybrid methods to benefit from both *expert systems* and *machine learning methods* in a *reactive* health monitoring system. Extended Hapicare has been validated by clinicians and medical systems engineers collaborating with Maidis in the context of the ITEA3 Medolution project[1] and recent proof-of-concepts projects. Doctors have been involved since the first design steps of the implemented systems. Due to legal and ethical restrictions, the validation of Extended Hapicare by involving patients in clinical tests is under investigation.

The remainder of the paper is organized as follows. First, we go through the state of the art as well as the required theoretical basis of this study in Section 2. Next, Extended Hapicare is described in detail in Section 3, while in Section 4, the evaluation and the experimental results are presented. The discussion regarding the observations in experiments is provided in Section 5. Lastly, we conclude the paper with the conclusion in Section 6.

2. Literature review

2.1. Telecare

Telecare has received massive attention; particularly after the emergence of IoT sensors, researchers tend to apply them for telecare and telemedicine. The telecare systems can be arguably divided into three groups in terms of functionality: (1) telemonitoring: collecting patient's information in real-time, (2) telemonitoring with alerts: in addition to telemonitoring, raising the alarm based on patients' situations, and (3) teletreatment: in addition to telemonitoring with alerts, suggesting treatments to patients based on their conditions. Most of the existing telecare systems are limited to telemonitoring systems with alerts.

Telecare systems are commonly targeted to the public, but some are specialized for patients with a specific type of chronic condition. For instance, Glykas and Chytas (2004), Parati et al. (2018), and Bucholc et al. (2019) have focused on high blood pressure, asthma, and Alzheimer's disease, respectively. Parati et al. (2018) have proposed a solution for patients to better control their condition and support the doctor for better follow-up. They point out that blood pressure telemonitoring allows improving the quality of lives of patients, improving the treatments, decreasing the face-to-face consultation sessions, and reducing the costs of healthcare (Parati et al., 2018). While Glykas and Chytas (2004) have introduced a web-based tool called AsthmaWeb. It accomplishes data gathering and monitoring to manage patients according to their personalized asthma action plan. Similarly, Bucholc et al. (2019) have introduced a decision support system to predict the severity of Alzheimer's disease based on biological and clinical measures.

In recent years, many researchers have started working on *teletreatment* systems. For instance, Xu et al. (2017) have proposed Cloud-MHMS, a monitoring system to help doctors diagnose patients' conditions better. In this system, patients' information is collected through their mobiles. Moreover, it uses process mining and alpha algorithm to propose a treatment plan based on similar patients' medical files. Even though Cloud-MHMS is useful for doctors' holistic diagnosis, it does not engage patients in their treatment. Likewise, Wong et al. (2018) have focused on reducing adverse drug events in intensive care units. This study discusses the effectiveness of decision systems regarding their rate of overridden alerts. Another approach in teletreatment is using the smart environment to help patients. For instance, Loreti et al. (2019) have proposed a framework intertwining a rule-based complex event processing with reactive event calculus to suggest a reaction based on patients' states.

Many studies focus on using IoT sensors and mobile for data collection; however, Habib et al. (2019) have designed a decision system based on wireless body sensor networks. As a subset of wireless sensor networks, the latter enables continuous monitoring of patients' vital signs, which is useful for patients in a critical state. In this study, the fuzzy inference system is used to calculate the weight of patients' risk. This study, like the mainstream of studies, overlooked personalization and customizability. On the contrary, Afzal et al. (2018) have introduced a mechanism to personalize wellness recommendations, using contextual information, e.g., location and weather, along with the recommendations. In this study, the general health recommendations are personalized to provide the ones that are best suited, based on the requirements, interests, and demands of the user.

Similarly, Rahimi and Wang (2013) have designed and implemented a framework to run a patient-specific clinical decision model. It enables collecting the patients' preferences and selecting one of the various options for them. This framework relies on decision trees, which allows a full customizable decision model; however, it requires patients to respond to multiple questionnaires, which is inconvenient for most patients.

[1] https://itea3.org/project/medolution.html.

Table 1
Features of hypercortisolism and their according probabilities (Friedman, 2015).

Feature	Percentage of patients
Fat redistribution	95
Menstrual irregularities	80
Thin skin and plethora	80
Moon facies	75
Increased appetite	75
Sleep disturbances	75
Nocturnal hyperarousal	75
Hypertension	75
Hypercholesterolemia and hypertriglyceridemia	70
Altered mentation	70
Diabetes mellitus and glucose intolerance	65
Striae	65
Hirsutism	65
Proximal muscle weakness	60
Psychological disturbances	50
Decreased libido and erectile dysfunction	50
Acne	45
Osteoporosis and pathological fractures	40
Easy bruisability	40
Poor wound healing	40
Virilization	20
Edema	20
Increased infections	10
Cataracts	5

2.2. Uncertainty

In the field of reasoning, uncertainty is classified into two categories: (1) *Aleatory* uncertainty is the intrinsic changing behavior, i.e., the observations differ in each experiment. (2) *Epistemic* uncertainty is rooted in the insufficiency of knowledge, i.e., principally, this type of uncertainty can be avoided with additional knowledge, and hence it is reducible (Billinton & Huang, 2008; Kiureghian & Ditlevsen, 2009). In the field of telecare, both types of uncertainties are possible. Faulty sensors, system failures, and human errors cause aleatory uncertainty while lacking enough information about patients, their medical files, and family history result in epistemic uncertainty. Since the medical rules are obtained during study and experiments, they are rarely absolute and deterministic. For instance, Table 1 depicts the features of hypercortisolism and their probabilities that Friedman (2015) has provided. It shows that the medical rule for hypercortisolism diagnosis using its features would consist of probabilistic relationships, and hence, it would be an uncertain rule. Several approaches exist for handling uncertainty, such as BN (Cooper & Herskovits, 1992), Dempster–Shafer (DS) (Dempster, 2008; Shafer, 1992), and fuzzy logic; (Zadeh, 1965); each of them is suitable for a specific purpose. DS is beneficial for gathering uncertain information from different sources and reasoning to a conclusion. However, fuzzy logic is a better fit when the states with low probability (membership values) are vital, e.g., diagnosing the early stage of a disease, since DS and BN mainly focus on the states with high probabilities. Verbert et al. (2017) have thoroughly compared the DS and BN; the overall summary of their comparison is shown in Table 2. As the BN is more suitable for the purpose of this paper, it is discussed in more detail.

2.2.1. Bayesian Network

Bayesian Network (BN) is a probabilistic graphical model via a directed acyclic graph. It embeds the conditional probabilities of various events and can predict their probability based on the evidence. In the graphical representation of a BN, each event is depicted as a vertex, and each edge shows a causal relationship between two events (Cooper & Herskovits, 1992). The relative dependence between two events is modeled into a conditional probability. For instance, if A and B be two binary events, and the event B causes event A to happen in one-fourth of times, this is depicted as $P(A|B) = 0.25$. A BN enables Bayesian inference, which deduces the probability of an event

Table 2
Comparison of DS and BN reasoning (Verbert et al., 2017).

Feature	BN	DS
Fit for causal and diagnostic reasoning	+	–
Fit for information fusion	–	+
Fit for making decision	+	+
Inference coherence	+	–
Adaptable	+	+

based on its causes and its effects as shown formally in the following equations; where e, c, and s respectively represent *Event*, *Cause*, and *effect* (Symptom) (Cooper & Herskovits, 1992). Eq. (1) is the formal equation to calculate causal inference. Informally, based on the law of total probability, the probability of an event is the sum of the probability of its conjunction with each of its causes.

$$P(e) = \sum_{c_i \in causes(e)} \left(P(e|c_i) \times P(c_i) \right) \tag{1}$$

Eq. (2) is the formal equation to calculate the inverse reference. Informally, based on Bayes' rule, the probability of an event regarding the observed effects is the ratio of the probability of their conjunction on the event; the numerator can be expanded to the product of the inverse conditional probability and the probability of the effect.

$$P(e|s) = \frac{P(s|e) \times P(e)}{P(s)} \tag{2}$$

$$\mu_A(x) \wedge \mu_B(x) = min[\mu_A(x), \mu_B(x)], x \in X \tag{3}$$

$$\mu_A(x) \vee \mu_B(x) = max[\mu_A(x), \mu_B(x)], x \in X \tag{4}$$

$$bel(A) = \sum_{B|B \subseteq A} m(B) \tag{5}$$

$$pl(A) = \sum_{B|B \cap A \neq \emptyset} m(B) \tag{6}$$

2.3. Ontology

Borst (1997) has defined ontology as a "formal specification of a shared conceptualization", thus allowing the formal depiction of information and their relationships (Díaz Rodríguez et al., 2014). In the medical field, ontology is attracting growing interest for formalizing and reasoning medical data. For example, Disease Ontology (DO) is an open-source ontology for biomedical data associated with human disease. Its vocabulary consists of 8757 terms with unique maximal cross-references with other terminologies like National Cancer Institute Thesaurus and the National Drug File - Reference Terminology (NDF-RT) (Kibbe et al., 2015). Systematized Nomenclature of Medicine (SNOMED) is a well-known general terminology widely used as a medical ontology with over 120,000 terms. *International Health Terminology Standards Development Organization* have freely provided Systematized Nomenclature of Medicine – Clinical Terms (SNOMED–CT). It includes four core components: (1) Concept Codes: numerical codes identifying clinical conditions organized in hierarchies, (2) Descriptions: text describing the concept codes, (3) Relationships between the concept codes, and (4) Reference Sets: limits and ranges for classification (El-Sappagh et al., 2018). Fig. 1 shows a sample visualization of concept *fever* in SNOMED–CT.

