

The role of cognitive science and artificial intelligence in supporting clinical diagnosis

CLAUDIO LUCCHIARI, MARIA ELIDE VANUTELLI AND RAFFAELLA FOLGIERI

University of Milan, Italy

Abstract

Research suggests that doctors are failing to make use of technologies designed to optimize their decision-making skills in daily clinical activities, despite a proliferation of electronic tools with the potential for decreasing risks of medical and diagnostic errors. This paper addresses this issue by exploring the cognitive basis of medical decision making and its psychosocial context in relation to technology. We then discuss how cognitive-led technologies – in particular, decision support systems and artificial neural networks – may be applied in clinical contexts to improve medical decision making without becoming a substitute for the doctor’s judgment. We identify critical issues and make suggestions regarding future developments.

Keywords: artificial neural networks; cognitive science; diagnosis; decision support systems; medical decision making; medical error

1. Introduction

Despite technological advances in health management systems, diagnostic errors and adverse events remain an urgent concern in medicine (Verghese *et al.* 2015; Ferrah *et al.* 2017; Shojania and Dixon-Woods 2017). Research suggests that technologies designed to support and improve decision making are not being adopted, even though such tools may assist with step-by-step routines and

procedures for making decisions, approaching clinical cases or performing a given therapeutic intervention (Ely and Graber 2015). In this paper, we present a synthetic overview of the problem by presenting a combined cognitive-science and artificial-intelligence approach, thus showing how cognitive-led smart decision support technologies may be useful in clinical practice.

The following section describes a cognitive approach to medical decision making. Then, in Section 3, we discuss examples that show how artificial intelligence (AI) technology – in particular, decision support systems and artificial neural networks – can assist with medical decision making, leading to improved outcome. The implications and possible applications of these technologies are discussed more broadly in relation to cognitive processes in Section 4. In Section 5 we outline some final considerations with regard to the frame of actual medical praxis.

2. The role of cognitive science in studying doctors’ medical decision making

A patient is an information carrier who enters medical organization routines in search of a diagnosis and/or appropriate treatment. The extraction of this clinical information entails information framing, a process that allows the integration of data from several environmental (the patient and

the organization) and internal (memory) sources within a single cognitive structure, a so-called mental model (Reason 1990).

This first mental model establishes what a doctor may or may not think about a specific patient, since it is the sole cognitive basis on which doctors are able to generate a clinical hypothesis. However, although it may seem that similar processes would follow a formally defined logical path, this is not always the case: numerous individual, interpersonal and organizational factors – e.g., working conditions, interruptions, time, availability of beds, diagnostic tools and general resources – can significantly affect the linearity of the process, and in some unfortunate cases the process can become a trap (Norman *et al.* 2017). Although biases may act at all decision-making stages, this initial information synthesis stage, when a physician needs to pool entry data, is in particular fraught with the risk of cognitive failures (Wilson *et al.* 1999; Klein 2005).

Further, the processes make physicians particularly inept at simultaneously evaluating different causes and factors, leading them to choose only one possible sense-making course and to simplify the clinical situation by filtering out ‘unnecessary’ information, instead of integrating information into a complex model. The result, though, is actually oversimplification, and taking into consideration only one possible pathway increases vulnerability to cognitive biases (Croskerry *et al.* 2013).

Physicians often fail to realize the impact of implicit cognitive mechanisms (e.g., heuristics and biases) in their decisions. When asked, physicians often assert that they always use logic, analytical reasoning and pre-existing guidelines in their medical decision making. The reality, though is that decision makers usually fail to recognize the factors that determine their choices. Instead, they often find *post hoc* explanations based on rational considerations about past events; finding a justification for a choice is easier and more convenient than analyzing why it was made. Naturally, this is another source of decision bias (Kahneman and Tversky 1973; Fischhoff 1975) and a potential hazard in medical decision making (Bornstein and Emler 2001).

