What does it mean to provide decision support to a responsible and competent expert?

Antoine Richard, Brice Mayag, François Talbot, Alexis Tsoukiàs, Yves Meinard

Noname manuscript No.
(will be inserted by the editor)

What does it mean to provide decision support to a responsible and competent expert?

The case of diagnostic decision support systems

Antoine Richard · Brice Mayag · François Talbot · Alexis Tsoukias · Yves Meinard

Received: date / Accepted: date

Abstract Decision support consists in helping a decision-maker to improve his/her decisions. However, clients requesting decision support are often themselves experts and are often taken by third parties and/or the general public to be responsible for the decisions they make. This predicament raises complex challenges for decision analysts, who have to avoid infringing upon the expertise and responsibility of the decision-maker. The case of diagnosis decision support in healthcare contexts is particularly illustrative. To support clinicians in their work and minimize the risk of medical error, various decision support systems have been developed, as part of information systems that are now ubiquitous in healthcare contexts. To develop, in collaboration with the hospitals of Lyon, a diagnostic decision support system for day-to-day customary consultations, we propose in this paper a critical analysis of current approaches to diagnostic decision support, which mainly consist in providing them with guidelines or even full-fledged diagnosis recommendations. We highlight that the use of such decision support systems by physicians raises responsibility issues, but also that it is at odds with the needs and constraints of customary consultations. We argue that the historical choice to favor guidelines or recommendations to physicians implies a very specific vision of what it means to support physicians, and we argue that the flaws of this vision partially ex-

A. Richard (ORCID: 0000-0001-8677-8910)
Université Paris-Dauphine, PSL Research University, CNRS, LAMSADE, UMR 7243, 75016 Paris, France
GIE Hopsis Lyon, 3 Quai de Célestins, 69229 Lyon, France

B. Mayag
Université Paris-Dauphine, PSL Research University, CNRS, LAMSADE, UMR 7243, 75016 Paris, France

F. Talbot
DSII Bron, Hospices Civils de Lyon, 61 Boulevard Pinel, 69672 Bron, France

A. Tsoukias (ORCID: 0000-0001-5772-3988)
Université Paris-Dauphine, PSL Research University, CNRS, LAMSADE, UMR 7243, 75016 Paris, France

Y. Meinard (ORCID: 0000-0001-5928-8959)
Université Paris-Dauphine, PSL Research University, CNRS, LAMSADE, UMR 7243, 75016 Paris, France
E-mail: yves.meinard@dauphine.lamsade.fr

plain why current diagnostic decision support systems are not accepted by physicians in their application to customary situations. Based on this analysis, we propose that decision support to physicians for customary cases should be deployed in an "adjustive" approach, which consists in providing physicians with the data on patients they need, when they need them, during consultations. The rationale articulated in this article has a more general bearing than clinical decision support and bears lessons for decision support activities in other contexts where decision-makers are competent and responsible experts.

Keywords Decision Analysis · Decision Support Systems · Diagnostic Decision Support Systems

1 Introduction

Decision support is an activity that consists in helping a decision-maker to improve his/her decisions, through a better understanding of the stakes of the decisions, a more thoughtful examination of the relevant data, or/and a more rigorous utilization of relevant theories and practices. Decision support is usually provided upon demand, but clients requesting decision support are often themselves knowledgeable, at least to some extent, about the topic concerning which they ask decision support. Moreover, clients requesting decision support are often taken by third parties and/or the general public to be responsible for the decisions they make. In such cases, the task of the decision analyst (or decision support provider) is delicate in the sense that s/he risks infringing upon the expertise and responsibility of the decision-maker. The case of attempts at providing decision support to physicians in customary consultations is paradigmatic. Physicians are experts in medical matters and they are responsible for the medical decisions they make, but numerous decision support tools are developed in the literature and in practice in hospitals to provide them with decision support. How can one make sure that these tools do not infringe upon physicians' expertise and responsibility? In this article, we set out to answer this question, based on a literature review and a critical methodological analysis of medical decision support approaches.

The decision support systems that we are about to analyze here are part of the larger set of information systems in healthcare environments, more commonly called Health Information Systems (HISs). HISs have been developed in the last decades mainly to support and improve healthcare processes, decisions, and outcomes of patients. Nowadays HISs are ubiquitous in hospitals and it is difficult to find a hospital without an information system. One can distinguish, among HISs, different kinds of systems dedicated to healthcare support. According to Shortliffe and Cimino (2014)'s review of computer applications in healthcare, one of the first systems developed in healthcare environments corresponded to systems allowing the recording of healthcare information. These are Electronic Health Records (EHRs), including databases, indexing systems, and research systems using healthcare information. With a similar objective, Computer Physician Order Entry (CPOE) (Kuperman and Gibson, 2003), are systems developed to digitize physician's orders.

Another subset of HISs is composed of Clinical Decision Support Systems (CDSSs) (Musen et al., 2014; Berner, 2016). CDSSs include all kinds of tools designed to transmit information to clinicians to help them to make decisions or simply to facilitate their daily processes. The main objective of CDSSs is to minimize the risk of medical errors. CDSSs themselves include a variety of systems. Alert Systems provide alert messages to clinicians when an emergency occurs, e.g. when a hospitalized patient undergoes a heart attack. Alert Systems are also integrated into some CPOEs to prevent mistakes in drug prescriptions and/or drug dosages (Van Der Sijs et al., 2006). Reminder Systems (Garg et al., 2005) are likewise developed to avoid omission errors.

Lastly, Diagnostic Decision Support Systems (or DDSSs) are a subset of CDSSs dedicated to providing support to physicians in their clinical diagnosis. These systems will be our main topic in the present article. According to a recent systematic survey of DDSSs (Yanase and Triantaphyllou, 2019), there are currently two main types of DDSSs:

- DDSSs based on "gold standard" rules or guidelines defined by experts of the domain or health authorities (thereafter: "Guideline-based DDSSs").
  Clinical practice guidelines, including diagnostic guidelines, are lists of instructions to follow in a specific situation. They are generally based on current best practices and can be represented by a flowchart. Fig. 1 shows an example of a flowchart from the MIMS website[1] and based on the guidelines for diabetes treatments produced by the UK's National Institute for Health and Care Excellence (NICE)[2]. Other examples of clinical guidelines can be found on the NICE website[3], on the website of the French "Haute Autorité de Santé" (HAS) [4] or in reports of International Classification of Diseases (World Health Organization et al., 1992).
  Guideline-based DDSSs encompass "expert systems", which integrate "gold-standard" flowcharts/rules into their process to produce full-fledged diagnosis recommendations to physicians (Yanase and Triantaphyllou, 2019), but also systems that prescribe to physicians the steps they should follow to abide by the "gold-standard" (this is the case, for example, of the systems found on the NICE website or the Quick Medical Reference (QMR) linked to the INTERNIST expert system (Miller et al., 1986; Miller, 2010)).
- DDSSs based on Machine Learning (ML) algorithms, or ML-based DDSSs, are used to support diagnoses of specific diseases, with the aim to minimize error rates by treating large amounts of data on patients(Dua et al., 2014; Yanase and Triantaphyllou, 2019).
  ML algorithms are methods used to learn how to approximate a classification function based on a learning dataset. Classification functions could be, for example, functions anticipating the value of an exogenous variable y depending of the value of an endogenous variable x, or functions distinguishing pictures of healthy from pictures of diseased organs by analyzing a matrix of pixels. ML problems are generally divided into three subclasses, depending of the degree of knowledge included in the learning dataset: supervised learning (full knowledge), semi-supervised learning (some pieces of information are not available) and unsupervised learning (no predefined class).
  Many ML algorithms have been proposed to handle these classification problems, from Naive Bayes algorithms to Artificial Neural Networks and Support Vector Machine algorithms. In this paper, we used the term "ML-based DDSSs" to refer to all the DDSSs using one of these ML algorithms.

[1] www.mims.co
[2] www.nice.org.uk/guidance/ng28
[3] www.nice.org.uk
[4] www.has-sante.fr

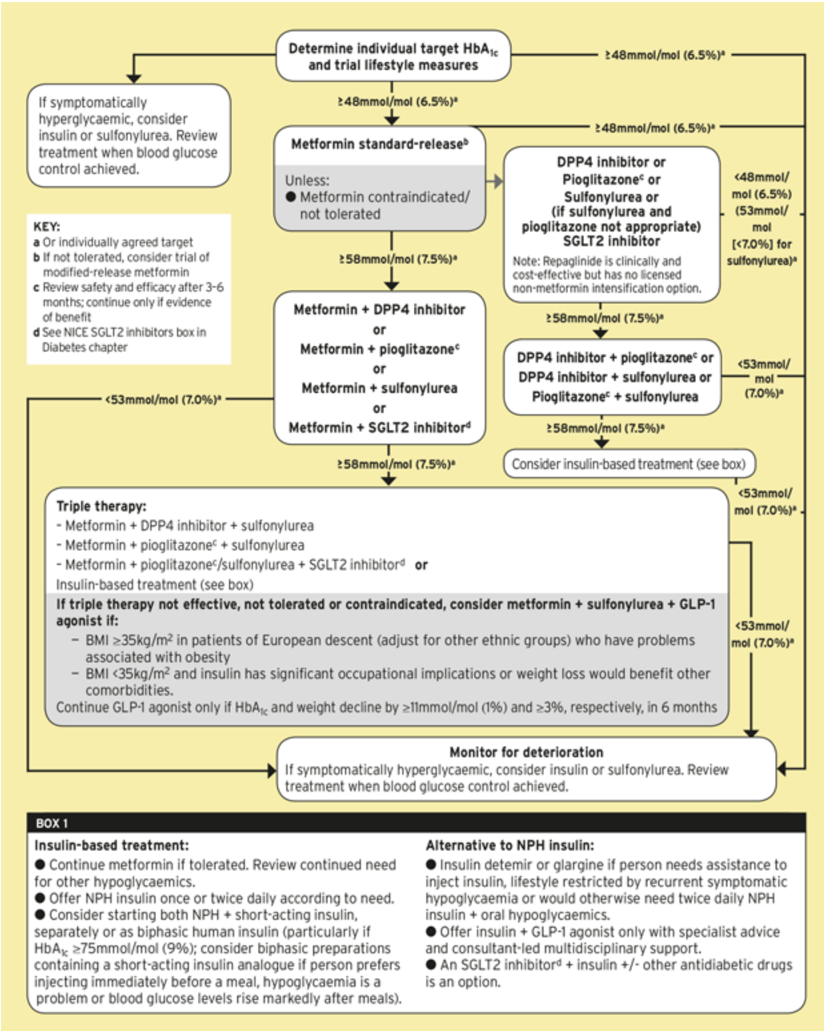

Fig. 1 Summary of NICE's guidance on treatment of type 2 diabetes proposed by the MIMS website[1]

As we will see in this paper, in their application to support customary diagnostic decisions, these DDSSs are currently in a paradoxical situation. On the one hand, their potential usefulness appears unquestionable, but on the other hand, they are generally poorly accepted by physicians. In addition, the use of DDSSs raises responsibility issues and involves patient safety risks. This paradoxical situation reflects, in our view, the more general difficulty to provide decision support to a competent, responsible decision-maker. By analyzing the specific case of DDSSs for customary consultations in detail, we aim to develop a new approach to address this general difficulty. To that end, we analyze here the reasons underlying the current failure of DDSS, and we draw the constructive lessons from this analysis.