SNOMED-CT is exploited by a growing number of medical applications, including clinical decision support systems, electronic health records, e-Prescription, and health research. For instance, the National Board of Health and Welfare of Sweden has implemented medical alert information using SNOMED-CT, which involves documentation

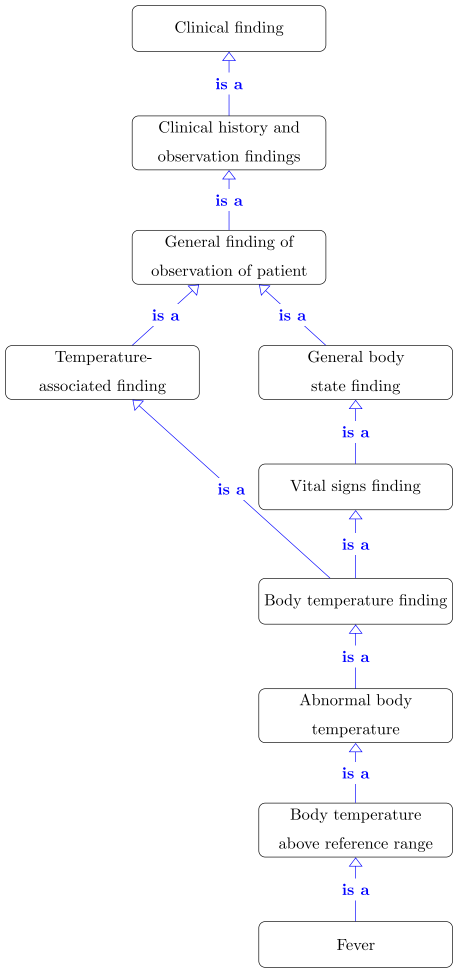


Fig. 1. Visualization of concept *Fever* in SNOMED-CT (Whetzel et al., 2011).

of patients' information regarding critical conditions, such as allergies and contagious disease (Socialstyrelsen, 2015). Moreover, *Snow Owl MQ* is a big-data platform that allows grouping the patients with similar characteristics, inspecting their health records for trends and correlation, and statically analyzing them for verification of clinical hypotheses (Ulrich, 2017).

2.4. Answer Set Programming

Answer Set Programming (ASP) is a form of declarative programming that focuses on severe search problems. ASP represents knowledge using logical phrases, and it derives new knowledge using automated reasoning. The concept of ASP is to represent a particular computational problem through a logic program and find the solutions, called answer sets, for that problem using automated reasoning performed by an ASP solver. ASP syntax is derived from Prolog, and its semantics are described using stable model semantics introduced by Michael Gelfond and Vladimir Lifschitz (Erdem et al., 2016). An ASP rule consists of two main parts: (i) Head and (ii) Body; if the body of an ASP rule is true, the ASP solver concludes that the head of that rule is also true. An ASP rule is formalized as follows:

$$Rule : Head \leftarrow Body.$$

$$l \leftarrow b_1, \dots, b_k, not\ b_{k+1}, \dots, not\ b_{k+n} \quad (k, n \geq 0). \tag{7}$$

where $b_1, \dots, b_k, not\ b_{k+1}, \dots, not\ b_{k+n}$ represents the rule's body and l represents its *head*. In ASP, a finite collection of ASP rules constructs an ASP program. Each rule in the ASP program can be seen as a limitation on answer sets of that ASP program. An answer set includes knowledge inferred using the reasoning on the ASP program. For instance, if an ASP program consists of the rule used in Eq. (7) and its answer set includes all of $b_1, \dots, b_k$ atoms and none of $b_{k+1}, \dots, b_{k+n}$ atoms, it should also include l. An answer set, also called stable model, is minimal and justified and composed of ground atoms, which are atoms with no variables. The formal definition of answer set is as follows (Lifschitz, 2010): Suppose the program Π consists of ASP rules. Grounding is performed on the program Π to replace the variables used in the program with all the constants appearing in the program. S is a set of ground atoms obtained using grounding. A Reduct Π^S, with no negated atoms, is obtained using two main steps: (i) for each atom $a \in S$, drop rules with *not a* in their body, (ii) drop literals *not a* from all other rules. The minimal model of the reduct (Π^S) is the answer set S.

Consider the illustrative example depicted in Fig. 2 with an ASP program Π with three facts and one rule, where a fact is a rule without body and with a single disjunct in the head. The predicate **condition**(`Episode`,`Pc`,`T`) represents the fact that there is a specific episode `Episode` with the probability `Pc` at the timestamp `T`. The predicate **suggestion**(`Episode`,`Action`,`Ps`) describes the fact that the suitable treatment for the specific episode `Episode` is the action `Action` with the probability `Ps`. The predicate **simpletreatment**(`Episode`, `Action`, `T`) depicts the fact that there is one possible treatment using the action `Action` at the timestamp `T` for the episode `Episode`. Where the `Episode` and timestamp are `"hypotension"` and `T`, respectively. Two actions as possible treatements are `"eating"` and `"takingpill"`. The inferred information, **simpletreatment**(`"hypotension"`, `"eating"`, `5`) and **simpletreatment**(`"hypotension"`, `"takingpill"`, `5`), are obtained using reasoning performed by answer set solvers.

The rich knowledge representation and efficient solvers are the main characteristics of ASP. Moreover, the non-monotonicity of ASP motivates us to use it for the treatment.

3. The proposed approach

3.1. General overview

As depicted in Fig. 3, Extended Hapicare operates in two phases for each medical condition: screening and monitoring. The medical condition has not been diagnosed during the former, and Extended Hapicare allows detecting potential medical conditions. In the case of any finding, Extended Hapicare will notify the doctor. With or without the notification of Extended Hapicare, the doctor can establish a diagnosis of the medical condition; and then prescribes monitoring for some specific episodes related to the diagnosed medical condition. Consequently, the phase of Extended Hapicare is changed from screening to monitoring for the diagnosed medical condition. Extended Hapicare provides telemonitoring and self-adaptive treatment; in this phase, the focus is to diagnose and react to episodes related to the diagnosed medical condition.

Since medical conditions are not exclusive and a patient might suffer from multiple medical conditions, Extended Hapicare can be in the monitoring phase for some medical conditions and in the screening phase for other medical conditions. For instance, in the case of a patient diagnosed with DM, Extended Hapicare is in the monitoring phase of DM and the screening phase of other medical conditions. In the former phase, it manages episodes related to DM, e.g., hypoglycemic episode, while in the latter phase, it diagnoses other medical conditions, e.g., chronic kidney diseases.

For both phases, Extended Hapicare applies IoT sensors, questionnaires, and manual inputs for capturing data, and ontology-based reasoning and probabilistic reasoning for enabling probabilistic diagnosis. The episodes are by definition acute and temporary; therefore, it is vital to diagnose them in real-time and react accordingly. On the other hand, the medical conditions last longer and usually are more complex to diagnose and react; therefore, in Extended Hapicare, in the case of diagnosing a medical condition, it notifies the doctor for further treatment.

```
%Facts
condition(" hypotension ", 50, 5).
suggestion(" hypotension ", " eating", 50).
suggestion(" hypotension ", " takingpill ", 40).

%Rule
simpletreatment(Episode, Action, T) :— condition(Episode,Pc,T),suggestion(Episode,
    Action,Ps).
```

Fig. 2. ASP example.

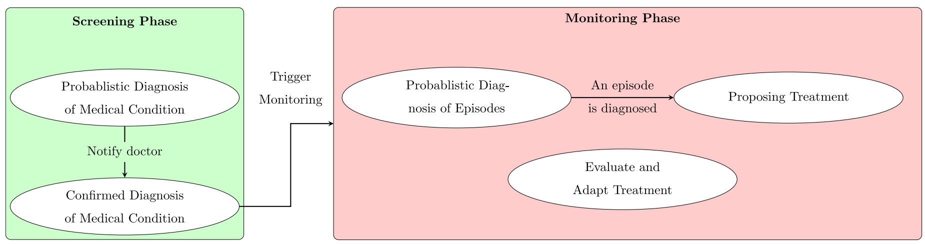


Fig. 3. General overview of Extended Hapicare.

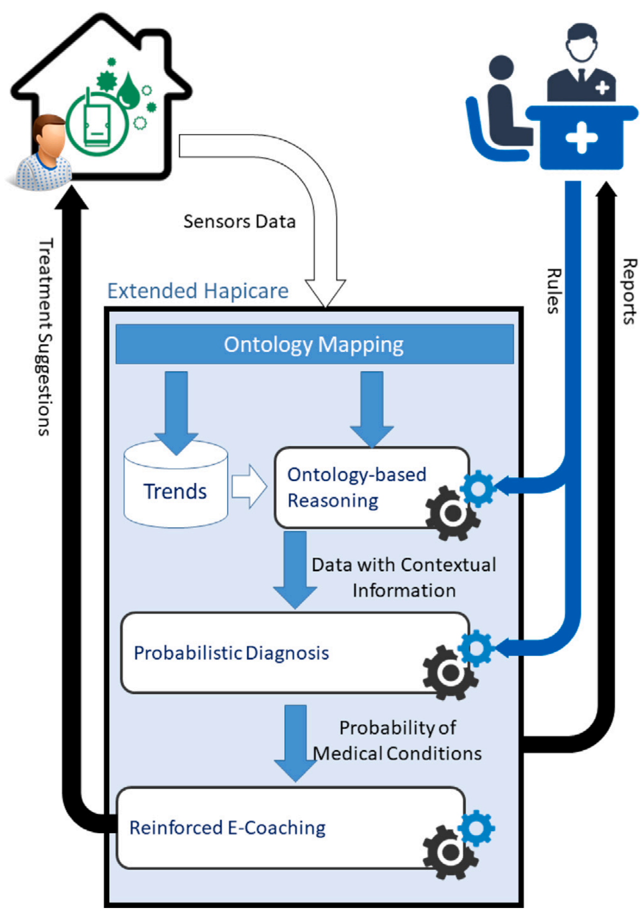


Fig. 4. Overall architecture of Extended Hapicare.