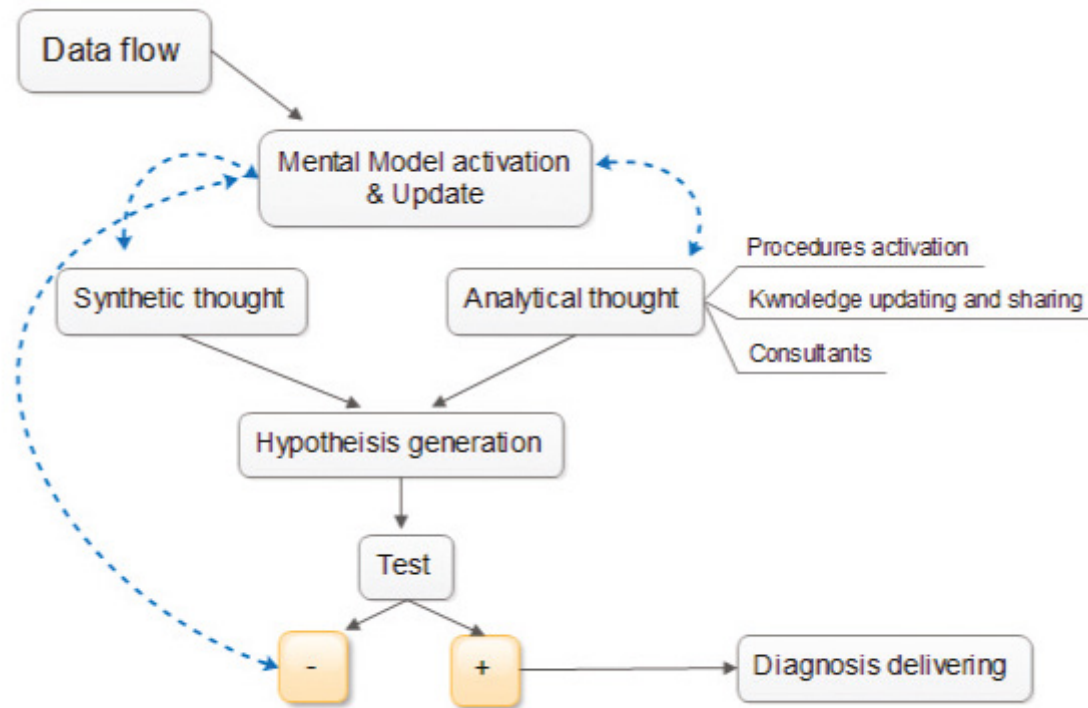
We argue that to optimize the diagnostic process and to prevent errors, doctors should exploit the

potential of computer science technology and AI, so as to better balance intuitive and analytical processes (Lucchiari and Pravettoni 2012; Marcum 2012) and thus avoiding falling into dangerous cognitive traps while using their knowledge and skills. Indeed, we argue that mind-friendly and human-centered decision-making support tools (Memon *et al.* 2014) are today equally important as technical and logical competencies in ensuring a rational diagnostic process. However, although use of technology is often evoked as a way to prevent medical errors, its impact is often limited by physicians’ difficulties in integrating it within clinical praxis. We thus believe that a fundamental step in the improvement of the diagnostic process consists of developing specific skills in using the methods provided by smart technologies, which requires not only technical investment but also cognitive and educational efforts.

3. The role of artificial intelligence

The systematic use of procedures and guidelines does not always guarantee the best intervention to be performed. This is not just the case when epidemiological data are inadequate, but because such an approach does not conform with the reality of how the human mind actually functions – decision making in fact often relies on inner guidelines, or so-called ‘mindlines’ (Gabbay and le May 2008). These procedures make experts fast and efficient in making critical decisions but increase the probability of committing errors in some decisional contexts. Thus, the problem of ‘immunizing’ medical decisions from errors cannot be addressed simply by education and training (Sherbino *et al.* 2014); various studies have suggested different approaches, but the use of cognitive-led technology to support medical decisions currently appears to be the most promising.

As noted in the previous section, physicians frequently need to simplify complex problems by automatic mental processes. In so doing, though, a diagnosis may be affected by cognitive biases (Holyoak 2012), even though decision-makers are rarely aware of this burden on their decisions (Figure 1). Further, physicians today work in the context of ‘big data’ (Lohr 2012): there is a great deal of data handling, not just to organize and

Figure 1. *Diagnostic schema*

categorize complex information but, foremost, to retrieve elements useful for medical diagnosis and prognosis. This context means that real-time adaptive systems such as decision support systems (DSSs) and artificial neural networks (ANNs – Baxt 1992), the latter involving machine learning algorithms (Kononenko 2001), are now crucial for making it possible for physicians to provide effective early interventions (Bertoni *et al.* 2005), to identify at-risk patients and to make diagnostic decisions (Lucchiari *et al.* 2014), as well as for optimizing the organization of healthcare systems (Berwick *et al.* 2008). These systems allow doctors to adapt to specific cases and needs, while taking account of psychological and cognitive factors.