By tackling this issue, this article aims to contribute to a broader research program devoted to analyzing the challenges facing decision support approaches and methodologies, as developed mainly in decision sciences and operational research, when they are applied to decisions involved in the design, implementation, and evaluation of public policies (Tsoukiàs et al., 2013; De Marchi et al., 2016). This research program has already produced applications to the evaluation of environmental policies (Jeanmougin et al., 2017), the design of policy options (Ferretti et al., 2019; Pluchinotta et al., 2018, 2019), the development of methodological tools for large scale environmental policies (Choulak et al., 2019), among others. In the wake of these contributions, we endorse the methodological and epistemological approach clarified in Tsoukiàs et al. (2013); Meinard and Tsoukiàs (2019); Meinard and Cailloux (2020)

Our reasoning unfolds in three steps. In section 2, we begin by reviewing historical choices that led to the current development policy of DDSSs and past experiences in the elaboration of DDSSs. Section 3 explores the adverse impact of HISs, CDSSs, and DDSSs, responsibility issues raised by the use of DDSSs, as well as gaps between DDSSs' design and the reality of customary consultations, to highlight potential reasons behind the failure of DDSSs in these situations. Section 4 discusses the conceptual approaches underlying the current DDSSs and sets out to determine which approach should be favored in the case of customary consultations. Section 5 briefly concludes the paper.

2 The paradoxical situation of Diagnostic Decision Support Systems

As introduced in section 1, HISs, such as EHRs and CPOEs, are now ubiquitous in hospitals. Due to this computerization of hospitals, works on CDSSs and DDSSs to support clinicians in their daily practices are on the rise. In this section, we develop a brief historical review of DDSSs and of the impact of the use of CDSSs in practice.

Our analysis is buttressed on a bibliographic review of the systems that have been developed to support physicians during consultations. In order to strengthen the purview of our analysis, we complemented this search by exploring the literature on support systems for clinicians in general. This analysis aims to capture the variety of systems that have been or can be used in practice to support physicians during customary consultations.

We made our research on PubMed with the following request: ("decision support system" or "computer-aided" or "artificial intelligence" or "machine learning" or "expert system") and "consultation". 393 articles were found with this request. This set of articles was to a large extent redundant for our purposes because it contained reviews and meta-analyses of DDSSs, their impact, and their acceptability, which synthesized the relevant information contained in other articles of this initial set. We, therefore, selected these reviews and meta-analyses. Because some fairly recent DDSSs might have been ignored in these reviews and meta-analyses, we also kept papers presenting specific DDSSs dedicated to physicians and published after 2017. We also keept papers including studies of the impact or acceptability of specific DDSS, because of the central role that these notions play in our study.

Applying these criteria to the content of titles and abstracts allowed filtering out 290 articles. Applying these criteria to the full content of the remaining papers then led to selecting 49 articles, including 12 (25%) reviews of systems used in general or in specific healthcare contexts, 27 (55%) papers presenting specific decision support systems (20 including clinical trials or feasibility studies), 10 (20%) studies of the impact of information systems on the performances of physicians, on patient safety or on the acceptability of systems.

2.1 A loss of confidence in physicians' diagnostic skills and know-how

According to Fieschi (1986)'s and Miller (1994)'s overviews of works on DDSSs from 1954 to 1993, early DDSSs were developed to try to reproduce, using computers, the behavior of physicians making a diagnosis. During these early stages of the history of DDSSs, from the 1950s to the late 1970s, most studies were devoted to representing physicians' behavior and possible uses of DDSSs in information systems. According to Miller (1994), the first studies devoted to developing information systems for diagnosis decision support purposes date back to the 1970s. In the early 1980s, the development of DDSSs in differents medical contexts, such as psychiatry (Morelli et al., 1987) or medical consultations (Kulikowski, 1988), was motivated by the development and proliferation of microcomputers, but also by innovations in user interfaces and networks systems. INTERNIST-1, developed by Miller et al. (1986), is an example of DDSSs developed during this period. These systems were generally designed to ask questions to physicians about the symptoms of patients in order to provide diagnostic suggestions to physicians. This "Greek Oracle" model of DDSSs, based on the idea that DDSSs are "magical tools" providing recommendations that physicians must follow, begun to be deprecated in the late 1980s (Miller and Masarie Jr, 1990). In the early 1990s, most studies on DDSSs had switched for explorations of AI methods such as neural networks or fuzzy logic systems, proposing new approaches for diagnosis decision support (Miller, 1994).

In 2000, the Institute of Medicine published the report *To Err Is Human: Building a Safer Health System*. This report, written by Donaldson et al. (2000), was a survey of multiple studies about medical errors, concluding that between 44000 and 98000 people die each year due to preventable medical errors. For comparison, Mokdad et al. (2004), who studied the causes of death in the U.S. in 2000, have reported an estimation of 43000 deaths due to motor vehicle crashes, 75000 deaths due to microbial agents and 29000 deaths due to incidents involving firearms. The *To Err Is Human* report pushed patient safety to the top of the agenda for governments and national healthcare policies.

According to Reider (2016), numerous national policies during the 2000s then set out to improve clinical practice guidelines, to improve education on patient safety, and to develop CDSSs. Other studies on medical and diagnosis errors, such as Leape (2000) and Berner and Graber (2008), bolstered governments in their efforts in this direction. Thereafter, many healthcare information systems were then developed to prevent potential medical errors. This is the case, in particular, of reminders, alert systems, and Guideline-based DDSSs. According to Miller (2016), at that time works on DDSSs also increasingly aimed at supporting physicians' diagnoses by giving them diagnostic recommendations, partly reinstating the deprecated "Greek Oracle" model of DDSSs.

Recent examples of Guideline-based DDSSs that were developed in this dynamic can be found in eIMCI (Bessat et al., 2019) and the ALMANACH project (Bernasconi et al., 2019), both dedicated to improving child health in primary care in developing countries by providing to physicians suggestions of diagnoses or actions according to guidelines of the Integrated Management of Childhood Illness (IMCI). The CHICA system (Anand et al., 2004) of Wishard Memorial Hospital in Indianapolis is another example dedicated to supporting child health in primary care, by generating forms based on patient's data and national guidelines. These forms are used to collect data on patients or to remind physicians of specific actions to do during consultations. Other applications of the CHICA system were developed, for the prevention of maternal depression (Carroll et al., 2013), prevention of suicidal behavior of adolescents (Etter et al., 2018), or prevention of obstructive sleep apnea (Honaker et al., 2018). López et al. (2017) presented a DDSS, called ophtalDSS, dedicated to supporting physicians in primary care to determine ocular diseases. OphtalDSS is based on decision trees and, once an ocular disease is confirmed, it provides adapted national guidelines. Kirby et al. (2018) proposed a DDSS based on guidelines of the American College of Cardiology/American Heart Association (ACC/AHA) to alert physicians when a patient meets criteria for severe aortic diseases and provide them with recommendations of the ACC/AHA. Similarly, Yang et al. (2018) proposed reminder systems, based on patients' allergy background, to prevent hypersensitivity reactions to radiocontrast media, according to the Korean health policy. Gonzalvo et al. (2017) proposed a DDSS based on the CONSORT guidelines, providing treatment recommendations for poly-medicated patients. Another well-known example is DxPlain (Barnett et al., 1987; Hoffer et al., 2005), an early DDSS dedicated to providing recommendations for primary care, which is still available[5].

A recent development in this history relates to the fact that, due to the generalized expansion of HISs, increasing volumes of data about patients are being recorded, flooding physicians under data (Pivovarov and Elhadad, 2015). This large amount of data quickly proved to be too difficult to analyze by human brains (Yanase and Triantaphyllou, 2019). Data mining and machine learning algorithms are better suited to this task, thanks to their distinctive efficiency when it comes to treating large amounts of data, unveiling correlations, approximating risk functions, or solving classification problems. According to Dua et al. (2014); Ozaydin et al. (2016); Miotto et al. (2017); Kulikowski (2019); Yanase and Triantaphyllou (2019), who surveyed ML-based DDSSs, many studies in the 2000s and the 2010s accordingly focussed on diagnoses assisted by machine learning algorithms.

Currently, ML-based DDSSs are being developed for numerous clinical situations. For example, Deig et al. (2019) surveyed ML-based DDSSs used in Radiation Oncology, mainly to assess the risk of bad reactions to treatments based on data on patients, allowing them to adapt treatments to improve outcomes. Peiffer-Smadja et al. (2019) reviewed ML-based DDSSs dedicated to supporting physicians for cases of infectious diseases by providing diagnoses/treatment recommendations, early detection of diseases, or predictions of responses to treatments. Gordon et al. (2018) surveyed the use of ML algorithms to support physicians in genetics, mainly in their analyzes of genetic risk, but also to recommend diagnoses to physicians. De Fauw et al. (2018); Zhang et al. (2018) recently proposed ML-based DDSSs to detect ocular diseases by analyzing retina images. Pearce et al. (2019) proposed a ML-based DDSS to evaluate the risk of emergency for a patient at the time of consultation. Titano et al. (2018) proposed a ML-based DDSS dedicated to anticipating neurological events by analyzing cranial radiographs. Numerous ML-based DDSSs are also dedicated to supporting the detection of tumors, such as breast tumors (Joo et al., 2004), brain tumors (Hollon et al., 2018), or skin tumors (Esteva et al., 2017). Elsner et al. (2018) and Pasquali

[5] `http://www.mghlcs.org/projects/dxplain`

et al. (2020) also reported the use of ML-based DDSSs in teledermatology to support physicians in the detection of skin tumors during teleconsultations. In a review of information systems, Kataria and Ravindran (2018) reported the use of ML-based DDSSs to anticipate responses to treatments or predict the propagation of diseases. In these examples, ML-based DDSSs appear to play the role of extensions of physicians, doing tasks that human physicians cannot perform with the same accuracy.

Based on this brief history of DDSSs, it appears that the *To Err Is Human* report, and the following works, have highlighted the limitations of physicians' diagnostic skills. In response, health authorities have financed the development of "gold-standard" guidelines and Guideline-based DDSSs dedicated to improving physicians' adherence to these guidelines. Early works on DDSSs, from the 1950s to the late 2000s, were mainly focused on these Guideline-based DDSSs. More recently, it appeared that machine-learning algorithms can be more performant than physicians for certain tasks (e.g., identify microscopic melanoma on images). This prompted the development of works on ML-based DDSSs in the last decades. Although works on Guideline-based DDSSs are still being developed, ML-based DDSSs started to dominate the field from the 2010s onwards[6]. In the subsections to come, we investigate whether these tools fulfill their promises by asking the following questions: Is the support provided by Guideline-based or ML-based DDSSs efficient in terms of patient safety? Are Guideline-based and ML-based DDSSs accepted by physicians and patients? Are current approaches of Guideline-based and ML-based DDSSs legitimately applicable to cases in which physicians can be considered to be "competent" and responsible for outcomes of patients?

2.2 Evidence that HISs, CDSSs, and DDSSs are potentially beneficial

In this sub-section, we start by reviewing studies of the impact of HISs on physicians' performances and patient safety, before zooming in on CDSSs and then on DDSSs. Patel et al. (2000) studied the impact of HISs, more specifically of the representation of knowledge in EHRs, not on clinicians' performances but on clinicians' reasoning and behaviors. They showed that a simple computer-based patient record system can have an important impact on physicians' behavior and working processes. In particular, they showed a standardization, through time, of physicians' working processes converging towards the EHR organization. Chaudhry et al. (2006) made a systematic review of the impacts of HISs on quality, efficiency, and costs of medical care, based on 257

[6] Some exceptions exist to the two predominant subsets of DDSSs (Guideline-based and ML-based DDSSs). Gräßer et al. (2017) proposed a DDSS dedicated to providing therapy recommendations, based not on expert guidelines or machine learning algorithms, but on similarity measures between the current case and previous ones, computed for each new cases, without any learning process involved. Whereas this system is akin to ML-based DDSSs, it does not use ML algorithms. Similarly, Giordanengo et al. (2019) proposed a DDSS dedicated to presenting self-collected data on patients and reminders of actions to do to physicians during the consultations of patients with diabetes. In this work, Giordanengo et al. (2019) didn't use the guidelines of any health authority but included physicians in the development process of the DDSS to establish rules to apply in specific situations. In addition, the recommendations established by consensus among the physicians involved are not intended for other physicians, but to developers adding needed features into the DDSS. Lastly, the ML-based DDSS proposed by Simon et al. (2019) does not use ML algorithms to make recommendations but to detect complex concepts in medical documents, facilitating access to information on patients or to reference documents. With this DDSS, Simon et al. (2019) showed that it is possible to use ML algorithms in other ways than by producing recommendations, while still providing support to physicians in practice.

studies. They concluded on the potentially beneficial impact of HISs on clinicians' performances. According to Leape and Berwick (2005) and Wachter (2004), who studied improvements in patient safety five years after *To Err is Human*, but also according to Clancy (2009), who proposed a similar analysis ten years after *To Err is Human*, the first impact of the Institute of Medicine report was the automation of medical error recording. With the introduction of HISs in hospitals, recording medical acts and results became more regulated. In addition, the development of reminders and alert systems helped to reduce potential mistakes. No doubt that such impacts of the ever-increasingly omnipresent HISs on physicians' work and on some aspects of patient outcomes, while not demonstrated before the 2000s, were to some extent perceived by physicians, medical authorities and the general public early on. In this context, the lost confidence epitomized by the *To Err Is Human* report provided a historical opportunity for CDSSs to entrench their usefulness.