3.2. Architecture

Extended Hapicare is designed in three primary layers taking advantage of three technologies. The first layer is ontology-based reasoning, which provides contextual information; the second layer is Bayesian reasoning, which provides probabilistic diagnosis; and the third one is the ASP layer, which provides the self-adaptive treatment. Fig. 4 shows the overall architecture of Extended Hapicare.

3.3. Data

A broad range of data can be beneficial in diagnosing medical conditions/episodes; therefore, in Extended Hapicare, various types of data are considered. The data used in Extended Hapicare can be classified into the following four categories:

- Medical information: The medical information provided by the patient, including vital signs, e.g., body temperature.
- Medical file: The medical information provided by an expert; blood test results.
- Non-medical information: The information which is not considered health or medical information, but might be useful, e.g., room temperature.
- Deduced knowledge: The information that is not primarily fed to Extended Hapicare, but deduced from the analysis and reasoning on the data in Extended Hapicare.

Various sources might provide the aforementioned data. The source of data used in Extended Hapicare can be classified into four following categories:

- IoT Sensors: IoT sensors capture the information from the patient either continuously or on-demand, and then the captured information is transferred to extended Hapicare.
- Manual Input: For capturing the information without sensor, e.g., pain, the data are provided manually by the patient.
- External system: For the medical file, the information are stored in a health information system, which can be transferred to Extended Hapicare.
- Extended Hapicare: The deduced knowledge is produced internally in Extended Hapicare; hence, the latter is the source of information resulting from reasoning on input data.

Any data, regardless of its type and source, are transmitted to the ontology-based reasoning component to be mapped to an ontology and then processed.

3.4. Ontology-based reasoning

The data within Extended Hapicare are modeled in ontological terms for uniform depiction, i.e., the collected data are transformed into an ontology for further investigations. To this end, from the ontologies discussed in Section 2.3, we have opted for SNOMED-CT ontology due to its prevalence in research.

The data perceived from sensors are raw and sometimes meaningless on their own; hence, the ontology-based part processes the collected data and provides contextual information. In Extended Hapicare, trends and thresholds are applied for composing the context of patients. The threshold-based context depicts how the data compare with predefined thresholds, while the trend-based context shows how the collected data compare with the previous readings for the same user. Both types of context are necessary in order to model the health situation of a patient. For instance, a healthy weight is defined using a threshold. However, sudden weight gain or weight loss is a symptom of many medical conditions, which is modeled as a trend context.

In Extended Hapicare, we have implemented the ontology-based reasoning using JBoss Drools, which is a business rule management system with a rich feature set (Bali, 2009). The rules used in this reasoning are based on ontological definitions of vital signs, medical conditions, and episodes by experts, i.e., doctors.

3.5. Probabilistic diagnosis

Reliable treatment is not possible except with a reliable diagnosis and reasoning; in other words, telecare can only provide a suggestion for a successfully diagnosed condition. To this end, we have used probabilistic reasoning to diagnose medical conditions and episodes. It analyzes the collected data which have undergone ontology-based reasoning. The challenge for using these data is their unreliability, i.e., the collected data from lay patients in their homes are not reliable because they might mismeasure their vital signs or make a mistake in the manual input. Moreover, some manufacturers produce sensors for personal use only and not for high precision measurement. Albeit a patient might not notice a faulty sensor for some time. Hence, we cannot afford to throw away the unreliable data as they might include valuable information about patients. Another challenge is collecting all the required information about patients at their homes. It would require many sensors and even many questions that are not convenient for patients; hence, missing data should deprive a diagnosis.

Moreover, each medical condition/episode has a set of causes or risk factors that affect the probability of that medical condition/episode. It causes some changes in patients, including symptoms. Because BN is successfully applied for causal inference and prediction and estimation when the data are missing or unreliable (Acid et al., 2004); it is selected for probabilistic reasoning for both phases of screening and monitoring in Extended Hapicare. Such that all undiagnosed medical conditions and the episodes of diagnosed medical conditions are each modeled in an individual BN, where the causes are modeled as parents and the symptoms as descendants of a medical condition/episode, to maintain the existent causal relation of features to the medical rule. Each medical condition/episode is modeled separately in order to avoid the complication of models and undesired interrelations between medical rules.

3.5.1. Creation of Bayesian Network

The creation of BNs has two steps: (1) creation of its structure, i.e., the shape of the graph; and (2) creation of the Conditional Probability Table (CPT), i.e., the probabilistic relationships between the nodes of the graph. These two steps can be performed as knowledge-driven, data-driven, or hybrid. In BN's knowledge-driven creation, both steps should be carried out by experts; who model the known probabilistic relationship between events (symptoms, causes, and medical conditions/episodes) and then use them to create a Bayesian network of each medical condition/episode. For example, the last row of the medical rule presented in Table 1, is modeled as $P(cataracts|hypercortisolism) = 0.05$. The prevalence of *cataracts* enables a backward inference for diagnosing *hypercortisolism* based on this conditional probability. In data-driven training, the BN is trained using data. In the hybrid training, experts provide the structure of the BN. At the same time, CPTs are extracted using data records of different patients regarding each medical condition/episode. In Extended Hapicare, data-driven training is not used in order to assure the expected structure that symptoms and causes of a medical condition/episode are respectively represented as descendants and parents of that condition in the BN. Therefore, in Extended Hapicare, the structure of BN is created by experts with respect to the description given in Section 3.5.2. Since the medical conditions are numerous and usually complex to be modeled by experts, the CPTs of BNs used in the screening phase are created using datasets. Moreover, as the accuracy of the diagnosis of episodes is critical, the CPTs of BNs used in the monitoring phase are created by experts (doctors) according to their experiences on the intended patient or other similar patients. In other words, the creation of BNs used in the screening phase and monitoring phase are, respectively, performed hybrid and knowledge-driven.

3.5.2. Bayesian inference for medical diagnosis

As established in Section 2.2 diagnostic rules for medical conditions/episodes are a set of cause–effect rules. The modeling of medical conditions and episodes are performed similarly; for the sake of illustration, we depict the modeling of Acute Kidney Injury (AKI), which is one of the possible episodes of CKD (Hatakeyama et al., 2017). AKI can dramatically increase the chance of mortality and morbidity (Khadzhynov et al., 2019); however, (Yang et al., 2015) have discussed that AKI can remain undetected in the majority of cases (74%). The cause–effect rules can be classified into three main categories:

- Immediate causes: Those are the events that affect the probability of a diagnosis in the short term. For example, hypotension and infection episodes increase the chance of AKI (Khadzhynov et al., 2019). Hence these medical conditions are considered as immediate causes of AKI.
- Background causes: Those are the underlying events that affect the probability of a diagnosis in an extended period of time. For example, Friedman (2015) has discussed that comorbidities are significant risk factors for AKI; i.e., patients with DM and heart diseases are more susceptible to AKI (Khadzhynov et al., 2019). Hence these medical conditions are considered as background causes of AKI.
- Symptoms: Those are the effects of the medical condition/episode, i.e., the events whose probability is affected by the occurrence of a medical condition/episode. For instance, reduced body weight, irregular heart rate, and swelling are more plausible in AKI (Friedman, 2015)

Fig. 5 depicts simplified modeling of cause–effect relationships for diagnosis of AKI episodes, where the simple red arrow, doubled green arrow, and dotted blue arrow show the cause–effect relationships of immediate causes, background causes, and symptoms, respectively. Each cause–effect relationship can hold a probability, and each event can have a probabilistic value. Hence, Bayesian inference can deduce the probability of this episode with respect to any information on the other events. The second step in modeling an episode is the creation of the CPTs. For modeling medical conditions in BN, the creation of CPTs can be done using the datasets, i.e., training a BN with the given structure using the existing datasets. In the case of modeling episodes in BN, experts provide CPTs. However, the expert might benefit from the explicit probabilities provided in the medical studies. For instance, regarding the background causes of AKI, Khadzhynov et al. (2019) have reported that $P(heart\ failure \cap AKI) = 6.12\%$, $P(DM \cap AKI) = 8.81\%$, $P(heart\ failure) = 11.49\%$, and $P(DM) = 18.50\%$, which result in the following conditional probabilities:

$$P(AKI|heart\ failure) = 53.26\% \tag{8}$$

$$P(AKI|DM) = 47.62\% \tag{9}$$

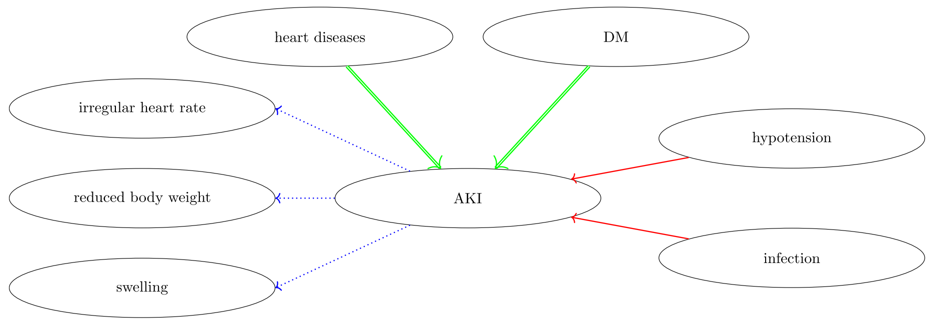


Fig. 5. BN modeling of cause–effect relationships for diagnosis of AKI.