As an example, one recent study (Olczak *et al.* 2017) compared the diagnostic accuracy of an AI system and human experts in an orthopedic setting when interpreting skeletal radiographs. Wrist, hand and ankle radiographs were selected and submitted to the ANN, as well as to two senior orthopedic surgeons. The results showed that the final accuracy for fractures was estimated at 83% by the best-performing ANN, similar to the rate of human diagnosis. Since training an ANN is much

easier, faster and cheaper than training doctors, and given the potentiality of the tool to improve its performance with time and experience (e.g. learning by processing new datasets), the study strongly suggests that in the near future similar ANNs could improve the accuracy of the diagnostic process. Furthermore, once developed and trained in one clinical context an ANN can be simply adapted to other hospitals, spreading the related clinical advantage even in contexts where expert doctors are not easily available (e.g., small hospitals).

Cognitive-led smart technologies may be used both to prevent errors and to optimize medical decision making. In what follows, we describe decision support systems and artificial neural networks in more detail.

3.1. *Decision support systems*

A DSS is a knowledge resource that, when fed with patients' data, can generate case-specific advice. In particular, the tool may elucidate the decision context, highlighting possible outcomes and consequences and thus promoting critical thinking about the decision to take.

Currently, there are many kinds of DSS available, including the following:

- *computerized reminder systems*, for increased effectiveness in preventive care settings;
- *computerized information tools*, for detailed information about pathogenesis, typical symptoms, treatment options and prognoses;
- *care pathways* for patients;
- *computerized patient interviewing*, consisting of a checklist for the patient to complete, after which the data are presented to the doctor in summary form;
- *telemedicine systems*;
- *summary sheet* with the results of special investigations and advice on the interpretation of results; and
- *distance learning materials*, used away from patients (e.g., in a course, or self-study).

Furthermore, decision support tools may be categorized based on their role in the decision process. Thus, DSSs may generate:

- inferences about diagnoses, prognoses or treatment effects;
- summaries of complex low-level data, e.g., through an intuitive display;
- a filter of cases similar to that of the individual patient at hand;
- alerts for rare events, e.g., high-risk cases; and
- audit compliance and reports about consistency in the use of clinical discretion within a clinical guideline.

Applying cognitive science to DSS implementation may help the diagnosis process in several ways. In particular, it may assist in providing a mind-friendly structuring of problem representation, e.g. of a clinical scenario.

An example of a DSS to improve patient care comes from an Italian project developed for an internal medicine department at a multi-site hospital system that comprised different health units. The protocol was aimed at evaluating the quality of care and the effectiveness of patient-specific reminders generated by the Medilogy Decision Support System (MediDSS), which was formed using international evidence-based guidelines and approved by an international panel of experts

(Medilogy Srl., Medilogy Decision Support System). The MediDSS includes a drug–drug interaction database containing evidence-based information about the harms and benefits of more than 18,000 drug interactions (Böttiger *et al.* 2009), and it may be used to integrate structured patient data from the electronic health records (EHR) to generate patient-specific reminders, therapeutic suggestions and diagnosis-specific links to full-text guidelines. It is conceived in a way that, when the clinician opens a patient’s record, the reminders pop up on the monitors.

It is important to note that the application of medical decision systems should always consider the specific aims at each stage.

3.1.1. *Barriers to adoption*

Although many studies have highlighted the importance of clinical DSSs in improving physicians’ decision making, such systems are not employed regularly. This may be due to several factors. First, health personnel often find it difficult to understand how to use a DSS efficiently in their working routine (it could be too slow or too complicated) or even why they should. Second, the DSS may not produce a timely and effective output able to influence physicians’ decisions. Sometimes, the output may be in a format that is not easy to understand or that is difficult to apply in a given situation. As such, the result may not be straightforward enough to encourage users to change a decision or to suggest new actions. Finally, health personnel may consider their performance already optimal.

A previous study (Lucchiari *et al.* 2015) showed that DSSs are often not considered particularly relevant as expert doctors are supposed to make a decision on their own. This confirms a classical misunderstanding of the role of decision support technologies, which is that they are intended to be a substitute for doctors’ responsibility rather than to extend their decision-making skills.