Anticipating the call for diagnostic decision support of the *To Err Is Human* report, Johnston et al. (1994) have studied 28 controlled trials of different kinds of CDSSs (computer-assisted dosing, DDSSs, preventive care reminder, and computer-aided quality assurance, etc.) to assess the impact of CDSSs on clinicians' performances. Clinicians' good performances are, in this study, defined as low error rates in drug dosage and diagnosis, but also as the respect of guidelines by clinicians. Based on the few studies they found, Johnston et al. (1994) reported that some CDSSs (especially drug dosage recommendation systems) seem to have a beneficial impact on clinicians' performance. Hunt et al. (1998) similarly studied the effects of CDSSs on physician performances through a systematic review of 68 controlled trials, updated by Garg et al. (2005) with 97 controlled trials. They concluded that many CDSSs can improve clinicians' performances. Kaushal et al. (2003) studied the effects of CDSSs, and more specifically of CPOEs, on medication safety. They showed a potential reduction in the rate of medication errors, due to the use of CDSSs. Slain et al. (2014) analyzed retrospectively one year of use of a CDSS dedicated to supporting nurses in an emergency department. The CDSS was integrated into the workflow of the emergency department and proposed the pre-screening of patients at their arrival. The authors reported a higher triage accuracy and a better transfer of information thanks to the use of the CDSS. Zier et al. (2017) analyzed the use of a CDSS for one year in comparison with three years without CDSS. This CDSS was dedicated to supporting organ donation by early detection of brain death. The authors mentioned an improvement in early detection of brain death and organ donation. There is an exception: Verdoorn et al. (2018), who studied one year of use of a Guideline-based CDSS dedicated to preventing drug-related problems, reported lower performances with the CDSS than without. The authors pointed out the need for improvements of the CDSS.

In the more specific case of DDSS, although there are exceptions, such as Eccles et al. (2002) and Poels et al. (2008), who analyzed controlled trials of Guideline-based DDSSs with scenarios based on customary situations for physicians and reported that DDSSs have no significant impact (either negative or positive, on physicians' performances or workflow), a majority of studies shows a beneficial impact on physicians performances. Heckerling et al. (1991); Chang et al. (1996); Murphy et al. (1996), and Elstein et al. (1996), who made controlled trials on the Iliad expert system (Warner et al., 1988; Warner Jr, 1989), showed that expert systems can improve physicians' diagnosis accuracy in complex cases, in particular in the case of students (Murphy et al., 1996). Taylor et al. (2008) made controlled trials of a Guideline-based DDSS dedicated to supporting physicians in asthma cases. They showed that the DDSS helped physicians to improve their decision process and to decrease the duration of consultations. Watrous et al. (2008) made controlled trials to evaluate the impact of a Guideline-based DDSS dedicated to supporting the detection of heart murmurs during auscultation. They showed an improvement in the sensitivity and specificity of physicians using the DDSS in the classification of murmurs. Carroll et al. (2013) proposed a clinical trial of their Guideline-based DDSS dedicated to supporting the prevention of maternal depression by alerting physicians when a patient meets some criteria. According to the authors, their DDSS showed a potential beneficial impact on patient safety. Kostopoulou et al. (2017) made a controlled trial of a Guideline-based DDSS dedicated to supporting general practitioners by providing a list of potential diagnoses according to data on patients. They showed an improvement in diagnostic accuracy with the DDSS. The authors also mentioned that physicians entered more data on patients when they used the DDSS. Kirby et al. (2018) analyzed the use of a Guideline-based DDSS, dedicated to supporting the prevention of aortic diseases, during one year in 13 hospitals. They showed that their DDSS improved physicians' accuracy but also the clinical outcomes of patients.

Concerning the use of DDSSs in developing countries, Dalaba et al. (2014) studied one year of implementation of a Guideline-based DDSS for child health in healthcare centers in Ghana. The authors reported a decrease in complications and a diminution of deaths after the introduction of the DDSS. Bessat et al. (2019) made clinical trials of a DDSS dedicated to supporting child health in primary care facilities in Burkina Faso. The authors reported improvements in patient safety due to the DDSS. Similarly, Bernasconi et al. (2019) analyzed the impact of the introduction of Guideline-based DDSSs dedicated to child health in hospitals in developing countries. Clinical trials showed that DDSSs improved physicians' accuracy in primary care.

Concerning ML-based DDSSs, even though they are individually able to outperform physicians during sensitivity and specificity tests (Esteva et al., 2017), their impact when used in clinical practices remains understudied (Yanase and Triantaphyllou, 2019). For example, according to Peiffer-Smadja et al. (2019), who reviewed ML-based DDSSs dedicated to infectious diseases, among 60 ML-based DDSSs only three included clinical trials. The feasibility study by Jaroszewski et al. (2019) on a ML-based DDSS dedicated to mental illness prevention showed good results in mental crisis detection. Currently, because ML-based DDSS is still an emerging domain, it remains hazardous to determine if ML-based DDSSs can improve patient safety or beneficially modify physicians' workflow.

To summarize, clinical trials of DDSSs showed a theoretically beneficial impact on physicians' performances and on patient safety. Guideline-based DDSSs are quite performant when used in primary care or in developing countries, situations where "gold-standard" guidelines for specific cases are welcomed. When it comes to ML-based DDSSs, they are currently mainly evaluated on their specificity/sensibility or precision/recall performances (Yanase and Triantaphyllou, 2019). It remains delicate to determine whether current ML-based DDSSs provide physicians with helpful support in practice.

2.3 A questionable acceptability

According to Shortliffe and Cimino (2014), the most ubiquitous tools are Alert Systems, Schedulers, and Electronic Health Records (EHRs). Studying the introduction of an alert system dedicated to HIV prevention, Chadwick et al. (2017) showed that despite "alert fatigue", alert systems are generally accepted by clinicians. However, introducing information systems in clinical contexts remains a difficult task. According to Heeks et al. (1999), who surveyed the potential causes of successes or failures of HISs, even though some HIS succeed, many of them fail. Keen (1994) studied information systems in healthcare contexts and concluded that for every documented success, there are myriads of failures. Pare and Elam (1998), who worked on the introduction of information systems in clinical contexts, argued that many health care institutions have consumed large amounts of money and frustrated countless people in wasted efforts to implement information systems.

Heeks et al. (1999) and Heeks (2006) surveyed different cases of successful or failed HISs' introduction in hospitals. An illustrative example they explored is Beynon-Davies and Lloyd-Williams (1998)'s study of the failure of the introduction of a computer-aided despatch system for the London ambulance service. In this case, failure arose because "*the speed and depth of change were simply too aggressive for the circumstances*". The cancellation of this system caused an estimated waste of £20 million (ca. US$33 million). Another telling example was Guah (1998)'s analysis of the introduction of an expert system for computerized coloscopy in the coloscopy unit of a university hospital, in the UK. This system produced non-significant statistical information for physicians and needed to learn new work processes. The tool was therefore abandoned.

Sittig et al. (2006) studied the factors influencing the acceptability of DDSSs. They reported that a high percentage of CDSS's guidelines and/or recommendations were overridden, or ignored, by physicians. According to Overhage et al. (1997); Tierney et al. (2003) and Weingart et al. (2003), the percentage of DDSSs recommendations overridden by physicians varies between 54% and 91%. Sittig et al. (2006) also reported that physicians were more willing to accept clinical decision support for elderly patients with multiple medications or chronic conditions. Onega et al. (2010), studied the acceptability of DDSSs by radiologists, in comparison with a double reading by another radiologist. The authors surveyed 257 radiologists from different hospitals across the USA. According to their results, the radiologists were more favorable to double reading, even though most of them perceived that DDSSs were better at improving recall rates than double reading. The meta-analysis proposed by Masud et al. (2019), on the use of DDSSs in radiology departments, showed similar results on the low acceptability of DDSSs despite an improvement of performances perceived by radiologists.

Only a handful of studies showed a good acceptability of Guideline-based DDSSs, in very specific situations. This is the case of Porat et al. (2017), who analyzed the acceptability by patients and physicians of a Guideline-based DDSS. 34 general practitioners participated in the study by consulting 12 standardized patients during controlled trials. The authors reported that 74% of GPs found the DDSS useful, even though the use of the DDSS required them to enter more data on patients while interacting with them. Developing countries also constitute a specific case in which guideline-based DDSSs appear to be largely accepted by both physicians and patients, as illustrated by (Dalaba et al., 2014; Bessat et al., 2019; Bernasconi et al., 2019).

Concerning ML-based DDSSs, just like their impact on patient safety or physicians' performances, their acceptability in practice remains understudied (Peiffer-Smadja et al., 2019). Jaroszewski et al. (2019) reported that, during clinical trials of their ML-based DDSS for mental illness prevention, only 28% of participants answered "very likely" to the question presented by the DDSS: "Be honest, how likely are you to try the resources I just shared?". Nadarzynski et al. (2020) studied the acceptability of information systems dedicated to sexual health prevention. The authors reported that, for the first contact, 70% of patients preferred face-to-face consultations. Only 40% of patients found AI-chatbot acceptable.

To sum-up, although there are exceptions in specific situations, it appears that DDSSs are generally poorly accepted in customary situations, where support appears to be redundant with physicians' capabilities (Masud et al., 2019). It hence appears that we are currently in a paradoxical situation. DDSSs appear to be able to improve physicians' performances and patient safety. However, in practice, DDSSs remain poorly accepted in many situations and difficult to integrate into physicians' workflow. It appears also that the intrinsic capacities of a DDSS are not the sole factor determining its usefulness. There is hence a need to better understand why some DDSSs are not well accepted and which features are likely to improve the acceptability of a DDSS in practice.

3 Explaining the paradoxical failure

In section 2, we saw that Clinical Decision Support Systems (CDSSs) are potentially beneficial to minimize medical errors in some cases. However, we also saw that the introduction of a CDSS in a hospital is not without risks or failure and that current Diagnostic Decision Support Systems (DDSSs) are generally not accepted by clinicians, who often ignore DDSS recommendations in their daily practice.

Early explorations of barriers to the use of guidelines contain useful indications on reasons why some decision support tools can be rejected by physicians. Cabana et al. (1999) made an early meta-analysis of 76 studies on the non-acceptability of clinical practice guidelines and reported 7 potential barriers, classified into three categories:

1. External barriers such as the presence of contradictory guidelines, the inability to reconcile patient's preferences with guidelines recommendations, and other environmental factors such as the lack of time or resources.
2. Barriers that affect the attitude of a physician towards guidelines, such as the lack of agreement with specific guidelines or guidelines in general, the inertia of previous practices, the belief that s/he cannot perform guideline recommendations and the belief that performance of guideline recommendations will not lead to desired outcomes.
3. Barriers linked with how knowledgeable physicians are about guidelines, due for example to problems of accessibility of the guidelines, or to the volume of information to compute and then the time needed to stay informed.