In some cases, the existing medical rules do not include explicit probabilities; e.g., Lehman et al. (2010) have reported that for each hour of severe hypotension, the probability of AKI increases by 22%. Both explicit and implicit conditional probabilities can help the experts in creating CPT of the BN of the medical condition/episode.

3.6. Self-adaptive treatment

It is vital to help patients regarding their medical episodes to improve their quality of life. However, a common pitfall is uniforming the treatment services for all patients. Each patient has a different set of requirements, and even for one patient, the treatment might vary through time (Vitacca et al., 2018). Hence, a treatment service should be developed that facilitates manual modification by doctors and automatic customization for patients' needs. To this end, in Extended Hapicare, we have implemented a self-adaptive treatment service using commonsense reasoning implemented by Answer Set Programming (ASP).

As discussed in Section 2.4, ASP finds the answers to achieve the defined objective, considering the facts and rules. The predefined facts include different episodes and their possible treatments; the input facts are the episodes with their associated probabilities. The experts, e.g., doctors and caregivers, provide the predefined facts, while the probabilistic diagnosis component computes the input facts. The objective is a state with no episodes with a probability above a threshold, which is a safe state in the patients' lives. In this study, self-adaptive treatment is formalized in ASP.

For the automatic customization of treatment, the ASP-solver is called in any diagnosis update to adapt the treatment service for patients. Each solver call results in answer sets that extract the following information: (i) the updated episodes and the most recent action, (ii) the possible actions at this point, and (iii) the obtained awards based on the updated episodes and the most recent action. When the system suggests the next action, and the patient performs this action, the system will observe the new state, i.e., the system monitors the episodes after performing the selected action. The system knowledge is then updated using changing the award of the selected action for that episode.

In ASP, a transition system $D_m = \{S_m, A, f_m\}$ is represented where S_m represents a set of states and always is discrete, A represents a set of actions, and f_m represents transition function, $f_m : S_m \times A \rightarrow S_m$ and always is deterministic. This kind of representation allows fast decision-making for treatment. The set of states is represented by predicates representing episodes that may change their true values at different times, such as `condition("hypotension", 50, 5)`, where `5` is the time step of the predicate. Actions, A, are selected actions for the treatment based on the episodes and possible treatments. For instance, the predicate `selectedAct("takingpill", "hypotension", 5)` is used to represent that the action `"takingpill"` is selected as action at time step `5` for treating the episode `"hypotension"`. The main ASP rules used in the self-adaptive treatment are shown in Fig. 6. The award value for a specific action is changed according to the effect of that action on episodes. After D time step of the conduction of the selected action for the treatment, observation is done to obtain updated episodes in order to update knowledge. The changing amount of award value in each iteration, `step(S)`, is defined as a predefined fact, see line 1 in Fig. 6, provided by the doctor. Line 3 is used to obtain all possible actions regarding the diagnosed episodes and their possible treatments, represented by `suggestion(Episode, Action, Ps)`. Line 4 is used to illustrate that if the probability of the episode decreases with performing the specific action, the award value of that action should increase, i.e., the selected action has been suitable to mitigate the episode. Therefore, the ASP rules are updated after each iteration. On the contrary, line 5 is used to illustrate that if the probability of episode increases with performing that action, the award value of that action should decrease, i.e., the selected action has been ineffective in mitigating the episode. Line 6 and line 7 are used to select an action with the maximum award value for the treatment. The rules show that in Extended Hapicare, self-adaptive treatment is online with choosing a suitable action, observing the consequences of that action, and changing the award value of that action.

3.7. Workflow

3.7.1. Workflow in screening phase

A typical workflow of screening phase in Extended Hapicare is presented in Fig. 7. First, the data are collected from the patient, either directly from the IoT sensors or manually input by the patient or his/her caregiver. Later, the data are processed using ontology-based reasoning to yield contextual information. Afterward, all the data and their context are processed within the probabilistic diagnosis to estimate the probability of different medical conditions. For each medical condition, the *sensing action state* is defined as when the probability of that medical condition is between TH_{wary} and TH_{diag}, where the former is a predefined threshold for suspecting a medical condition, while the latter is a predefined threshold for diagnosing a medical condition. *Sensing action state* signifies that the medical condition is possible, but more evidence is required to confirm or reject this diagnosis. To this end, Extended Hapicare selects a cause/symptom related to that medical condition, where no recent information exists about that cause/symptom. Then Extended Hapicare requests the measurement of the selected cause/symptom. In order to minimize sensing actions, in Extended Hapicare, the most effective cause/symptom of that medical condition is selected, i.e., they are close to being considered as *pathognomonic* or *sine qua non* causes/symptoms. The former case is used to confirm the presence of the medical condition, i.e., if the response is positive, it is most likely that the medical condition is present; however, the latter case is exploited to confirm the absence of the medical condition, i.e., if the response is positive, it is most likely that the medical condition is absent. This cycle continues until either the probabilities fall below TH_{wary}, which shows it was a false doubt, or one surpasses TH_{diag}, which signifies detection of a possible medical condition. In the former case, the flow of continuous telemonitoring continues, while in the latter case, the patient's doctor is notified.

If the doctor establishes the diagnosis of a medical condition, which can be as a result of a notification from Extended Hapicare; the doctor

```
step(2).
timeStep(4).
possibleAction( Action , award(Action, Episode,Ps), T) :— condition(Episode,Pc,T),
    suggestion(Episode,Action,Ps).
possibleAction( Action , award(Action, Episode,NewValue), T+ D):—possibleAction(Action
    , award(Action, Episode, Value), T), condition(Episode, Pc, T), condition(
    Episode, PcNew , T+ D), PcNew<Pc, timeStep(D), NewValue =Value+S, step(S).
possibleAction( Action , award(Action, Episode,NewValue), T+ D):—possibleAction(Action
    , award(Action, Episode, Value), T), condition(Episode, Pc, T), condition(
    Episode, PcNew , T+ D), PcNew>Pc, timeStep(D), NewValue =Value—S, step(S).
1{max_sel_weight(X)}1 :— possibleAction(_, award(_,_, X), _), #max {V :
    possibleAction(Action, award(Action, Episode, V), T)} = X.
selectedAct( Action , Episode, T):—max_sel_weight(X), possibleAction( Action , award(
    Action, Episode, X), T).
#show selectedAct/3.
```

Fig. 6. Main ASP rules for self-adaptive treatment.

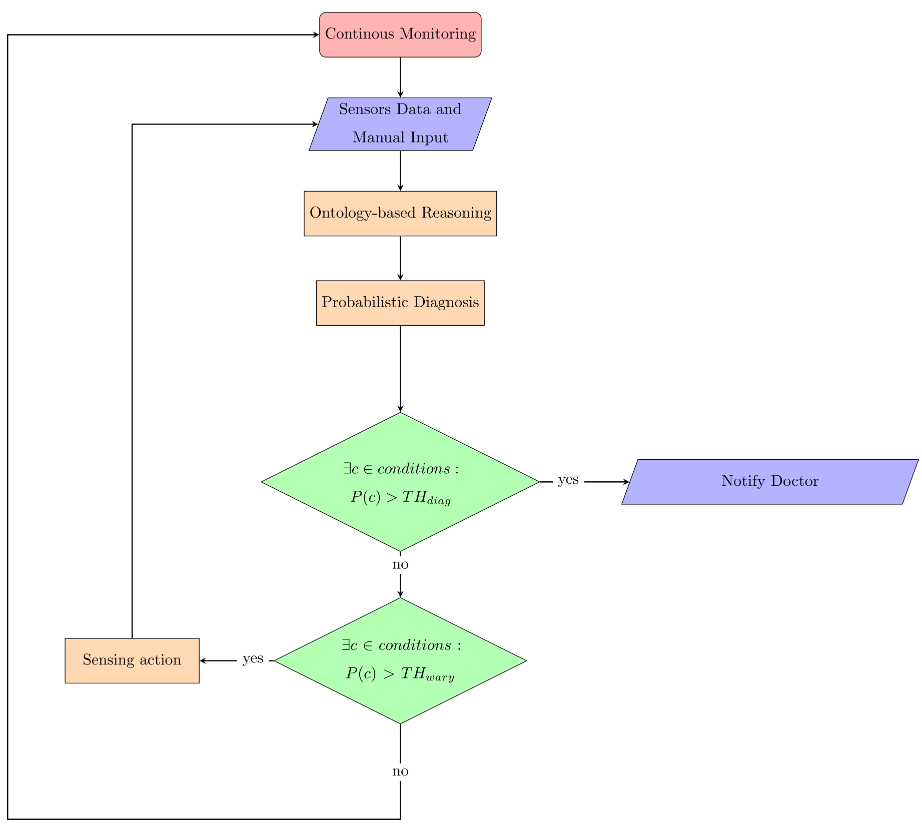


Fig. 7. Flowchart of screening phase in Extended Hapicare.

can prescribe telemonitoring for managing that medical condition. To this end, the doctor triggers the monitoring phase for the episodes related to the diagnosed medical condition.