We suggest that to overcome such barriers to the widespread use of DSSs, technical aspects, cognitive issues and psychosocial factors should be considered together within a single decision support model. Currently, some developers appear to design technologically advanced systems with little relevance to the real world, while others create DSSs without first determining whether a clinical

need exists (Randolph *et al.* 1999). It is instead first necessary to identify the most effective method to improve the actual praxis, this being the critical variable for deploying a cognitive-led technology.

3.2. *Neural networks to support clinical decisions*

An artificial neural network (ANN) is a mathematical tool able to simulate human information processes through calculations. There are many types of problems that ANNs can deal with, e.g., prediction, network fitting, classification and pattern recognition, clustering and dynamic time series (Demuth and Beale 1993). Each ANN is made up of several nodes (or neurons) divided into layers and connected to each other. A specific mathematical function provides the strength of these connections, and their values vary across time depending on training (see Figure 2). The output is the result of these connections, which point to a final node that will contain the network outcome. Through proper training, an ANN learns how to connect several variables to produce the desired outcome; e.g., an assessment of whether a person with a set of symptoms has a particular clinical condition.

A typical ANN is a multiple-layer network with a single input layer but multiple hidden and output layers. At first, it is necessary to train the network in which input should match a particular output (Hagan *et al.* 1996). During this training phase, the information goes through some hidden layers that have random initial weights, which helps the network to learn the input/output association. Using feed-forward and back-propagation algorithms, the weights of the hidden layers are re-calculated to support the network output in reducing the gap to the desired target values. In this way, given that an ANN model fits specific needs, it is possible to obtain a trained neural network able to provide the desired outcome with an estimated error.

Many types of neural networks may be differentiated, based on some fundamental characteristics that include application, learning method, learning algorithms and the architecture of the connections. One fundamental classification here is based on the type of learning, either supervised (when there is a gold-standard database to train the network

to obtain a specific output to an input dataset) or unsupervised. We then may distinguish three basic categories:

1. *Associative memories*: ANNs can learn associations between patterns (complex sets of data as the pixels of a computer image) so that the presentation of pattern 'A' gives as output pattern 'B', even if the former is inaccurate or partial (resistance to noise). It is then possible to use an associative memory to give a full pattern as output in response to a partial pattern in the input.
2. *Simulators of mathematical functions*: ANNs can understand the function that links output with input, based on the examples provided during the learning phase. After this phase, the network can provide an output in response to an input which was not used during the training. It follows that the network can interpolate and extrapolate rules from a training dataset. Such capability is easily verifiable by training a network with a sequence of input/output data coming from a known function, and it is particularly useful for the treatment and prediction of phenomena with unclear mathematical relationships between input and output. In any case, the network works like a 'black box,' since the function implemented is not intelligible. Error back-propagation networks, which are currently the most used training algorithm for effectiveness and flexibility, are part of this class.
3. *Classifiers*: ANNs classify data into specific categories based on similarity. In this last type, the network makes use of unsupervised learning, or self-organizing learning, in which the input data are not distributed on a default set of categories. The learning algorithm of a neural network depends on the type of application, as well as on the architecture of the connections. The classical architecture of a classifier is based on Kohonen self-organizing maps, which refer to unsupervised neural networks that represent their distribution as a map, where similar samples are mapped closely together.

Neural networks are built to work in parallel and can process large amounts of data simultaneously. They are sophisticated statistical systems with

good noise immunity: if system units present a malfunction, although the network as a whole would experience a degradation of performance it is unlikely to face a system crash. Consequently, ANNs are particularly robust. However, a model produced by a neural network, although very efficient, is not explainable in human symbolic language. Outcomes must be accepted as if they were coming from a black box. In other words, and differently from an algorithmic system, you cannot follow step-by-step the path from the input to the output, since the system achieves the result without showing the process. This fact may limit the use of ANNs in some cases.

Furthermore, as with any modeling algorithm, the neural networks are efficient only if the predictor variables are chosen with care. They cannot effectively treat categorical variables (for example, the patient's residence address) with many different values. The training phase sets the weights of individual neurons, and this may take some time if the number of records and variables analyzed is particularly large. There are no theorems or models to determine the optimal network, and the success of a network then depends greatly on the experience of the designer.