In this section, we enlarge and update this analysis of potential reasons for non-acceptability, applying it more broadly to different aspects of HISs, CDSSs, and DDSSs, with a special focus on customary diagnostic.

3.1 Adverse impacts of HISs, CDSSs and DDSSs

Tsai et al. (2003) studied the impact of wrong diagnostic suggestions given by a DDSS on physicians' performance. They thereby questioned a commonly accepted postulate: if a DDSS does a mistake or a wrong proposal, the physician will detect it. This study was based on 83 simulations adapted from real clinical cases of cardiology. The subjects were 30 internal medicine residents in their second or third years of training and the DDSS was controlled to produce sometimes proposals that did not fit with *"gold standards"*. Tsai et al. (2003) reported that, when the DDSS produced good proposals, the accuracy of subjects increased. By contrast, the subject's accuracy dropped down when the DDSS proposal was incorrect. These authors also reported that subjects followed the DDSS's proposal more often when it was presented with a good confidence index. Povyakalo et al. (2013) developed a similar study on the impact of computer-aided detection of cancer on the performance of 50 radiologists. In this study, they evaluated the discriminating ability of radiologists on 180 mammograms with and without computer support. They reported that computer-aided detection helped less discriminating radiologists, but hindered the more discriminating radiologists by reducing their sensitivity. Bowman (2013), who worked on safety implications of electronic health record (EHR) systems, reported that poor design, improper use, and EHR-related errors, such as bugs or errors in the data, can lead to errors that endanger patients and decrease the quality of care. The risk of poor design and programming errors actually concerns all kinds of HIS, including CDSSs.

Bertillot (2016) studied HISs' attempts at rationalizing and standardizing clinicians daily practices, based on a set of interviews of clinicians (physicians, nurses, etc.) in several hospitals in France. Bertillot (2016) thereby showed that the introduction, in the last decades, of different HISs in hospitals improved the traceability of hospitalized patients and allowed for better transmission of information, but it also set the stage for the introduction of evaluation systems in these hospitals. These evaluation systems allowed comparing performances between hospital services, which led to the introduction of *"competitive managerial practices in public hospital"*. Bertillot (2016) also reported an additional administrative workload for clinicians, who had to enter information in the software. This time spent doing administrative work, though necessary for different reasons, is not a time devoted to patients. Mitchell et al. (2016)'s results highlight the same aspect of the impacts of HISs. They interviewed patient safety experts about their perceptions of works on patient safety incident reporting. This qualitative study highlights that clinicians, mainly due to a lack of time, perceived systematic reviews of patient safety incidents as an additional workload. Hall et al. (2016) reviewed 46 studies on wellbeing and patient safety to determine if there was an association between clinicians' wellbeing, burnout, and patient safety. They reported that clinicians' poor wellbeing was significantly correlated with higher risks of burnout, worse patient safety, and higher risks of medical errors. West et al. (2018) made a similar work on clinicians, burnout, their reasons, and their consequences. They reported the use of HISs as one of the factors leading to clinicians' burnout. One can hence see that, by trying to reduce the risk of medical errors, current HISs increase clinicians' workload. This additional workload reduces the wellbeing of clinicians and, by collateral effect, potentially increase the risk of medical errors in practice.

An associated risk was studied by Cabitza et al. (2017): the unintended consequences of Machine Learning in medicine. They reported that ML systems, due to their efficiency but also their opacity, could amplify the loss of clinicians' skills reported by Tsai et al. (2003) and Povyakalo et al. (2013). They also reported that the intrinsic uncertainty of healthcare contexts affects the performances of ML systems, reducing their accuracy. Similarly, Challen et al. (2019) studied the potential impact of artificial intelligence on clinical safety. They reported potential causes of errors due to AI tools in healthcare contexts. For example, ML systems are generally trained in a specific context and lose their accuracy when the context is changed. The opacity of some ML systems and the automation complacency were also reported as factors increasing the risk of medical errors. Authors also argued that reinforcement-based ML systems for decision support are potentially dangerous in the long run, by making unsafe exploration or reinforcing only short term behaviors.

If physicians are the only ones in charge of detecting potential errors of tools supposed to support them, it simply creates an additional workload and appears counterproductive. Not to mention the fact that, in the case of ML-based DDSS, physicians are supposed to be less "competent" than the DDSS to do the same tasks, and are therefore unlikely to be able to detect if the DDSS has made an error[7].

3.2 Responsibility issues

Itani et al. (2019), who studied the use of data mining algorithms for decision support, showed that social factors, such as patients' and physicians' values, are an important aspect to take into account to understand the acceptability or rejection of DDSS. These values refer to social perceptions and ethical implications of the use of DDSSs, but also to the social pressure on the responsibility of physicians with respect to the consequences of their decisions.

According to Goodman (2016), who surveyed the ethical and legal issues surrounding CDSSs, there is a need to define legal responsibilities in the use of CDSSs. Indeed, if one uses a DDSS and the DDSS is wrong, who is responsible? (De Dombal, 1987) The answer clearly depends on how the DDSS was developed or used.

For example, a technical error in programming could lead to an ill-advised recommendation. In such a case, one might argue that the true responsible is the programmer. But medical errors could also come from mistakes that a physician made when using the DDSS. The method on which the DDSS was based can also be a source of error. In the case of a Guideline-based DDSS using rules defined by experts, these "experts" might have provided rules that can be considered to be "dangerous" or "foolish" by the rest of the medical community.

One might argue that ML-based DDSSs are more trustworthy than Guideline-based DDSSs, due to the high performances of ML algorithms, outperforming physicians (Esteva et al., 2017), and using large amounts of data. However, responsibility issues are not different in the case of ML-based DDSSs: if the ML-based DDSS's recommendation was wrong and led to a medical error, who was responsible? ML-based DDSSs are trained and evaluated on datasets that might fail to encompass all the variety of possible use cases. Even if a trained ML-based DDSS had high sensitivity and specificity on a test dataset, these criteria of performances are not enough when we talk about patient safety in real situations. Moreover, supervised ML algorithms can only reproduce the behaviors they learned. Therefore, just like in the case of Guideline-based DDSSs, if the learning dataset was based on the behaviors of physicians whose behavior can be considered to be "dangerous" or "foolish" by the rest of the medical community, the trained ML algorithm will reproduce, and even amplify, this "dangerous" behavior (Garcia, 2016; Sandvig et al., 2016; Zou and Schiebinger, 2018). The main difference with Guideline-based DDSSs is that it is more difficult for physicians to detect if a ML-based DDSS had an unwanted behavior, especially if the process

[7] The emerging field of Explainable AI (Doran et al., 2017; Gunning, 2017; Rudin and Radin, 2019) holds promises to mitigate this problem.

of the ML-based DDSS is opaque to physicians. This is all the more worrying when physicians have high confidence in the ML-based DDSS's recommendations because the latter outperformed them (Tsai et al., 2003; Povyakalo et al., 2013). In such cases, responsibility problems are all the more worrying.

To summarize, both guideline-based and ML-based DDSSs create problems when physicians are considered to be responsible for patient outcomes. As introduced in subsection 3.1, physicians cannot be the only ones responsible for preventing potential medical errors due to the use of a DDSSs supposed to support them. According to the Asilomar AI Principles[8], developed during a workshop organized by the FutureOfLife Institute and dedicated to guiding institutions and designers to build beneficial Artificial Intelligence (AI), designers of AI systems and institutions must take up their share of responsibility in preventing errors or misuses of AI systems. These principles concern ML-based DDSSs, but also some Guideline-based DDSSs such as expert systems. The emergence of legislative instruments aimed at regulating the use of HISs, and more specifically the use of personal data, and to encourage the transparency of algorithms (e.g. GDPR in Europe (Voigt and Von dem Bussche, 2017)), witnesses the growing public awareness of such problems, pinpointing the fact that current decision support systems fall short of expectations.

3.3 A reality-design gap in customary situations

In addition to the adverse impacts of HISs and to responsibility issues, the literature suggests another reason potentially explaining the non-acceptability of some decision support tools by physicians: the so-called "reality-design gap problem".

This concept was introduced by Heeks et al. (1999) and Heeks (2006) in an attempt to explain why HISs succeed or fail. They argued that the bigger the gap between how a HIS was designed and the reality of daily practices, the higher the risk that the system will fail. To formalize this problem of design-reality gap, Heeks et al. (1999) proposed the ITPOSMO framework, formalizing seven dimensions that could create a gap: Information (Are physicians accustomed to using such kind of information?), Technology (Do the hospital have the technological capacities to run this system?), Processes (How does the system integrate itself into physicians' workflow?), Objectives and values (Do objectives of the system match with physicians' objectives and values?), Staffing and Skills (Does the system necessitate high technical skills to be used?), Management systems (Does the system necessitate additional structures to manage it?) and Other resources (Is it time-costing to use the system? Does the system create any additional workload?).

According to Heeks et al. (1999) and Heeks (2006), if the introduction of a HIS requires too many and/or too profound changes in clinicians' current daily practices, then the risk of non-acceptability is high. However, the goal of the introduction of a HIS is to improve clinical processes and/or healthcare outcomes, and accordingly to induce changes in clinical practices. If a HIS is too close to clinicians' daily practices, no improvement is possible. The difficulty in designing HISs is therefore to find a convenient equilibrium between minimizing the risk of non-acceptability of the HIS and maximizing the potential improvements of clinicians' practices.

[8] `https://futureoflife.org/ai-principles/`

To flesh out the meaning of this reasoning for our investigation, let us detail the content of the seven dimensions of ITPOSMO in the case that we focus on in this article: the one of customary consultations:

– Information: Guideline-based DDSSs generally provide actions/treatment recommendations based on "gold-standard" guidelines adapted to the situation. Physicians and clinicians are accustomed to the use of such "gold-standard" guidelines, part of their work being to be aware of the new "gold-standard" for the cases they treat regularly.
– Technology: Guideline-based DDSSs are generally integrated into already existing HISs and do not necessitate more technological resources than access to a database.
– Processes: Guideline-based DDSSs generally necessitate that physicians enter symptoms of the patient or other additional data asked by the DDSS. In customary situations, the process of Guideline-based DDSSs can be redundant with the physicians' process during a consultation.
– Objectives and values: the objective of Guideline-based DDSSs is generally to improve adherence to "gold-standard" guidelines. In customary situations, this objective is confronted with physicians' values, such as their free will, or the acceptability of "gold standard" guidelines (Cabana et al., 1999). Guideline-based DDSSs can also automatize too many things in physicians' workflow, leading potentially to a sensation of lack of control (Heeks, 2006)
– Staffing and Skills: Guideline-based generally do not necessitate additional skills to be used by physicians.
– Management systems: because "gold-standards" are evolving continuously, guideline-based DDSSs generally need to be managed regularly by an external agent to keep their recommendations up to date with the most recent "gold-standard" guidelines.
– Other resources: the use of Guideline-based DDSSs can be time-consuming for physicians, who can spend more time on the tools than interacting with the patient (Porat et al., 2017). In addition, physicians have to understand the reasons behind recommendations to prevent medical mistakes, creating an additional workload for physicians.

Concerning ML-based DDSSs, the task is a bit more difficult than for Guideline-based DDSSs, mainly because it is still an emerging domain, and ML-based DDSSs are still rarely used in practice (Peiffer-Smadja et al., 2019).

– Information: ML-based DDSSs generally provide recommendations or risk degrees. However, the reasons underlying a given recommendation can be unintelligible for physicians, depending on the ML algorithm used.
– Technology: Some learning algorithms, such as neural networks, can necessitate powerful technological resources. However, the classifier produced by a learning algorithm, such as a decision tree or a trained neural network, does not generally necessitate powerful resources to be used. In the case of online learning or continuous learning, powerful resources might be necessary (Kulikowski, 2019).
– Processes: current ML-based DDSSs are generally based on data already entered in the system and do not necessitate additional actions to be done by physicians. They can provide their support quickly in specific points. They can then easily be integrated into physicians' workflow as the display of an additional piece of information about a patient.