3.7.2. *Workflow in monitoring phase*

A typical workflow of the monitoring phase in Extended Hapicare is presented in Fig. 8. The data gathering and ontology-based and probabilistic reasoning parts are the same as those in the screening phase. The chance of occurrence of various episodes is low, as they are temporary and acute; on the other hand, the occurrence of multiple medical conditions is not rare, but many medical conditions are more susceptible in the presence of other medical conditions. Hence, in the *sensing action state* of the monitoring phase, differential diagnosis is performed; that is, the selection of sensing action is affected by all the susceptible episodes with their probabilities between TH_{wary} and TH_{diag}. Similar to the screening phase, when no episode is susceptible, Extended Hapicare pursues continuous telemonitoring. However, if one or multiple episodes are detected, i.e., their probabilities surpass TH_{diag}; these diagnosed episodes alongside their probabilities are forwarded to the self-adaptive treatment component to process them using ASP and propose a customized treatment to the patient. Furthermore, during this treatment, the patient's state is closely monitored to modify the patient's profile based on changes in the probabilities of episodes. If the suggested treatment were effective, the patient would start feeling better, and the probability of the episode would decrease; hence, the probability of the ASP rule related to that suggested treatment should increase. When the recommended treatment is not beneficial for the patient, the ASP reduces the recommended treatment probability.

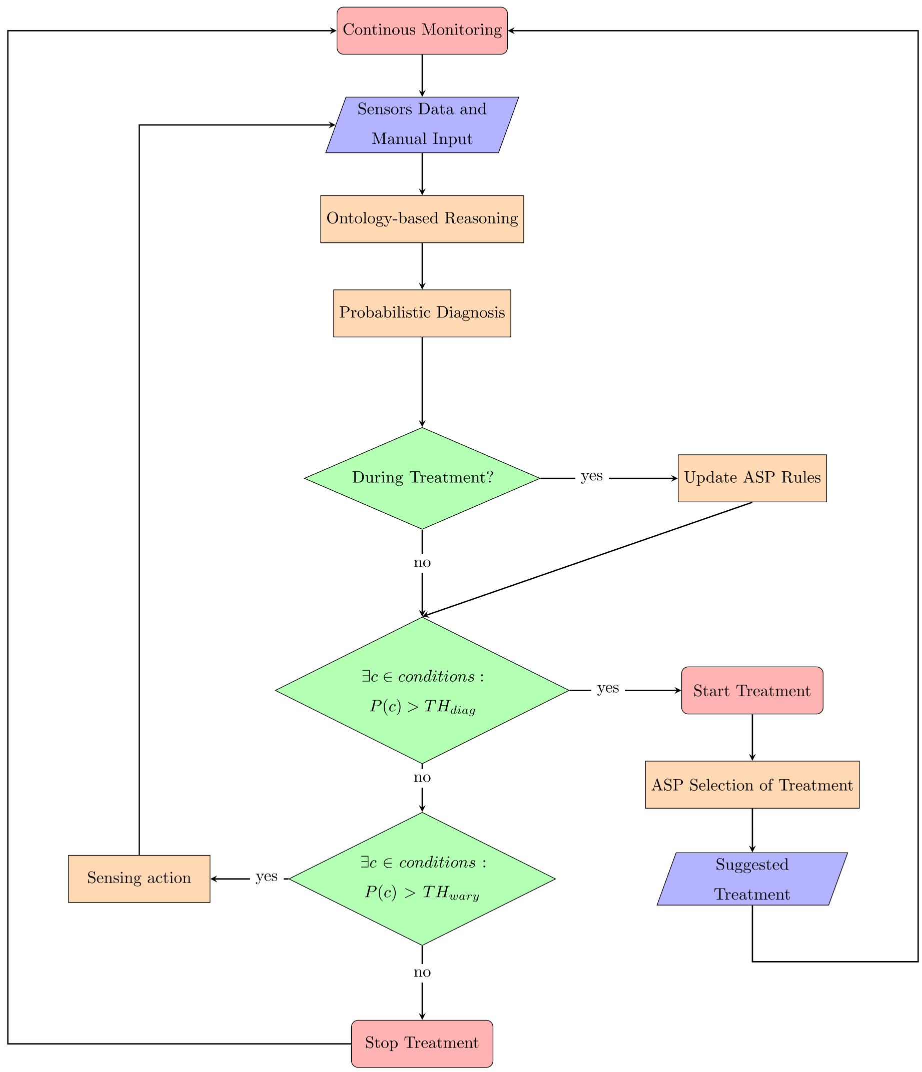


Fig. 8. Flowchart of telemonitoring phase in Extended Hapicare.

4. Evaluation

A comprehensive evaluation of Extended Hapicare requires access to monitor real-world patients to verify whether the proposed solution is fully adapted to their needs. The aforementioned environment was not accessible due to legal and ethical restrictions; however, Extended Hapicare is validated by clinicians and medical systems engineers collaborating with the Maidis company in the context of the ITEA3 Medolution project.

Probabilistic diagnosis is an integral component in screening and monitoring phases; in the screening phase, a reliable screening is only possible with a powerful diagnosis component; moreover, in the monitoring phase, as the self-adaptive treatment component relies on the result of the probabilistic diagnosis component, the performance of the latter can depict the expected performance of the whole system. Hence, a comparative study of the diagnosis component is provided to illustrate its strengths, specifically in the case of patients' inadequate information. Moreover, four scenarios are provided to validate Extended Hapicare.

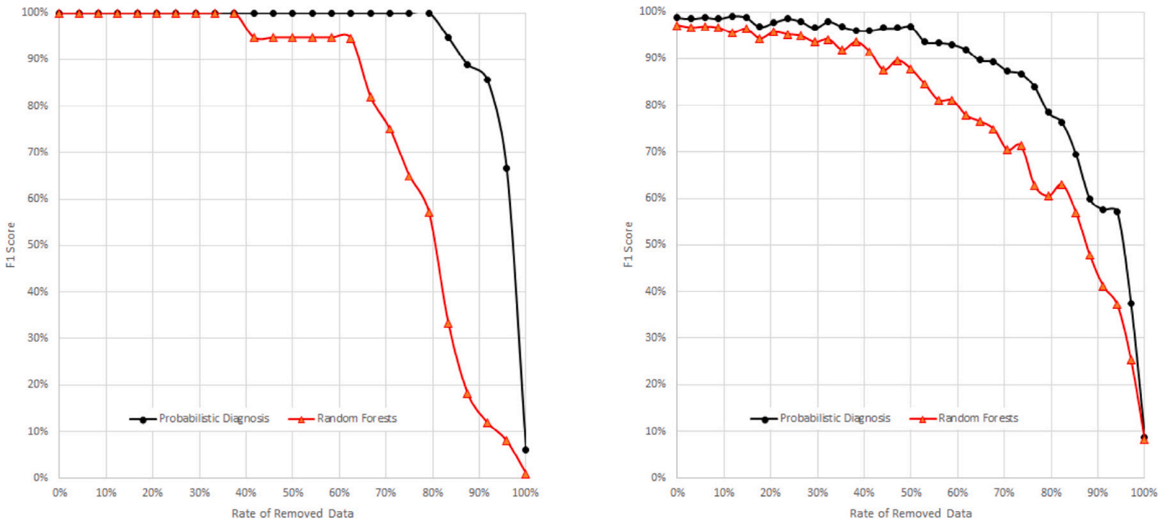
4.1. Comparative study

Since the use of probabilistic diagnosis is similar in both phases, in the rest of this section, without losing generality, we focus on the use of probabilistic diagnosis in the screening phase, i.e., for detecting plausible medical conditions. The performance of the probabilistic diagnosis component is evaluated based on a classical classifier method, namely random forest, to demonstrate how the loss of data affects their respective performance. First, the numeric features were processed using the ontology-based reasoning component of Extended Hapicare to produce a nominal context. Then both random forest and BN are trained using the dataset. We have created BN using hybrid creation for a valid comparison (see Section 3.5.1). The evaluation is made by comparing the performance of both methods to predict while increasingly removing random data features.

4.1.1. Dataset description

In order to demonstrate the performance of the probabilistic diagnosis in Extended Hapicare, we have implemented the evaluation using two datasets on medical conditions, namely, chronic kidney disease and dermatology datasets.

Chronic kidney disease Dataset: The first dataset is for the prediction of chronic kidney diseases; this dataset includes 11 numeric and 13 nominal features to predict chronic kidney disease, from which 4 are causes and 20 are symptoms. The causes in this model are *age*, *hypertension*, *DM*, and *coronary artery disease*; while the symptoms include the measurements of *blood pressure*, *hemoglobin*, *red blood cell count*, and *white blood cell count*. The dataset is collected from 400 patients in a hospital in India and labeled regarding the presence of



(a) CKD dataset **(b)** Dermatology dataset

Fig. 9. Comparison of F1 score between Extended Hapicare diagnosis and random forest.

chronic kidney disease. This dataset is available on the machine learning repository of the University of California, Irvine (Dua & Graff, 2017).