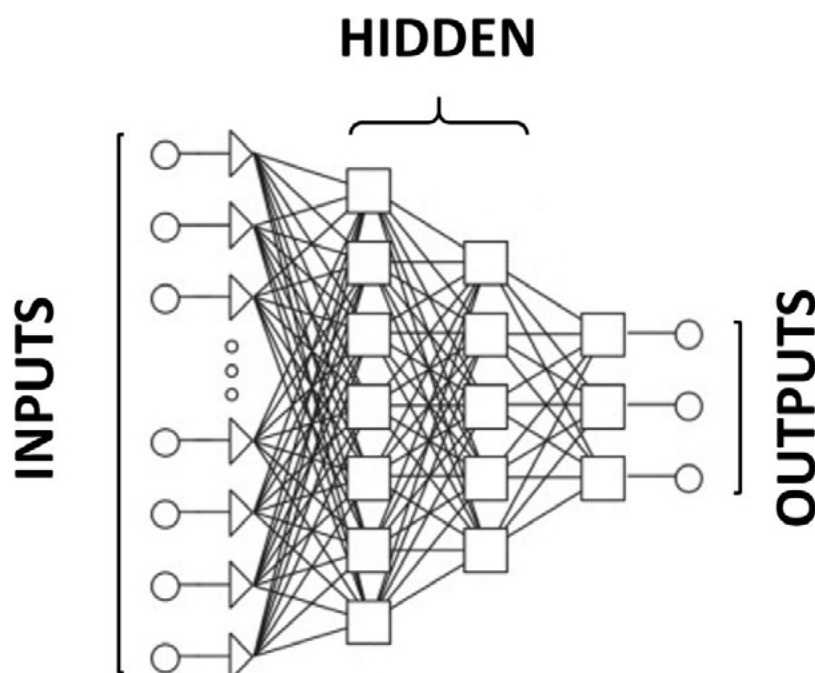
Developing a prediction system by neural networks requires following a series of operations that

affect the result and the overall effectiveness of the forecast. The main steps are as follows:

- designating the objective to forecast;
- collecting data and creating a learning database to train the neural network;
- defining the architecture and parameters necessary for the definition of the neurons connection weights; and
- generalization of the output.

The dataset preparation is a particularly important phase of a neural network implementation. The first consideration here is data selection: this is intuitively decisive, as the learning of any biological or artificial system depends on the available information. The reason the network needs a training phase is to better approximate a function, and to produce accurate forecasting. This process implies comparing the response provided by the network to a given input (independent variable) with the already known value of a dependent variable (output). By changing the weights of connections, it is possible to minimize the difference between the desired output (the value of the function for specific input values) and the actual output (the output value of the neural network for the same input values). In this phase, a continuous calibration of

Figure 2. *A back-propagation artificial neural network*



the database is necessary to eliminate data that do not modify the network functioning.

This procedure should follow a series of steps:

1. a broad definition of the first database;
2. implementation of first learning network;
3. evaluation of the information content of the individual variables;
4. analysis of the correlation matrix between the input variables;
5. elimination of less significant and more self-interrelated variables; and
6. operation of subsequent learning network with a reduced and optimized database.

In this way, it is possible to identify the optimal training dataset according to an iterative scheme that should maximize the learning ability of the network and minimize the information (input) needed to feed the network. Once the dataset is ready, a part of it is used to train the network (the training set), while the remaining part is required to test the predictive capabilities of the network (testing set).

Neural networks are particularly suitable when the relationship between input and output is very complicated in terms of the number of variables, and when the relationship is non-linear. ANNs perform an iterative and empirical exploration of data, to identify 'soft rules' that adequately describe the data interdependence; therefore, they are particularly powerful when we have many variables without a good explanatory hypothesis to organize them, and this is the primary use of neural networks in the medical field. Each stage that traditionally characterizes the medical decision-making process may benefit from the use of ANNs. In particular, the approach allows fast computing of the success probability of each alternative, and then an assessment of the related utility value. Furthermore, it can be used to test decision validity (Lisboa and Taktak 2006).