- Objectives and values: the main objective of current ML-based DDSSs is to provide highly performant tools to guide physicians in tasks they are not able to do alone with the same accuracy. In customary situations, for which physicians can be considered to be "competent", this objective may seem superfluous and can arouse their suspicion.
- Staffing and Skills: the use of current ML-based DDSSs might require additional training by physicians at least in terms of know-how to interpret the DDSS's results and to better understand how they work, their strengths and limitations.
- Management systems: It is possible to implement continuous learning by updating regularly the training dataset and rerunning the learning algorithm. This might also require to continuously test the performances of the DDSS, to prevent errors. However, none of these necessarily requires the intervention of an external agent, and everything can be automated.
- Other resources: the understanding of recommendations by physicians, when it is possible, may generate additional workload.

There certainly are exceptions to the general characteristics we explored above in our application of reality/design gaps analysis to Guideline-based DDSSs and ML-based DDSSs. Our goal was simply to highlight general trends in current ways to support physicians during their practices and see if they are applicable or not for customary consultations. Concerning Guideline-based DDSSs, for customary cases, physicians are often already aware of "gold-standard" to follow. Using Guideline-based DDSSs is generally time-consuming for physicians and redundant with their existing workflow. Concerning ML-based DDSSs, the main gap comes from the technology used. If physicians do not understand how the system works and how to interpret its results, the system will be seen as a "black-box", generating distrust. This will be reinforced if the objective of the ML-based DDSS is to outperform physicians in customary situations for which they feel competent.

Besides, for both Guideline-based and ML-based DDSSs, physicians are entrusted with the responsibility to make sure that the DDSS did not mislead her/him with ill-advised recommendations, creating an additional workload.

To sum-up our exploration so far, it appears that current tools used to support physicians are plagued by important drawbacks (adverse impacts, responsibility issues, and reality-design gaps), which are exacerbated in customary consultations. New approaches to support physicians in such situations are hence needed.

4 The way forward: the quest for "the right information"

According to Osheroff et al. (2012), the goal of CDSSs is to improve healthcare decisions and outcomes, including patient safety, by giving physicians the "right information". Osheroff's definition proved successful in the literature because it provides a synthetic formula that looks unquestionable. It also conveniently encompasses the immense diversity of CDSSs. But this successfulness of the formula also lies to a large extent in the indeterminacy of the phrase "the right information". In the case of current DDSSs, the "right information" is embodied by guidelines and/or diagnosis recommendations. In this section, we explore the idea that a crucial reason underlying the lack of acceptability of current DDSSs by physicians in customary consultations might be that this "right information" is not that right after all, and we set out to identify the truly "right information".

4.1 What is "information" in healthcare contexts?

At first glance, one might think that the notion of "information" in our context is unequivocal. A piece of information, one might think, is a raw data formatted to be readable by a physician. The interpretation of a piece of information by a physician gives her/him pieces of knowledge about a situation and allows her/him to make a decision. Collected by EHRs and CPOEs, hospital databases are rich in such raw data on patients, including: weights, ages, symptoms, reviews of hospitalizations, drugs took, allergies, etc. All these raw data can give clinicians a first layer of information.

With the same logic, the evolution of such data through time, their interconnection in patient care processes, gives a second layer of information. The notion of information hence appears more complex after all, since there are several layers of information.

A third layer of information can be found in guidelines summarizing "gold standards" to follow in a specific situation or for a specific operation. As mentioned in section 1, clinical practice guidelines are a list of instructions to follow in a specific situation. Guidelines include various formats such as pathways or algorithms to follow, and/or appropriateness criteria or parameters to check and instructions concerning how to interpret them (Field et al., 1990). But by admitting that "information" can refer to that third kind of entity, one admits that interpretation frameworks thanks to which data are interpreted, such as theories or sets of practices and know-how, are also pieces of "information" in a sense.

We see here that, in healthcare contexts, the term "information" can refer to a large diversity of entities, including raw data, interpreted data, and interpretation frameworks.

This analysis of the notion of "information" in healthcare contexts shows that the current approach, which consists in giving guidelines to physicians, is a particular kind of decision support approach, anchored in a very particular understanding of the notion of "information". This approach reflects a desire to standardize diagnosis processes, based on the presupposed idea that such a standardization will lead to minimizing medical errors. However, as mentioned by Woolf (1993), who studied the impacts of guidelines on patient care, such standardization could harm patients and interfere with the individualization of care. In clinical contexts, physicians' adaptability can, in many cases, be more important than conformism.

This suggests that, instead of clinging to the standard reductive view of "information" aimed at standardizing diagnosis processes, one should strive to identify the kind of information that physicians need when they proceed to make a diagnosis.

4.2 Identifying the constraints binding the decision support process to determine what is the "right information"

We claim that, in order to identify what counts as the "right information" in customary diagnosis decision support, we need to analyze, at a methodological and epistemological level, the true meaning and significance of the activity that consists in providing decision support in this context. Our analysis so far has highlighted the numerous specificities associated with the context of customary diagnosis. As in most medical contexts, this specific context raises responsibility issues, but another marked specificity is that, in this context, physicians are competent, and are not easily outcompeted by sophisticated tools.

These specificities reflect constraints that bind the interaction between decision support providers, developing decision support tools, and physicians, which are decision-makers benefiting from decision support. Decision support providers concerned with providing relevant and *acceptable* decision support have no choice but to take these constraints into account to choose the kind of approach to unfold in their interactions with physicians.

Meinard and Tsoukiàs (2019) showed the pivotal role of an analysis of the constraints binding decision support processes, which is key to choose a relevant decision support approach, which plays, in turn, a decisive role to entrench the validity and the legitimacy of the decision support provided. This framework sheds useful light on our analysis of the various drawbacks plaguing various current DDSSs, developed above.

As explained in section 1, Guideline-based DDSSs include systems providing recommendations based on "gold-standard" guidelines, but also systems providing directly these "gold-standard" guidelines. This approach is relevant when the interaction between the decision-maker and the decision support provider is constrained by a requirement to homogenize decision processes and make them converge towards "gold-standards" that are collectively recognized, by expert institutions and the general public. Such decision support interactions correspond to what Meinard and Tsoukiàs (2019) call situations in which an "irrevocable governance pattern" binds the decision support process. Still according to Meinard and Tsoukiàs (2019), in such situations, decision support providers should endorse a "conformist" approach, striving to identify the tools that will be most acceptable to the members of the governance pattern. In situations in which physicians' skills are deficient due to problems in physician training, as observed in some cases in developing countries, and in which health institutions play a key role in a powerful governance scheme aimed at reinforcing the quality of medical treatment, Guideline-based DDSSs are relevant. In such cases, "the right information" that decision support should provide to physicians really is encapsulated in guidelines.

ML-based DDSSs include all recommendation systems based on supervised ML algorithms. These tools assume that there exists a function linking data on patients to a specific class (e.g., disease, set of treatments, risk degree, etc.), independently of the beliefs and knowledge of the decision-maker, or her/his context of decisions. They also assume that this function can be approximated by machine learning, and more specifically by deep learning because multi-layered neural networks are known to be universal approximators (Hornik et al., 1989). The goal is then to find the best approximation of this function, generally evaluated by its sensitivity and specificity. An approach based on such tools is relevant in situations in which the objectivity and truthfulness of the underlying theories and algorithms can be taken for granted and considered unquestionable. In such situations, in which a given theoretical framework is considered to be unquestionable, Meinard and Tsoukiàs (2019) argue that the relevant approach to decision support is "objectivist". In healthcare contexts, situations bound by this constraint are those in which the data to collect, the efficiency of existing tools to collect them, and their capacity to outperform all other forms of expertise, are clearly established. This is the case, for example, for the detection of ocular diseases (Zhang et al., 2018) or the evaluation of the risk of infection (Peiffer-Smadja et al., 2019) or of treatment reaction (Deig et al., 2019). These are tasks for which we can suppose that a classification function exists, but we cannot suppose that any decision-maker is "competent" enough to approximate it closely.

The decision support context that we are mainly interested in here, customary consultations, is not characterized by patterns similar to those presented above, for which Guideline-based DDSSs and ML-based DDSSs appear relevant, respectively. In customary consultations, physicians are competent: they do not need to be monitored by authorities verifying their compliance with gold-standards, and they do not need tools to replace them. In such contexts, the main constraint is that the conditions should be met for physicians to be able to exercise their responsibilities. This echoes situations that Meinard and Tsoukiàs (2019) refer to by talking about a constraint to respect a "sanctified spirit of initiative" of the decision-maker. This phrase is arguably rather vague, but the case of physicians performing customary diagnosis might be case in point to clarify it. The two distinctive features of customary diagnosis, competence and responsibility, point to the need for decision support providers to leave the decision-maker to make her/his own choices and to take responsibility for them. The role of the decision support provider in such cases is to make all efforts to facilitate, smoothen, and speed-up the processes favored by the decision-maker, and to adjust to her/his needs. This approach to decision support is called "adjustive" by Meinard and Tsoukiàs (2019).

4.2.1 An example of application

In order to flesh out in concrete terms what such an "adjustive" approach consists in, we introduce here a practical example, referring to Richard et al. (2018), a work developed in collaboration with the public hospitals of Lyon to propose a DDSS dedicated to supporting customary consultations. Clinicians of the hospitals of Lyon have a software, called Easily®, at their disposal. This software allows clinicians to access different kinds of HISs. During consultations, physicians have access to the hospital's EHR and to CPOEs, but not to any DDSSs for now.

Richard et al. (2018) proposed an analysis focussed on interactions between physicians and patients, but also between physicians and HISs, during customary consultations of endocrinologists, to identify what kind of tools should provide relevant support in such situations. To do so, Richard et al. (2018) made practical observations and analyzed event logs of customary consultations (more than 12.000 event logs, divided into 2.700 traces).

Based on these analyses, Richard et al. (2018) built a synthetic model of the decision process of physicians during a medical consultation (Fig. 2) This model shows interactions between possible actions of physicians during medical consultations. Richard et al. (2018) highlighted that key points of the decision process, such as the choice of a prescription or the choice to put end to the consultations, are highly dependent on the accumulation of data on patients. Accordingly, the authors concluded that the search for raw data on patients, and then the choice of the raw data to look at, constitute a central key point of consultations.

However, due to the huge quantities of data accumulated on patients in the last decades, physicians are nowadays flooded by medical data (Pivovarov and Elhadad, 2015). Even though most of the data on patients needed by physicians during their medical consultations are available, these data are not always easily accessible.

The model proposed by (Richard et al., 2018) suggests that physicians, during medical consultations, spend more time searching for data about the patient than analyzing them to reach a diagnostic. This analysis shows that, contrary to what most DDSSs currently available assume, what physicians need during customary consultations is not recommendations of diagnoses or diagnostic guidelines. As concluded by Richard

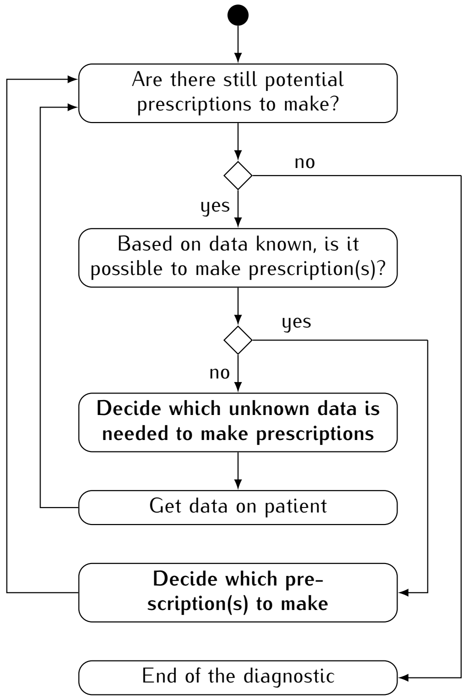

Fig. 2 Graphical representation of the model of physicians' diagnostic decision process proposed by Richard et al. (2018)

et al. (2018), what physicians need are tools that can anticipate, retrieve, and summarize data needed by physicians about patients. A relevant tool is accordingly one that would speed up the search for data. This idea echoes Sittig et al. (2006), who argued that guidelines and/or diagnostic recommendations are useless but for complex cases. It also appears all the more relevant in the light of studies on the summarization of electronic health records (EHRs) such as Pivovarov and Elhadad (2015), showing that there is an increasing need for EHRs summarizers.