Dermatology Dataset: The second dataset focuses on the diagnosis of different types of erythemato-squamous illnesses in dermatological patients. They all share clinical characteristics of erythema and scaling, and hence it is challenging to distinguish them. Psoriasis, seborrheic dermatitis, lichen planus, rosea pityriasis, chronic dermatitis, and pityriasis rubra pilaris are the illnesses in this group. Unfortunately, the exact diagnosis requires biopsy in most cases. To this end, Güvenir et al. (1998) have collected this dataset to decrease the cost of prediction among these illnesses. The dataset has 33 linear and one nominal features and include 366 records, of which 2 are causes, and the rest are symptoms. This dataset is accessible on the machine learning repository of the University of California Irvine (Dua & Graff, 2017).

4.1.2. Comparison

Since the random forest model obtains a well-established prediction performance for the aforementioned datasets, it was selected as a rival for comparison. Fig. 9 depicts how removing the data from the records affects the performance of Extended Hapicare and random forest. They perform almost similarly when the data are complete. However, as the number of missing data increases, the drop of performance in the random forest model is enormous, which confirms the selection of BN for handling missing data in the core of reasoning in Extended Hapicare.

4.1.3. Robustness analysis

A BN is robust when the values of its target class slightly depend on the inputs of its causes (Surhone et al., 2011). In a robust BN, the changes in one of the causes hardly result in dramatic changes in the target class. Chan and Darwiche (2004) have implemented a tool integrating multiple approaches for quantification of robustness, also known as sensitivity analysis. We have used their tool to conduct sensitivity analysis based on the algorithm of Shenoy and Shafer (1986), on the target class, representing the medical condition. The results show that the three causes can raise the probability of the disease to 37.04% in the case of the *chronic kidney diseases dataset*. For the *dermatology dataset*, there are only two causes and they can raise the probability to 68.72% on average for the six illnesses of this dataset. This high sensitivity is due to the unbalanced values of the two causes as they are similar in 66.56% of data records.

Table 3
Summary of medical file of patient.

Item	Value
Name	Frank Smith
Gender	Male
Year of Birth	1960
BMI	36 kg/m^2 (10-Jan-2021)
Smoking	Yes
Chronic Diseases	DM Type II (5-Dec-2016)
Medical History	Heart Attack (7-Jun-2018)
Family Doctor	Dr. Anna Doe

4.2. Use case

In this use case, the patient is *Frank Smith*, a user of the Extended Hapicare application; the summary of his medical file is shown in Table 3. In Extended Hapicare, Frank Smith is under screening for various chronic diseases including high blood pressure, chronic kidney disease, and COVID-19. He is also under monitoring for diabetic and heart attack episodes. In this use case, we assume $TH_{wary} = 70\%$ and $TH_{diag} = 90\%$.

4.2.1. Scenario 1

Once Frank felt shortness of breath, he consulted the Extended Hapicare application. In the latter, the new information, *shortness of breath*, is processed as follows:

- The data are mapped into SNOMED-CT ontology; hence, Dyspnea (Concept Id: 267036007) is obtained and then added to the knowledge base.
- Ontology-based reasoning processes the ontology-mapped data; since no recent activity is present in the knowledge base, Dyspnea at rest (Concept Id: 161941007) is deduced using the ontology-based reasoning and then added to the knowledge base.
- Medical conditions/episodes that are related to the new finding are narrowed down to COVID-19 medical condition.
- The probability of COVID-19 is recalculated based on the knowledge base (including the new finding), which results in P(COVID-19) = 73%.
- Since the obtained probability is between TH_{wary} and TH_{diag}, the application selects Fever (Concept Id: 386661006) as missing information for sensing action.

- The sensing action is mapped to the following patient-friendly statement: "Please measure your body temperature".
- Frank uses an IoT sensor to measure his body temperature, which is 40°C.
- The obtained temperature is analyzed and then mapped into SNOMED-CT; hence, Fever (Concept Id: 386661006) is obtained and then added to the knowledge base.
- Medical conditions/episodes that are related to the new finding are narrowed down to COVID-19 and hypoglycemia medical conditions.
- The probabilities of COVID-19 and hypoglycemia are recalculated based on the knowledge base (including the new finding), which results in P(COVID-19) = 91% and P(Hypoglycemia) = 64%.
- Since the probability of COVID-19 exceeds TH_{diag}, the diagnosis is added to the knowledge base. Extended Hapicare notifies Frank's doctor about the possible presence of COVID-19 medical condition and encourages Frank to book an appointment with his doctor.

Frank visits his doctor for further diagnosis and possible treatments of his medical condition.

4.2.2. Scenario 2

In the second scenario, Dr. Anna Doe, Frank's family doctor, notices an alert from Extended Hapicare reminding a complete blood test for Frank as the previous one is old; hence, she orders a blood test for him. The results of his blood tests are processed in Extended Hapicare similar to the steps discussed in Section 4.2.1. Given the medical file and the new information from the blood test, using the model described in Section 4.1, the probability of CKD is calculated, and the obtained probability exceeds the diagnosis threshold; hence, Dr. Doe is notified. She examines Frank and confirms CKD diagnosis. Dr. Doe decides to start the monitoring phase for Frank regarding CKD. Frank is currently following peritoneal dialysis treatment at home that uses the lining of the abdomen to filter the blood inside his body; moreover, he is under continuous monitoring for the related episodes, e.g., Urinary Tract Infection (UTI) and Acute Kidney Injury (AKI), and further complications due to comorbidities. The doctor also prescribes him some antibiotics to take in the case of a UTI.

4.2.3. Scenario 3

Few months after diagnosis of CKD, Frank displays symptoms of *hypotension* and consequently Extended Hapicare diagnoses a hypotension episode. It follows the treatment discussed in Section 3.6. As mentioned in Section 3.5.2, episodes of hypotension affect the probability of AKI. Hence, after the diagnosis of hypotension, the probability of AKI is increased, but it is still under TH_{wary}; so no further action is performed.

4.2.4. Scenario 4

Frank feels sick and consults Extended Hapicare; the application selects the sensing action *measuring body temperature*. Ontology-based reasoning deduces *fever* and then adds it to the knowledge base. Given the knowledge base, probabilistic reasoning results in increasing the probabilities of COVID-19 and UTI, which are now both surpassing TH_{wary}. For differential diagnosis of the two possible choices, Extended Hapicare selects *cough* as a sensing action. Once Frank responds that he does not cough, the probability of COVID-19 is reduced. Since the probability of UTI is still over TH_{wary}, Extended Hapicare asks Frank for *a burning sensation while urinating* as a sensing action; his positive response increases the probability of UTI. For validation of this diagnosis, Extended Hapicare selects a pathognomonic sensing action and asks Frank to use test strips for UTI detection.[2] As Frank reports the color of the UTI test, Extended Hapicare confirms the presence of UTI, and based on the prescribed antibiotics in the treatment rules, the self-adaptive treatment module selects one of them and recommends Frank to start taking it. After a few days, Extended Hapicare asks Frank to take UTI test strips to see the changes in the state of infection. As Frank's health state regarding UTI is not improving, Extended Hapicare reduces the selected antibiotic's probability and recommends the patient take another antibiotic. After a few days, Extended Hapicare increases the probability of the second antibiotic as Frank recovers from UTI. Dr. Doe is notified at each step, and she can directly take over the medical treatment or change the treatment rules on the fly.

5. Discussion

In telecare, most sensors require an action from the patient, e.g., it needs to put the hand inside the cuff of a sphygmomanometer for measuring blood pressure. Moreover, for some of the critical information, there are no sensors; for example, no sensor can measure the pain's location. However, the more sensing actions, the more inconvenient for the user, and patients will eventually opt out if a system asks too many questions. On the other hand, holistic treatment is only possible with all the information. Hence, in Extended Hapicare, we have applied BN for diagnosis, which resulted in semi-holistic treatment with minimal data.

The evaluation results support the selection of the BN; they show the performance of diagnosis with minimum information; for both datasets, the results are hugely in favor of the BN. Albeit having more than half of the data, BN and random forest performed similarly.

On the other hand, one of the common pitfalls is depending too much on specific causes for diagnosis, which results in a biased prediction. In medical diagnosis, the causes only affect the general probability of a medical condition, and a diagnosis is only possible using the effects, also known as symptoms. The robustness analysis shows that the BN's training is efficient and learning about all the causes is not enough for the diagnosis of CKD in the first dataset. However, for the second dataset, there are only two causes, and they have similar values for most of the records; hence the BN training is not robust.

The scenarios presented in Section 4.2 demonstrate how Extended Hapicare can help patients and doctors with screening and monitoring of medical conditions/episodes. In the first two scenarios, the screening of a patient is shown. In the first scenario, the proposed framework captures the patient's information at his home for screening an urgent medical condition. The step-by-step interactions of Extended Hapicare illustrate the workflow of the system. In the second scenario, the medical file and external data are used for screening a medical condition. These two scenarios depict how Extended Hapicare can help a doctor in his/her diagnosis by providing insights into possible medical conditions.

In the last two scenarios, the monitoring phase is shown. In the third scenario, the treatment process and self-adaptive treatment are discussed. Moreover, we have demonstrated that a diagnosis of an episode might affect the probabilities of other medical conditions/episodes. In the last scenario, we have shown that Extended Hapicare can help patients diagnose and treat medical conditions/episodes even when the first treatment is not sufficient. This process might take several weeks with traditional methods of diagnosis and treatment. The result of the self-adaptive treatment can also help doctors treat other medical conditions/episodes; because the doctor can see which treatments were more effective.