The power of these tools for predicting outcomes and supporting clinical decisions is widely recognized today, and the use of ANNs within medical settings is increasing rapidly – a recent PubMed search reveals that since 2000 more than 150 scientific papers have been dedicated to this issue. The specific contribution that an ANN may provide is due to its ability to integrate a large

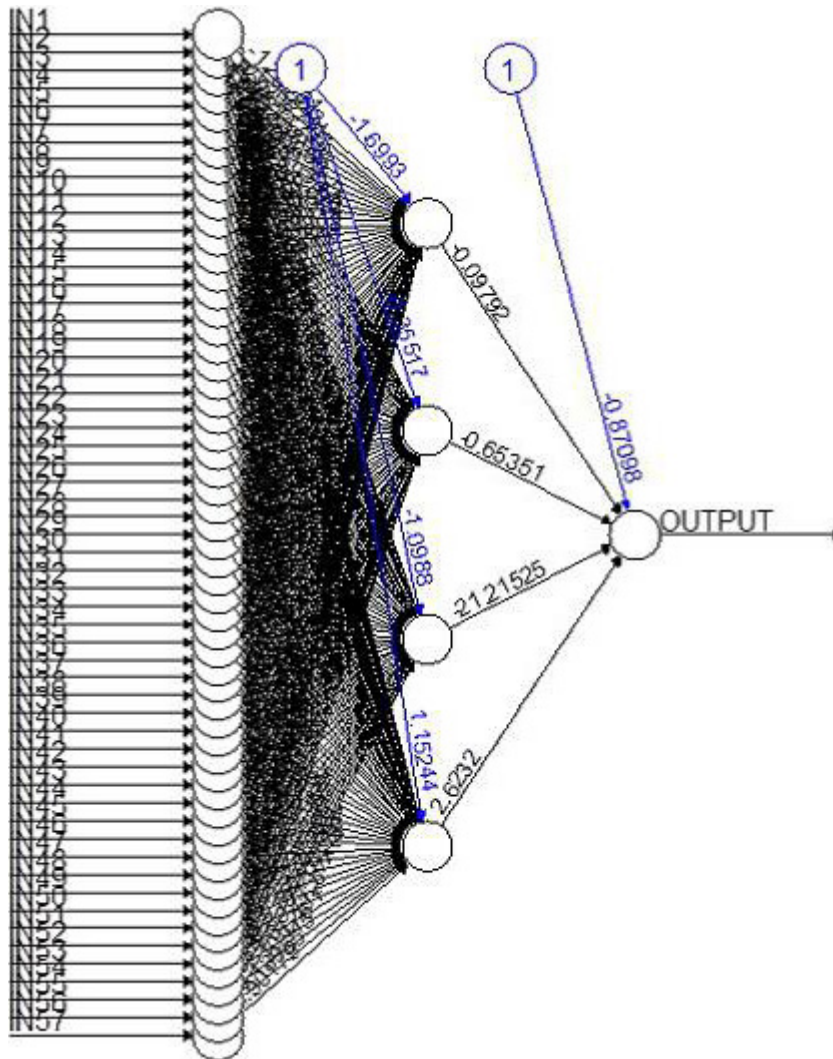
number of parameters to predict a clinical outcome without the kinds of cognitive constraints that lead doctors to prune the data. Whereas doctors use heuristics that simplify a situation based on their individual experience, ANNs prune variables based only on mathematic computing. Furthermore, ANNs can predict outcomes in a fast and reliable way, being able to learn with training. With ANNs, the choice of the best parameters to consider in a given context is explicit, evidence-based and standardized, since it is not the consequence of individual mindlines. Furthermore, since an ANN learns once a new item is entered into the system, tool validity may improve with time and adapt to new inputs. As such, these tools also allow the construction of a shared cognition that may support all medical decisions in the same context.

3.2.1. A network for surgical site infection (SSI) prediction

We propose here an example we have developed for prediction in a surgical setting. In particular, we projected an ANN aimed at estimating the probability of the occurrence of a surgical site infection (SSI) within one year after hip or knee prosthetic implant surgery. In this context, there is sufficient data about patients, the hospital stay and intervention characteristics that may serve as input for the ANN.

The literature emphasizes how risk factors are difficult to identify and how patients' characteristics should be used with caution to predict SSI, since they vary from patient to patient. It is not clear which combination of signs, symptoms and other parameters gives rise to a good or bad post-surgery pathway. The AI implemented for this pilot study is a feed-forward network, which uses a gradient descent-learning algorithm with supervised learning and back-propagation training. The initial database was a 58×690 table, containing 58 variables for each of the 690 patients of the sample considered. In the first phase, 552 subjects (80% of the whole sample) were used to train the network. The remaining 138 patients were used to feed the network in the testing phase. The sensitivity of the network was 76.9% (CI 53.9%; 99.8%), with a 100% of specificity (CI 99.9%; 100%); the positive predictive value of the network was 100%, and the negative predictive value was 98.1%; finally, the positive likelihood ratio turned out to be far

Figure 3. The 58 input neurons and four hidden neurons neural network used to predict SSI



greater than 10 (> 769), and the negative one was equal to 0.231.