Nevertheless, physicians need different data depending on their medical specialty or the pathology of the patient. This can be learned by questioning physicians and by creating a set of rules, but expert systems are generally difficult to build and to maintain through time (Shortliffe, 2012; Miller, 2010). In addition, questioning physicians would be against our aim to prevent any increase in their workload. Richard et al. (2018) therefore set out to learn what data are needed by physicians by analyzing their searches and their entry in the hospital's database, so as to anticipate their needs and provide them with a subset of data about their current patient at the beginning of medical consultation. By doing so, the searching phase of medical consultations should be minimized by handing over to the information system the task for which it is more efficient than human beings: searching data in a large database. In this approach, the aim of decision support is to ensure that physicians have all the data they need on their patients, and the interpretation of these data is then left to physicians.

4.3 Promises and limits

As mentioned before, the conclusions reached by Richard et al. (2018) were based only on observations and analyses of consultations in endocrinology at the HCL. However, the approach proposed in this case-study, illustrating the more largely applicable reasoning developed above, holds promises in light of the above analysis of the reasons underlying the non-acceptability of current DDSSs.

A first strength of the proposed approach is that it draws on the competence and the cumulated experience of the physician. We have seen that current approaches used to support physicians suppose that physicians are not competent enough. Whereas such approaches can be relevant in complex situations, for customary consultations they are inappropriate and they can arouse distrust towards the DDSS among physicians or a feeling of being put aside by the DDSS. With an adjustive approach, physicians keep the leadership of the decision process. Moreover, their competences in drawing diagnosis and interacting with patients are highlighted.

A second important strength of the proposed approach is that it does not increase the workload of physicians, and rather decreases it. We have seen above that the increase of clinicians' workload, due to the introduction of decision support systems, has been reported as a barrier to their acceptability. Current approaches tend to tell physicians how they should work, without taking into account their current decision processes or the impact of the format of decision support. With our approach, the first aim is to understand physicians' current decision processes, to establish on which point of their workflow physicians need support in priority and what kind of decision support is more relevant to provide.

A third strength is that, as compared with numerous other approaches, our approach does not involve a risk to decrease physicians' performances or capacities. According to Povyakalo et al. (2013), the use of current DDSSs tends to decrease the performances of physicians with good diagnosis skills. In addition, Tsai et al. (2003) have reported that wrong recommendations of current DDSSs are less detected by inexperienced physicians. This loss of diagnosis skills is often cited as an important barrier to diagnosis decision support. Focussing on providing data on patients prevents this problem by refusing to prescribe what to do during the diagnosis decision process. The interpretation of data on patients is left to the physicians. The impact of our approach on physicians' diagnosis skills is therefore minimized.

Lastly, and arguably most importantly, a major strength of our approach is that it does not infringe upon the responsibility of the physician. Indeed, the responsibility issues raised by the use of guidelines, described in section 3.2, are no longer a problem if we focus on providing data on patients. As mentioned above, our approach does not prescribe what to do during the diagnosis decision process, it only focusses on providing to physicians with what they need during their decision processes.

5 Conclusion

In this paper, we have developed a reflection on the current approaches to supporting customary diagnostic decisions, which consist mainly of giving guidelines and/or diagnosis recommendations. We have explored the historical reasons that led to the choice of this approach and we have highlighted its drawbacks. In particular, we have stressed the fact that DDSSs tend to put physicians at the background on their own decisions, raise various responsibility issues, and are generally not accepted by physicians.

We have then argued that giving guidelines or recommendations reflects a strong choice on how to support decisions, which ignores the current decision-maker process or the impact of recommendations on this process. In the case of customary medical consultations, the "sanctified spirit of the initiative" of physicians is currently a binding constraint. Current DDSSs are not relevant in such situations. We have argued that DDSSs dedicated to supporting customary consultations must endorse an adjustive approach, which consists in ensuring that physicians have all the data they need about patients to reach a diagnosis. The interpretation and final decisions are then left to the decision-maker and her/his expertise, avoiding responsibility issues raised by Guideline-based and ML-based DDSSs in such situations.

Decision support systems developed in an adjustive approach can be seen as "personal assistants" that provide support during all the decision process and adjust themselves by interacting with decision-makers. However, just like "conformist" and "objectivist" approaches, adjustive approaches are not adapted to all situations. In cases in which conformist and objectivist approaches are relevant, guideline-based and ML-based DDSS undoubtedly have a role to play, and one should certainly not replace them by adjustive approaches. Analyzing the features of decision processes, the constraints binding interactions between decision-makers and decision support providers, and other aspects of the context, is always needed to choose the most relevant approach. Identifying these points during the development of new DDSSs could help designers to have a better understanding of the kind of support needed and to propose more adapted systems to physicians. Works that include physicians or clinicians in the development offer interesting promises in this respect (Giordanengo et al., 2019; Horrocks et al., 2018).

The reasoning developed in this article is focussed on diagnostic decision support for customary medical consultations, tasks for which physicians are considered to be competent and responsible. However, it bears lessons for other contexts in which decision support has to be provided to competent, responsible decision-makers. For example, in the context of the implementation of environmental policies, Meinard and Thébaud (2019) argued that environmental management schemes are currently crippled in France by the lack of a large-scale database on vegetation types, while environmental institutions spend considerable time and money to produce ill-adapted guidelines unusable by experts in the field. Decision support in this area could largely benefit from an analysis developed along the lines of our analysis of DDSSs.

Acknowledgements This work was made in collaboration with employees of the hospitals of Lyon. Thanks to all of them. Special thanks to Pr. Moulin and Dr. Riou for their suggestions and instructive discussions. Special thanks also to Juliette Rouchier, Olivier Cailloux, and Philippe Grill for their advices and comments on earlier versions of this manuscript. We also thank two anonymous reviewers of the journal for their powerful and exacting comments and criticisms.

References

Anand V, Biondich PG, Liu GC, Rosenman MB, Downs SM, et al. (2004) Child health improvement through computer automation: the chica system. In: Medinfo, pp 187–191, DOI 10.3233/978-1-60750-949-3-187

Barnett GO, Cimino JJ, Hupp JA, Hoffer EP (1987) Dxplain: an evolving diagnostic decision-support system. Jama 258(1):67–74, DOI 10.1001/jama.1987.03400010071030

Bernasconi A, Crabbé F, Adedeji AM, Bello A, Schmitz T, Landi M, Rossi R (2019) Results from one-year use of an electronic clinical decision support system in a post-conflict context: An implementation research. PloS one 14(12), DOI 10.1371/journal.pone.0225634

Berner ES (2016) Clinical decision support systems, 3rd edn. Springer

Berner ES, Graber ML (2008) Overconfidence as a cause of diagnostic error in medicine. The American journal of medicine 121(5):S2–S23, DOI 10.1016/j.amjmed.2008.01.001

Bertillot H (2016) Comment l'évaluation de la qualité transforme l'hôpital. les deux visages de la rationalisation par les indicateurs. Cahiers internationaux de sociologie de la gestion 15:11–48

Bessat C, Zonon NA, D'Acremont V (2019) Large-scale implementation of electronic integrated management of childhood illness (eimci) at the primary care level in burkina faso: a qualitative study on health worker perception of its medical content, usability and impact on antibiotic prescription and resistance. BMC public health 19(1):449, DOI 10.1186/s12889-019-6692-6

Beynon-Davies P, Lloyd-Williams M (1998) Health information systems,'safety'and organizational learning. Health Informatics Journal 4(3-4):128–137, DOI 10.1177/146045829800400303

Bowman S (2013) Impact of electronic health record systems on information integrity: quality and safety implications. Perspectives in health information management 10(Fall)

Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PAC, Rubin HR (1999) Why don't physicians follow clinical practice guidelines?: A framework for improvement. Jama 282(15):1458–1465, DOI 10.1001/jama.282.15.1458

Cabitza F, Rasoini R, Gensini GF (2017) Unintended consequences of machine learning in medicine. Jama 318(6):517–518, DOI 10.1001/jama.2017.7797

Carroll AE, Biondich P, Anand V, Dugan TM, Downs SM (2013) A randomized controlled trial of screening for maternal depression with a clinical decision support system. Journal of the American Medical Informatics Association 20(2):311–316, DOI 10.1136/amiajnl-2011-000682

Chadwick D, Hall C, Rae C, Rayment M, Branch M, Littlewood J, Sullivan A (2017) A feasibility study for a clinical decision support system prompting hiv testing. HIV medicine 18(6):435–439, DOI 10.1111/hiv.12472

Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K (2019) Artificial intelligence, bias and clinical safety. BMJ Qual Saf pp bmjqs–2018, DOI 10.1136/bmjqs-2018-008370

Chang PL, Li Yc, Huang ST, Wang TM, Hsien ML (1996) Effects of a medical expert system on differential diagnosis of renal masses: A prospective study. Computerized medical imaging and graphics 20(1):43–48, DOI 10.1016/0895-6111(96)00029-8

Chaudhry B, Wang J, Wu S, Maglione M, Mojica W, Roth E, Morton SC, Shekelle PG (2006) Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. Annals of internal medicine 144(10):742–752, DOI 10.7326/0003-4819-144-10-200605160-00125

Choulak M, Marage D, Gisbert M, Paris M, Meinard Y (2019) A meta-decision-analysis approach to structure operational and legitimate environmental policies–with an application to wetland prioritization. Science of the Total Environment 655:384–394, DOI 10.1016/j.scitotenv.2018.11.202

Clancy CM (2009) Ten years after to err is human. American Journal of Medical Quality 24(6):525–528, DOI 10.1177/1062860609349728

Dalaba MA, Akweongo P, Williams J, Saronga HP, Tonchev P, Sauerborn R, Mensah N, Blank A, Kaltschmidt J, Loukanova S (2014) Costs associated with implementation of computer-assisted clinical decision support system for antenatal and delivery care: case study of kassena-nankana district of northern ghana. PLoS One 9(9), DOI 10.1371/journal.pone.0106416

De Dombal F (1987) Ethical considerations concerning computers in medicine in the 1980s. Journal of medical ethics 13(4):179–184, DOI 10.1136/jme.13.4.179

De Fauw J, Ledsam JR, Romera-Paredes B, Nikolov S, Tomasev N, Blackwell S, Askham H, Glorot X, O'Donoghue B, Visentin D, et al. (2018) Clinically applicable deep learning for diagnosis and referral in retinal disease. Nature medicine 24(9):1342–1350, DOI 10.1038/s41591-018-0107-6

De Marchi G, Lucertini G, Tsoukiàs A (2016) From evidence-based policy making to policy analytics. Annals of Operations Research 236(1):15–38, DOI 10.1007/s10479-014-1578-6

Deig CR, Kanwar A, Thompson RF (2019) Artificial intelligence in radiation oncology. Hematology/Oncology Clinics 33(6):1095–1104, DOI 10.1016/j.hoc.2019.08.003

Donaldson MS, Corrigan JM, Kohn LT, et al. (2000) To err is human: building a safer health system, vol 6. National Academies Press

Doran D, Schulz S, Besold TR (2017) What does explainable ai really mean? a new conceptualization of perspectives. arXiv preprint arXiv:171000794

Dua S, Acharya UR, Dua P (2014) Machine learning in healthcare informatics, vol 56. Springer

Eccles M, McColl E, Steen N, Rousseau N, Grimshaw J, Parkin D, Purves I (2002) Effect of computerised evidence based guidelines on management of asthma and angina in adults in primary care: cluster randomised controlled trial. Bmj 325(7370):941, DOI 10.1136/bmj.325.7370.941