6. Conclusion

The number of patients struggling with chronic diseases is growing; the traditional treatments are inefficient and massively costly. Efficient and holistic treatment is required to enhance their quality of life. However, comprehensive data collection is intrusive and expensive. In this paper, we have introduced Extended Hapicare as a compromise to achieve the semi-holistic diagnosis with the least possible intrusion.

[2] UTI test strips are diagnosis kits that change color in contact with urine and can be used at home. The presented color shows the presence and type of infection.

Additionally, we have used ASP to adapt treatment to the specific needs of each patient. Extended Hapicare collects data from IoT-based sensors as well as self-assessment; these data are processed through ontology-based reasoning for contextual information of collected data. The probabilistic diagnosis is responsible for diagnosing medical conditions/episodes. The diagnosis is based on the patients' contextual information of collected data. Hence, it creates a list of episodes and their associated probabilities and forwards it to the ASP component to suggest the most suitable treatment to each patient. Our experiments have shown the performance of probabilistic diagnosis in comparison with random forest model. Moreover, the validation of Extended Hapicare using four scenarios demonstrates its effectiveness in diagnosing and treating medical conditions and episodes of patients. However, a comprehensive experiment regarding the whole package of Extended Hapicare requires real patients with real-world rules and situations, which can be conducted as future works. Moreover, research on patients' security and privacy in this system is also planned for future studies.

CRediT authorship contribution statement

Hossain Kordestani: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Roghayeh Mojarad:** Software, Writing – review & editing, Visualization. **Abdelghani Chibani:** Writing – review & editing, Supervision. **Kamel Barkaoui:** Formal analysis, Writing – review & editing, Supervision, Project administration. **Yacine Amirat:** Supervision. **Wagdy Zahran:** Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We express our sincere appreciation to Dr. Paul VERDIEL, MD, for providing us medical information that enriched our work.

Funding

This study was supported by the French National Research and Technology Agency (ANRT) [CIFRE Reference number 2018/0284].

References

Acid, S., de Campos, L. M., Fernández-Luna, J. M., Rodrıguez, S., Marıa Rodrıguez, J., & Luis Salcedo, J. (2004). A comparison of learning algorithms for Bayesian networks: a case study based on data from an emergency medical service. *Artificial Intelligence in Medicine*, *30*(3), 215–232. http://dx.doi.org/10.1016/j.artmed.2003.11.002, URL: https://linkinghub.elsevier.com/retrieve/pii/S0933365703001325.

Afzal, M., Ali, S. I., Ali, R., Hussain, M., Ali, T., Khan, W. A., Amin, M. B., Kang, B. H., & Lee, S. (2018). Personalization of wellness recommendations using contextual interpretation. *Expert Systems with Applications*, *96*, 506–521. http://dx.doi.org/10.1016/j.eswa.2017.11.006, URL: https://linkinghub.elsevier.com/retrieve/pii/S0957417417307480.

Bali, M. (2009). *From technologies to solutions, Drools JBoss rules 5.0 developer's guide: Develop rules-based business logic using the Drools platform*. Birmingham: Packt Publ, OCLC: 845547425.

Bhatti, A., Siyal, A. A., Mehdi, A., Shah, H., Kumar, H., & Bohyo, M. A. (2018). Development of cost-effective tele-monitoring system for remote area patients. In *2018 international conference on engineering and emerging technologies* (pp. 1–7). Lahore, Pakistan: IEEE, http://dx.doi.org/10.1109/ICEET1.2018.8338646, URL: http://ieeexplore.ieee.org/document/8338646/.

Billinton, R., & Huang, D. (2008). Aleatory and epistemic uncertainty considerations in power system reliability evaluation. In *Proceedings of the 10th international conference on probablistic methods applied to power systems* (pp. 1–8). Puerto Rico: IEEE.

Borst, P. (1997). *Construction of engineering ontologies for knowledge sharing and reuse*. (Ph.D. thesis), Enschede: Centre for Telematics and Information Technology, OCLC: 68151508.

de Bruin, J. S., Schuh, C., Seeling, W., Luger, E., Gall, M., Hütterer, E., Kornek, G., Ludvik, B., Hoppichler, F., & Schindler, K. (2018). Assessing the feasibility of a mobile health-supported clinical decision support system for nutritional triage in oncology outpatients using Arden Syntax. *Artificial Intelligence in Medicine*, *92*, 34–42. http://dx.doi.org/10.1016/j.artmed.2015.10.001, URL: https://linkinghub.elsevier.com/retrieve/pii/S0933365715001396.

Bucholc, M., Ding, X., Wang, H., Glass, D. H., Wang, H., Prasad, G., Maguire, L. P., Bjourson, A. J., McClean, P. L., Todd, S., Finn, D. P., & Wong-Lin, K. (2019). A practical computerized decision support system for predicting the severity of Alzheimer's disease of an individual. *Expert Systems with Applications*, *130*, 157–171. http://dx.doi.org/10.1016/j.eswa.2019.04.022, URL: https://linkinghub.elsevier.com/retrieve/pii/S0957417419302520.

Bujnowska-Fedak, M., & Grata-Borkowska, U. (2015). Use of telemedicine-based care for the aging and elderly: promises and pitfalls. *Smart Homecare Technology and TeleHealth*, 91. http://dx.doi.org/10.2147/SHTT.S59498.

Chan, H., & Darwiche, A. (2004). Sensitivity analysis in Bayesian networks: From single to multiple parameters. In *UAI '04, Proceedings of the 20th conference on uncertainty in artificial intelligence* (pp. 67–75). Arlington, Virginia, USA: AUAI Press.

Cooper, G. F., & Herskovits, E. (1992). A Bayesian method for the induction of probabilistic networks from data. *Machine Learning*, *9*(4), 309–347. http://dx.doi.org/10.1007/BF00994110, URL: http://link.springer.com/10.1007/BF00994110.

Dempster, A. P. (2008). Upper and lower probabilities induced by a multivalued mapping. In *Classic works of the Dempster–Shafer theory of belief functions* (pp. 57–72). Springer.

Díaz Rodríguez, N., Cuéllar, M. P., Lilius, J., & Delgado Calvo-Flores, M. (2014). A fuzzy ontology for semantic modelling and recognition of human behaviour. *Knowledge-Based Systems*, *66*, 46–60. http://dx.doi.org/10.1016/j.knosys.2014.04.016, URL: https://linkinghub.elsevier.com/retrieve/pii/S0950705114001385.

Dua, D., & Graff, C. (2017). *UCI machine learning repository*. University of California, Irvine, School of Information and Computer Sciences, URL: http://archive.ics.uci.edu/ml.

El-Sappagh, S., Franda, F., Ali, F., & Kwak, K.-S. (2018). SNOMED CT standard ontology based on the ontology for general medical science. *BMC Medical Informatics and Decision Making*, *18*(1), 76. http://dx.doi.org/10.1186/s12911-018-0651-5, URL: https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-018-0651-5.

Erdem, E., Gelfond, M., & Leone, N. (2016). Applications of answer set programming. *AI Magazine*, *37*(3), 53. http://dx.doi.org/10.1609/aimag.v37i3.2678, URL: https://aaai.org/ojs/index.php/aimagazine/article/view/2678.

Friedman, T. (2015). Adrenal gland. In *Andreoli and Carpenter's cecil essentials of medicine* (9th ed.). (pp. 642–652). Saunders.

Glykas, M., & Chytas, P. (2004). Technological innovations in asthma patient monitoring and care. *Expert Systems with Applications*, *27*(1), 121–131. http://dx.doi.org/10.1016/j.eswa.2003.12.007, URL: https://linkinghub.elsevier.com/retrieve/pii/S0957417403002148.

Güvenir, H., Demiröz, G., & İlter, N. (1998). Learning differential diagnosis of erythemato-squamous diseases using voting feature intervals. *Artificial Intelligence in Medicine*, *13*(3), 147–165. http://dx.doi.org/10.1016/S0933-3657(98)00028-1, URL: https://linkinghub.elsevier.com/retrieve/pii/S0933365798000281.

Habib, C., Makhoul, A., Darazi, R., & Couturier, R. (2019). Health risk assessment and decision-making for patient monitoring and decision-support using Wireless Body Sensor Networks. *Information Fusion*, *47*, 10–22. http://dx.doi.org/10.1016/j.inffus.2018.06.008, URL: https://linkinghub.elsevier.com/retrieve/pii/S156625351730790X.

Harris, R. E. (2019). *Epidemiology of chronic disease: Global perspectives*. Jones & Bartlett Learning.

Hatakeyama, Y., Horino, T., Kataoka, H., Matsumoto, T., Ode, K., Shimamura, Y., Ogata, K., Inoue, K., Taniguchi, Y., Terada, Y., & Okuhara, Y. (2017). Incidence of acute kidney injury among patients with chronic kidney disease: a single-center retrospective database analysis. *Clinical and Experimental Nephrology*, *21*(1), 43–48. http://dx.doi.org/10.1007/s10157-016-1243-2, URL: https://doi.org/10.1007/s10157-016-1243-2.

Khadzhynov, D., Schmidt, D., Hardt, J., Rauch, G., Gocke, P., Eckardt, K.-U., & Schmidt-Ott, K. M. (2019). The incidence of acute kidney injury and associated hospital mortality. *Deutsches Aerzteblatt Online*, http://dx.doi.org/10.3238/arztebl.2019.0397, URL: https://www.aerzteblatt.de/10.3238/arztebl.2019.0397.

Kibbe, W. A., Arze, C., Felix, V., Mitraka, E., Bolton, E., Fu, G., Mungall, C. J., Binder, J. X., Malone, J., Vasant, D., Parkinson, H., & Schriml, L. M. (2015). Disease Ontology 2015 update: an expanded and updated database of human diseases for linking biomedical knowledge through disease data. *Nucleic Acids Research*, *43*(D1), D1071–D1078. http://dx.doi.org/10.1093/nar/gku1011, URL: http://academic.oup.com/nar/article/43/D1/D1071/2435381/Disease-Ontology-2015-update-an-expanded-and.

Kiureghian, A. D., & Ditlevsen, O. (2009). Aleatory or epistemic? Does it matter? *Structural Safety*, *31*(2), 105–112. http://dx.doi.org/10.1016/j.strusafe.2008.06.020, URL: https://linkinghub.elsevier.com/retrieve/pii/S0167473008000556.

Knowles, B., Smith-Renner, A., Poursabzi-Sangdeh, F., Lu, D., & Alabi, H. (2018). Uncertainty in current and future health wearables. *Communications of the ACM*, *61*(12), 62–67. http://dx.doi.org/10.1145/3199201, URL: http://dl.acm.org/citation.cfm?doid=3293542.3199201.

Kordestani, H., Mojarad, R., Chibani, A., Osmani, A., Amirat, Y., Barkaoui, K., & Zahran, W. (2019). Hapicare: A healthcare monitoring system with self-adaptive coaching using probabilistic reasoning. In *2019 IEEE/ACS 16th international conference on computer systems and applications* (pp. 1–8). Abu Dhabi, United Arab Emirates: IEEE, http://dx.doi.org/10.1109/AICCSA47632.2019.9035291, URL: https://ieeexplore.ieee.org/document/9035291/.

Lehman, L.-w., Saeed, M., Moody, G., & Mark, R. (2010). Hypotension as a risk factor for acute kidney injury in ICU patients. *Computing in Cardiology*, 11.

Lifschitz, V. (2010). Thirteen definitions of a stable model. In A. Blass, N. Dershowitz, & W. Reisig (Eds.), *Fields of logic and computation, Vol. 6300* (pp. 488–503). Berlin, Heidelberg: Springer Berlin Heidelberg, http://dx.doi.org/10.1007/978-3-642-15025-8{_}24.

Loreti, D., Chesani, F., Mello, P., Roffia, L., Antoniazzi, F., Cinotti, T. S., Paolini, G., Masotti, D., & Costanzo, A. (2019). Complex reactive event processing for assisted living: The habitat project case study. *Expert Systems with Applications*, *126*, 200–217. http://dx.doi.org/10.1016/j.eswa.2019.02.025, URL: https://linkinghub.elsevier.com/retrieve/pii/S0957417419301381.

Miller, B., & O'Toole, M. (2003). episodes. In *Miller-Keane encyclopedia and dictionary of medicine, nursing, and allied health, No. 7*. Saunders, URL: https://medical-dictionary.thefreedictionary.com/episodes.

Miller, B., & O'Toole, M. (2003). telecare. In *Miller-Keane encyclopedia and dictionary of medicine, nursing, and allied health, No. 7*. Saunders, URL: https://medical-dictionary.thefreedictionary.com/telecare.

OMG (2009). *Production rule representation (PRR)*. Object Management Group, Inc, URL: https://www.omg.org/spec/PRR/1.0/PDF.

Paget, J., Spreeuwenberg, P., Charu, V., Taylor, R. J., Iuliano, A. D., Bresee, J., Simonsen, L., & Viboud, C. (2019). Global mortality associated with seasonal influenza epidemics: New burden estimates and predictors from the GLaMOR Project. *Journal of Global Health*, *9*(2), Article 020421. http://dx.doi.org/10.7189/jogh.09.020421, URL: http://jogh.org/documents/issue201902/jogh-09-020421.pdf.

Parati, G., Dolan, E., McManus, R. J., & Omboni, S. (2018). Home blood pressure telemonitoring in the 21st century. *The Journal of Clinical Hypertension*, *20*(7), 1128–1132. http://dx.doi.org/10.1111/jch.13305, URL: http://doi.wiley.com/10.1111/jch.13305.

Pramanik, M. I., Lau, R. Y., Demirkan, H., & Azad, M. A. K. (2017). Smart health: Big data enabled health paradigm within smart cities. *Expert Systems with Applications*, *87*, 370–383. http://dx.doi.org/10.1016/j.eswa.2017.06.027, URL: https://linkinghub.elsevier.com/retrieve/pii/S095741741730444X.

Rahimi, S. M., & Wang, X. (2013). Location recommendation based on periodicity of human activities and location categories. In D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, G. Weikum, J. Pei, V. S. Tseng, L. Cao, H. Motoda, & G. Xu (Eds.), *Advances in knowledge discovery and data mining, Vol. 7819* (pp. 377–389). Berlin, Heidelberg: Springer Berlin Heidelberg.

Shafer, G. (1992). Dempster-Shafer theory. *Encyclopedia of Artificial Intelligence*, *1*, 330–331.

Shenoy, P. P., & Shafer, G. (1986). Propagating belief functions with local computations. *IEEE Expert*, *1*(3), 43–52. http://dx.doi.org/10.1109/MEX.1986.4306979, URL: https://ieeexplore.ieee.org/document/4306979/.

Socialstyrelsen (2015). SNOMED CT-grant of license of the Swedish National Release. URL: https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/dokument-webb/ovrigt/snomed-ct-swedish-national-release-affiliate-licence.pdf.

Sultan, M., Kuluski, K., McIsaac, W. J., Cafazzo, J. A., & Seto, E. (2018). Turning challenges into design principles: Telemonitoring systems for patients with multiple chronic conditions. *Health Informatics Journal*, Article 146045821774988. http://dx.doi.org/10.1177/1460458217749882, URL: http://journals.sagepub.com/doi/10.1177/1460458217749882.

Surhone, L., Tennoe, M., & Henssonow, S. (Eds.), (2011). *Robust Bayes analysis*. Betascript Publishing.

Ulrich, S. (2017). Snow Owl MQ. B2i Healthcare, URL: http://b2i.sg/snow-owl-mq/.

Verbert, K., Babuška, R., & De Schutter, B. (2017). Bayesian and Dempster–Shafer reasoning for knowledge-based fault diagnosis–A comparative study. *Engineering Applications of Artificial Intelligence*, *60*, 136–150. http://dx.doi.org/10.1016/j.engappai.2017.01.011, URL: https://linkinghub.elsevier.com/retrieve/pii/S0952197617300118.

Vitacca, M., Montini, A., & Comini, L. (2018). How will telemedicine change clinical practice in chronic obstructive pulmonary disease? *Therapeutic Advances in Respiratory Disease*, *12*, Article 175346581875477. http://dx.doi.org/10.1177/1753465818754778, URL: http://journals.sagepub.com/doi/10.1177/1753465818754778.

Whetzel, P. L., Noy, N. F., Shah, N. H., Alexander, P. R., Nyulas, C., Tudorache, T., & Musen, M. A. (2011). BioPortal: enhanced functionality via new web services from the National Center for Biomedical Ontology to access and use ontologies in software applications. *Nucleic Acids Research*, *39*(suppl_2), W541–W545.

WHO (Ed.), (2005). *Preventing chronic diseases: a vital investment*. Geneva, Ottawa: World Health Organization; Public Health Agency of Canada, OCLC: ocm62288123.

Wong, A., Amato, M. G., Seger, D. L., Rehr, C., Wright, A., Slight, S. P., Beeler, P. E., Orav, E. J., & Bates, D. W. (2018). Prospective evaluation of medication-related clinical decision support over-rides in the intensive care unit. *BMJ Quality & Safety*, *27*(9), 718–724. http://dx.doi.org/10.1136/bmjqs-2017-007531, URL: http://qualitysafety.bmj.com/lookup/doi/10.1136/bmjqs-2017-007531.

Xu, B., Xu, L., Cai, H., Jiang, L., Luo, Y., & Gu, Y. (2017). The design of an m-Health monitoring system based on a cloud computing platform. *Enterprise Information Systems*, *11*(1), 17–36. http://dx.doi.org/10.1080/17517575.2015.1053416, URL: https://www.tandfonline.com/doi/full/10.1080/17517575.2015.1053416.

Yanase, J., & Triantaphyllou, E. (2019). A systematic survey of computer-aided diagnosis in medicine: Past and present developments. *Expert Systems with Applications*, *138*, Article 112821. http://dx.doi.org/10.1016/j.eswa.2019.112821, URL: https://linkinghub.elsevier.com/retrieve/pii/S0957417419305238.

Yang, L., Xing, G., Wang, L., Wu, Y., Li, S., Xu, G., He, Q., Chen, J., Chen, M., Liu, X., Zhu, Z., Yang, L., Lian, X., Ding, F., Li, Y., Wang, H., Wang, J., Wang, R., Mei, C., ... Wang, H. (2015). Acute kidney injury in China: a cross-sectional survey. *The Lancet*, *386*(10002), 1465–1471. http://dx.doi.org/10.1016/S0140-6736(15)00344-X, URL: https://linkinghub.elsevier.com/retrieve/pii/S014067361500344X.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, *8*(3), 338–353. http://dx.doi.org/10.1016/S0019-9958(65)90241-X, URL: http://www.sciencedirect.com/science/article/pii/S001999586590241X.