Elsner P, Bauer A, Diepgen TL, Drexler H, Fartasch M, John SM, Schliemann S, Wehrmann W, Tittelbach J (2018) Position paper: Telemedicine in occupational dermatology–current status and perspectives. JDDG: Journal der Deutschen Dermatologischen Gesellschaft 16(8):969–974, DOI 10.1111/ddg.13605

Elstein AS, Friedman CP, Wolf FM, Murphy G, Miller J, Fine P, Heckerling P, Miller T, Sisson J, Barlas S, et al. (1996) Effects of a decision support system on the diagnostic accuracy of users: a preliminary report. Journal of the American Medical Informatics Association 3(6):422–428, DOI 10.1136/jamia.1996.97084515

Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542(7639):115–118, DOI 10.1038/nature21056

Etter DJ, McCord A, Ouyang F, Gilbert AL, Williams RL, Hall JA, Tu W, Downs SM, Aalsma MC (2018) Suicide screening in primary care: use of an electronic screener to assess suicidality and improve provider follow-up for adolescents. Journal of Adolescent Health 62(2):191–197, DOI 10.1016/j.jadohealth.2017.08.026

Ferretti V, Pluchinotta I, Tsoukiàs A (2019) Studying the generation of alternatives in public policy making processes. European Journal of Operational Research 273(1):353–363, DOI 10.1016/j.ejor.2018.07.054

Field MJ, Lohr KN, et al. (1990) Clinical practice guidelines: directions for a new program. National Academies Press

Fieschi M (1986) Expert systems for medical consultation. Health Policy 6(2):159–173, DOI 10.1016/0168-8510(86)90005-9

Garcia M (2016) Racist in the machine: The disturbing implications of algorithmic bias. World Policy Journal 33(4):111–117

Garg AX, Adhikari NK, McDonald H, Rosas-Arellano MP, Devereaux P, Beyene J, Sam J, Haynes RB (2005) Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. Jama 293(10):1223–1238, DOI 10.1001/jama.293.10.1223

Giordanengo A, Årsand E, Woldaregay AZ, Bradway M, Grottland A, Hartvigsen G, Granja C, Torsvik T, Hansen AH (2019) Design and prestudy assessment of a dashboard for presenting self-collected health data of patients with diabetes to clinicians: Iterative approach and qualitative case study. JMIR diabetes 4(3):e14002, DOI 10.2196/14002

Gonzalvo CM, de Wit HA, van Oijen BP, Hurkens KP, Janknegt R, Schols JM, Mulder WJ, Verhey FR, Winkens B, van der Kuy PHM (2017) Supporting clinical rules engine in the adjustment of medication (scream): protocol of a multicentre, prospective, randomised study. BMC geriatrics 17(1):35, DOI 10.1186/s12877-017-0426-3

Goodman KW (2016) Ethical and legal issues in decision support. In: Clinical decision support systems, Springer, pp 131–146, DOI 10.1007/978-3-319-31913-1_8

Gordon ES, Babu D, Laney DA (2018) The future is now: Technology's impact on the practice of genetic counseling. In: American Journal of Medical Genetics Part C: Seminars in Medical Genetics, Wiley Online Library, vol 178, pp 15–23, DOI 10.1002/ajmg.c.31599

Gräßer F, Beckert S, Küster D, Schmitt J, Abraham S, Malberg H, Zaunseder S (2017) Therapy decision support based on recommender system methods. Journal of healthcare engineering 2017, DOI 10.1155/2017/8659460

Guah M (1998) Evaluation and analysis of multimedia information system design and implementation at the coloscopy unit of st. james university hospital, leeds, uk. PhD thesis, MSc Dissertation, School of Management, UMIST, Manchester

Gunning D (2017) Explainable artificial intelligence (xai). Defense Advanced Research Projects Agency (DARPA), nd Web 2

Hall LH, Johnson J, Watt I, Tsipa A, O'Connor DB (2016) Healthcare staff wellbeing, burnout, and patient safety: a systematic review. PloS one 11(7):e0159015, DOI 10.1371/journal.pone.0159015

Heckerling P, Elstein A, Terzian C, Kushner M (1991) The effect of incomplete knowledge on the diagnoses of a computer consultant system. Medical Informatics 16(4):363–370, DOI 10.3109/14639239109067658

Heeks R (2006) Health information systems: Failure, success and improvisation. International journal of medical informatics 75(2):125–137, DOI 10.1016/j.ijmedinf.2005.07.024

Heeks R, Mundy D, Salazar A (1999) Why health care information systems succeed or fail. Information Systems for Public Sector Management DOI 10.2139/ssrn.3540062

Hoffer EP, Feldman MJ, Kim RJ, Famiglietti KT, Barnett GO (2005) Dxplain: patterns of use of a mature expert system. In: AMIA Annual Symposium Proceedings, American Medical Informatics Association, vol 2005, p 321

Hollon TC, Lewis S, Pandian B, Niknafs YS, Garrard MR, Garton H, Maher CO, McFadden K, Snuderl M, Lieberman AP, et al. (2018) Rapid intraoperative diagnosis of pediatric brain tumors using stimulated raman histology. Cancer research 78(1):278–289, DOI 10.1158/0008-5472.CAN-17-1974

Honaker SM, Dugan T, Daftary A, Davis S, Saha C, Baye F, Freeman E, Downs SM (2018) Unexplained practice variation in primary care providers' concern for pediatric obstructive sleep apnea. Academic pediatrics 18(4):418–424, DOI 10.1016/j.acap.2018.01.011

Hornik K, Stinchcombe M, White H, et al. (1989) Multilayer feedforward networks are universal approximators. Neural networks 2(5):359–366, DOI 10.1016/0893-6080(89)90020-8

Horrocks M, Michail M, Aubeeluck A, Wright N, Morriss R (2018) An electronic clinical decision support system for the assessment and management of suicidality in primary care: Protocol for a mixed-methods study. JMIR research protocols 7(12):e11135, DOI 10.2196/11135

Hunt DL, Haynes RB, Hanna SE, Smith K (1998) Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review. Jama 280(15):1339–1346, DOI 10.1001/jama.280.15.1339

Itani S, Lecron F, Fortemps P (2019) Specifics of medical data mining for diagnosis aid: A survey. Expert systems with applications 118:300–314, DOI 10.1016/j.eswa.2018.09.056

Jaroszewski AC, Morris RR, Nock MK (2019) Randomized controlled trial of an online machine learning-driven risk assessment and intervention platform for increasing the use of crisis services. Journal of consulting and clinical psychology 87(4):370, DOI 10.1037/ccp0000389

Jeanmougin M, Dehais C, Meinard Y (2017) Mismatch between habitat science and habitat directive: Lessons from the french (counter) example. Conservation Letters 10(5):634–644, DOI 10.1111/conl.12330

Johnston ME, Langton KB, Haynes RB, Mathieu A (1994) Effects of computer-based clinical decision support systems on clinician performance and patient outcome: a critical appraisal of research. Annals of internal medicine 120(2):135–142, DOI 10.7326/0003-4819-120-2-199401150-00007

Joo S, Yang YS, Moon WK, Kim HC (2004) Computer-aided diagnosis of solid breast nodules: use of an artificial neural network based on multiple sonographic features. IEEE transactions on medical imaging 23(10):1292–1300, DOI 10.1109/TMI.2004.834617

Kataria S, Ravindran V (2018) Digital health: a new dimension in rheumatology patient care. Rheumatology international 38(11):1949–1957, DOI 10.1007/s00296-018-4037-x

Kaushal R, Shojania KG, Bates DW (2003) Effects of computerized physician order entry and clinical decision support systems on medication safety: a systematic review. Archives of internal medicine 163(12):1409–1416, DOI 10.1001/archinte.163.12.1409

Keen J (1994) Information management in health services. Open University Press, DOI 10.1016/j.jacr.2017.08.033

Kirby AM, Kruger B, Jain R, Daniel P, Granger BB (2018) Using clinical decision support to improve referral rates in severe symptomatic aortic stenosis: a quality improvement initiative. CIN: Computers, Informatics, Nursing 36(11):525–529, DOI 10.1097/CIN.0000000000000471

Kostopoulou O, Porat T, Corrigan D, Mahmoud S, Delaney BC (2017) Diagnostic accuracy of gps when using an early-intervention decision support system: a high-fidelity simulation. Br J Gen Pract 67(656):e201–e208, DOI 10.3399/bjgp16X688417

Kulikowski CA (1988) Artificial intelligence in medical consultation systems: A review. IEEE Engineering in Medicine and Biology Magazine 7(2):34–39, DOI 10.1109/51.

1972

Kulikowski CA (2019) Beginnings of artificial intelligence in medicine (aim): Computational artifice assisting scientific inquiry and clinical art–with reflections on present aim challenges. Yearbook of medical informatics 28(01):249–256, DOI 10.1055/s-0039-1677895

Kuperman GJ, Gibson RF (2003) Computer physician order entry: benefits, costs, and issues. Annals of internal medicine 139(1):31, DOI 10.7326/0003-4819-139-1-200307010-00010

Leape LL (2000) Institute of medicine medical error figures are not exaggerated. Jama 284(1):95–97, DOI 10.1001/jama.284.1.95

Leape LL, Berwick DM (2005) Five years after to err is human: what have we learned? Jama 293(19):2384–2390, DOI 10.1001/jama.293.19.2384

López MM, López MM, de la Torre Díez I, Jimeno JCP, López-Coronado M (2017) Mhealth app for ios to help in diagnostic decision in ophthalmology to primary care physicians. Journal of medical systems 41(5):81, DOI 10.1007/s10916-017-0731-6

Masud R, Al-Rei M, Lokker C (2019) Computer-aided detection for breast cancer screening in clinical settings: Scoping review. JMIR medical informatics 7(3):e12660, DOI 10.2196/12660

Meinard Y, Cailloux O (2020) On justifying the norms underlying decision support. European Journal of Operational Research DOI 10.1016/j.ejor.2020.02.022

Meinard Y, Thébaud G (2019) L'identification syntaxonomique dans les démarches de gestion et/ou de restauration d'espaces naturels en france: pour ou contre? Naturae 2019(6):165–173

Meinard Y, Tsoukiàs A (2019) On the rationality of decision aiding processes. European Journal of Operational Research 273(3):1074–1084, DOI 10.1016/j.ejor.2018.09.009

Miller R, Masarie Jr F (1990) The demise of the "greek oracle" model for medical diagnostic systems. Methods of information in medicine 29(01):1–2

Miller RA (1994) Medical diagnostic decision support systems—past, present, and future: a threaded bibliography and brief commentary. Journal of the American Medical Informatics Association 1(1):8–27, DOI 10.1136/jamia.1994.95236141

Miller RA (2010) A history of the internist-1 and quick medical reference (QMR) computer-assisted diagnosis projects, with lessons learned. Yearbook of medical informatics 19(01):121–136, DOI 10.1055/s-0038-1638702

Miller RA (2016) Diagnostic decision support systems. In: Clinical Decision Support Systems, Springer, pp 181–208, DOI 10.1007/978-3-319-31913-1_11

Miller RA, McNeil MA, Challinor SM, Masarie Jr FE, Myers JD (1986) The internist-1/quick medical reference project—status report. Western Journal of Medicine 145(6):816

Miotto R, Wang F, Wang S, Jiang X, Dudley JT (2017) Deep learning for healthcare: review, opportunities and challenges. Briefings in bioinformatics 19(6):1236–1246, DOI 10.1093/bib/bbx044

Mitchell I, Schuster A, Smith K, Pronovost P, Wu A (2016) Patient safety incident reporting: a qualitative study of thoughts and perceptions of experts 15 years after 'to err is human'. BMJ Qual Saf 25(2):92–99, DOI 10.1136/bmjqs-2015-004405

Mokdad AH, Marks JS, Stroup DF, Gerberding JL (2004) Actual causes of death in the united states, 2000. Jama 291(10):1238–1245, DOI 10.1001/jama.291.10.1238

Morelli RA, Bronzino JD, Goethe JW (1987) Expert systems in psychiatry. Journal of medical systems 11(2-3):157–168, DOI 10.1007/BF00992350

Murphy GC, Friedman CP, Elstein AS, Wolf FM, Miller T, Miller J (1996) The influence of a decision support system on the differential diagnosis of medical practitioners at three levels of training. In: Proceedings of the AMIA Annual Fall Symposium, American Medical Informatics Association, p 219

Musen MA, Middleton B, Greenes RA (2014) Clinical decision-support systems. In: Biomedical informatics, Springer, pp 643–674, DOI 10.1007/978-1-4471-4474-8_22

Nadarzynski T, Bayley J, Llewellyn C, Kidsley S, Graham CA (2020) Acceptability of artificial intelligence (ai)-enabled chatbots, video consultations and live webchats as online platforms for sexual health advice. BMJ Sexual & Reproductive Health DOI 10.1136/bmjsrh-2018-200271

Onega T, Bowles EJA, Miglioretti DL, Carney PA, Geller BM, Yankaskas BC, Kerlikowske K, Sickles EA, Elmore JG (2010) Radiologists' perceptions of computer aided detection versus double reading for mammography interpretation. Academic radiology 17(10):1217–1226, DOI 10.1016/j.acra.2010.05.007

Osheroff JA, Teich JM, Levick D, Saldana FT Luis ans Velasco, Sittig DF, Rogeers KM, Jenders RA (2012) Improving outcomes with clinical decision support: an implementer's guide. HIMSS

Overhage JM, Tierney WM, Zhou XH, McDonald CJ (1997) A randomized trial of "corollary orders" to prevent errors of omission. Journal of the American Medical Informatics Association 4(5):364–375, DOI 10.1136/jamia.1997.0040364

Ozaydin B, Hardin JM, Chhieng DC (2016) Data mining and clinical decision support systems. In: Clinical Decision Support Systems, Springer, pp 45–68, DOI 10.1007/978-3-319-31913-1_3

Pare G, Elam JJ (1998) Introducing information technology in the clinical setting: Lessons learned in a trauma center. International Journal of Technology Assessment in Health Care 14(2):331–343, DOI 10.1017/S0266462300012290

Pasquali P, Sonthalia S, Moreno-Ramirez D, Sharma P, Agrawal M, Gupta S, Kumar D, Arora D, et al. (2020) Teledermatology and its current perspective. Indian Dermatology Online Journal 11(1):12, DOI 10.4103/idoj.IDOJ_241_19

Patel VL, Kushniruk AW, Yang S, Yale JF (2000) Impact of a computer-based patient record system on data collection, knowledge organization, and reasoning. Journal of the American Medical Informatics Association 7(6):569–585, DOI 10.1136/jamia.2000.0070569

Pearce C, McLeod A, Rinehart N, Patrick J, Fragkoudi A, Ferrigi J, Deveny E, Whyte R, Shearer M (2019) Polar diversion: using general practice data to calculate risk of emergency department presentation at the time of consultation. Applied clinical informatics 10(01):151–157, DOI 10.1055/s-0039-1678608

Peiffer-Smadja N, Rawson TM, Ahmad R, Buchard A, Pantelis G, Lescure FX, Birgand G, Holmes AH (2019) Machine learning for clinical decision support in infectious diseases: a narrative review of current applications. Clinical Microbiology and Infection DOI 10.1016/j.cmi.2019.09.009

Pivovarov R, Elhadad N (2015) Automated methods for the summarization of electronic health records. Journal of the American Medical Informatics Association 22(5):938–947, DOI 10.1093/jamia/ocv032

Pluchinotta I, Pagano A, Giordano R, Tsoukiàs A (2018) A system dynamics model for supporting decision-makers in irrigation water management. Journal of environmental management 223:815–824, DOI 10.1016/j.jenvman.2018.06.083

Pluchinotta I, Kazakçi AO, Giordano R, Tsoukiàs A (2019) Design theory for generating alternatives in public decision making processes. Group Decision and Negotiation

28(2):341–375, DOI 10.1007/s10726-018-09610-5
Poels PJ, Schermer TR, Schellekens DP, Akkermans RP, de Vries Robbé PF, Kaplan A, Bottema BJ, Van Weel C (2008) Impact of a spirometry expert system on general practitioners' decision making. European Respiratory Journal 31(1):84–92, DOI 10.1183/09031936.00012007
Porat T, Delaney B, Kostopoulou O (2017) The impact of a diagnostic decision support system on the consultation: perceptions of gps and patients. BMC medical informatics and decision making 17(1):79, DOI 10.1186/s12911-017-0477-6
Povyakalo AA, Alberdi E, Strigini L, Ayton P (2013) How to discriminate between computer-aided and computer-hindered decisions: a case study in mammography. Medical Decision Making 33(1):98–107, DOI 10.1177/0272989X12465490
Reider JM (2016) Impact of national policies on the use of clinical decision support. In: Clinical Decision Support Systems, Springer, pp 111–130, DOI 10.1007/978-3-319-31913-1_7
Richard A, Mayag B, Meinard Y, Talbot F, Tsoukiàs A (2018) How AI could help physicians during their medical consultations: An analysis of physicians' decision process to develop efficient decision support systems for medical consultations. In: PFIA 2018, Nancy, France
Rudin C, Radin J (2019) Why are we using black box models in ai when we don't need to? a lesson from an explainable ai competition. Harvard Data Science Review 1(2), DOI 10.1162/99608f92.5a8a3a3d
Sandvig C, Hamilton K, Karahalios K, Langbort C (2016) When the algorithm itself is a racist: Diagnosing ethical harm in the basic components of software. International Journal of Communication 10:19
Shortliffe E (2012) Computer-based medical consultations: MYCIN, vol 2. Elsevier
Shortliffe EH, Cimino JJ (2014) Biomedical informatics, 4th edn. Springer
Simon G, DiNardo CD, Takahashi K, Cascone T, Powers C, Stevens R, Allen J, Antonoff MB, Gomez D, Keane P, et al. (2019) Applying artificial intelligence to address the knowledge gaps in cancer care. The oncologist 24(6):772–782, DOI 10.1634/theoncologist.2018-0257
Sittig DF, Krall MA, Dykstra RH, Russell A, Chin HL (2006) A survey of factors affecting clinician acceptance of clinical decision support. BMC medical informatics and decision making 6(1):6, DOI 10.1186/1472-6947-6-6
Slain T, Rickard-Aasen S, Pringle JL, Hegde GG, Shang J, Johnjulio W, Venkat A (2014) Incorporating screening, brief intervention, and referral to treatment into emergency nursing workflow using an existing computerized physician order entry/clinical decision support system. Journal of Emergency Nursing 40(6):568–574, DOI 10.1016/j.jen.2013.10.007
Taylor B, Dinh M, Kwok R, Dinh D, Chu M, Tang E (2008) Electronic interface for emergency department management of asthma: a randomized control trial of clinician performance. Emergency Medicine Australasia 20(1):38–44, DOI 10.1111/j.1742-6723.2007.01040.x
Tierney WM, Overhage JM, Murray MD, Harris LE, Zhou XH, Eckert GJ, Smith FE, Nienaber N, McDonald CJ, Wolinsky FD (2003) Effects of computerized guidelines for managing heart disease in primary care. Journal of General Internal Medicine 18(12):967–976, DOI 10.1111/j.1525-1497.2003.30635.x
Titano JJ, Badgeley M, Schefflein J, Pain M, Su A, Cai M, Swinburne N, Zech J, Kim J, Bederson J, et al. (2018) Automated deep-neural-network surveillance of cranial images for acute neurologic events. Nature medicine 24(9):1337–1341, DOI 10.1038/

s41591-018-0147-y

Tsai TL, Fridsma DB, Gatti G (2003) Computer decision support as a source of interpretation error: the case of electrocardiograms. Journal of the American Medical Informatics Association 10(5):478–483, DOI 10.1197/jamia.M1279

Tsoukiàs A, Montibeller G, Lucertini G, Belton V (2013) Policy analytics: an agenda for research and practice. EURO Journal on Decision Processes 1(1-2):115–134, DOI 10.1007/s40070-013-0008-3

Van Der Sijs H, Aarts J, Vulto A, Berg M (2006) Overriding of drug safety alerts in computerized physician order entry. Journal of the American Medical Informatics Association 13(2):138–147, DOI 10.1197/jamia.M1809

Verdoorn S, Kwint HF, Hoogland P, Gussekloo J, Bouvy ML (2018) Drug-related problems identified during medication review before and after the introduction of a clinical decision support system. Journal of clinical pharmacy and therapeutics 43(2):224–231, DOI 10.1111/jcpt.12637

Voigt P, Von dem Bussche A (2017) The EU general data protection regulation (GDPR). Springer, DOI 10.1007/978-3-319-57959-7

Wachter RM (2004) The end of the beginning: Patient safety five years after 'to err is human' amid signs of progress, there is still a long way to go. Health affairs 23(Suppl1):W4–534, DOI 10.1377/hlthaff.W4.534

Warner HR, Haug P, Bouhaddou O, Lincoln M, Warner Jr H, Sorenson D, Williamson JW, Fan C (1988) Iliad as an expert consultant to teach differential diagnosis. In: Proceedings of the Annual Symposium on Computer Application in Medical Care, American Medical Informatics Association, p 371

Warner Jr H (1989) Iliad: moving medical decision-making into new frontiers. Methods of information in medicine 28(04):370–372, DOI 10.1055/s-0038-1636792

Watrous RL, Thompson WR, Ackerman SJ (2008) The impact of computer-assisted auscultation on physician referrals of asymptomatic patients with heart murmurs. Clinical Cardiology: An International Indexed and Peer-Reviewed Journal for Advances in the Treatment of Cardiovascular Disease 31(2):79–83, DOI 10.1002/clc.20185

Weingart SN, Toth M, Sands DZ, Aronson MD, Davis RB, Phillips RS (2003) Physicians' decisions to override computerized drug alerts in primary care. Archives of internal medicine 163(21):2625–2631, DOI 10.1001/archinte.163.21.2625

West CP, Dyrbye LN, Shanafelt TD (2018) Physician burnout: contributors, consequences and solutions. Journal of internal medicine 283(6):516–529, DOI 10.1111/joim.12752

Woolf SH (1993) Practice guidelines: a new reality in medicine: Iii. impact on patient care. Archives of internal medicine 153(23):2646–2655, DOI 10.1001/archinte.1993.00410230060008

World Health Organization, et al. (1992) The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines. Geneva: World Health Organization

Yanase J, Triantaphyllou E (2019) A systematic survey of computer-aided diagnosis in medicine: Past and present developments. Expert Systems with Applications p 112821, DOI 10.1016/j.eswa.2019.112821

Yang MS, Choi SI, Song WJ, Kim SH, Cho SH, Min KU, Kim JH, Chang YS (2018) Electronic consultation support system for radiocontrast media hypersensitivity changes clinicians' behavior. Allergy, asthma & immunology research 10(2):165–171, DOI 10.4168/aair.2018.10.2.165

Zhang K, Liu X, Liu F, He L, Zhang L, Yang Y, Li W, Wang S, Liu L, Liu Z, et al. (2018) An interpretable and expandable deep learning diagnostic system for multiple ocular diseases: qualitative study. Journal of medical Internet research 20(11):e11144, DOI 10.2196/11144

Zier J, Spaulding A, Finch M, Verschaetse T, Tarrago R (2017) Improved time to notification of impending brain death and increased organ donation using an electronic clinical decision support system. American Journal of Transplantation 17(8):2186–2191, DOI 10.1111/ajt.14312

Zou J, Schiebinger L (2018) Ai can be sexist and racist—it's time to make it fair. Nature 559:324–326, DOI 10.1038/d41586-018-05707-8