The best result was obtained by using four hidden neurons (see Figure 3). The error threshold is more than acceptable (4.1), the time required to go through the steps (288 seconds) and for training (15 seconds) is in both cases excellent, the convergence of the network is very close to 100% (99.9994%) and the predictive capability (intended as general correspondence between the real output and the output calculated by the network) is 76%.

Our artificial neural network showed a good predictive value concerning an exclusively physician-led decision process that requires careful and costly tests (bacterial culture, tissue ultrasound, scintigraphy and so on) to be confirmed. This neural network, therefore, proved to be a

potentially useful cognitive support mechanism for the decision-making process.

Also, the neural network designed was proved to be a SpPin diagnostic test due to high specificity (100%). This means that the trained system does not give rise to false positive cases and can confirm the presence of infection with absolute precision (all persons labeled positive became infected). Also, the value of sensitivity (76.9%) is good, as it limits the possibility that a subject potentially affected by SSI is considered at risk by the network. Under the positive predictive value of 1 this approach confirmed the capability of a network to identify and confirm infection, even though SSI prevalence is quite low (about 1.88%); this means that, when the data for a new patient are entered into the network, there is maximum

capacity for alert. Earlier SSI diagnosis also makes it easier to plan a patient's therapeutic journey inside and outside the hospital, with fewer hospital stays and thus less cost for hospitals.

The neural network approach has extraordinary potential because of its fault tolerance and the ability to extract new solutions efficiently and in the shortest time possible. Further, whatever decision doctors take will be based on overt and sharable considerations, making them fully accountable for their choices.

4. Implications and applications

Physicians today can collect a great deal of data about their patients, ranging from personal to clinical information, including pharmacological treatments and their effects, different prognoses, imaging scans and other reports. During the diagnostic journey, such data is processed, evaluated and assigned towards the most likely diagnostic pathway. However, as discussed above, humans tend to solve complexity by using experience-based heuristics. These cognitive shortcuts can simplify a situation, so that doctors select only essential data to reach the best solution within a reasonable time. However, this process can also exclude potentially relevant information. Moreover, some data are also discarded because traditional statistical techniques are not robust enough to handle the amounts of data now generated (Sharpe and Caleb 1994). The use of cognitive-led smart technologies can overcome this possible loss of information, as well as oversimplification. They make it possible to process several variables together and to present them in an accessible format so as to be easily applied to a specific clinical picture.

Physicians should consider such tools an essential part of their clinical praxis, not just as an aid for avoiding medical errors, but also as a way to promote optimization, thus reducing risks for patients and increasing clinical and economic efficiencies. Cognitive-led smart technologies should be part of their daily routine, but not considered as being in competition with their experience and professionalism. However, despite the considerable attention that has been given to this topic over the last 20 years, doctors still need support in using them. They need to focus not only on DSS

or other similar tools but also on making specific and customized decision tools, for instance, based on the use of ANNs. As such, the construction of such tools requires an interdisciplinary effort. Doctors should be involved in the definition of the decision model to be implemented, be in charge of the data collection procedures and decide how and when to integrate the tools' outcomes with their daily activities. We argue that the use of these technologies is the most promising way to reduce medical errors, but that this will be possible only with the full involvement of doctors in their actual application.

Being involved in this way will increase physicians' trust in decision-making technology, thus allowing them to focus their attention on aspects of their role specifically related to their expertise while delegating mere data processing to computers. This way, automatic processing will support the human mind toward making balanced decisions – i.e., decisions that incorporate characteristics of both natural and artificial cognition.

5. Conclusion

In this paper, we have addressed the ability of cognitive science to explain which mental mechanisms lead doctors to take diagnostic decisions and the enormous potential of AI to support physicians' decision making at all stages of this process. AI confirms that a cognitive approach represents the primary method for improving the reliability of making a diagnosis, and, as a corollary, provides a monitoring system to test the appropriateness of decisions, thus helping health personnel as well as health organizations to learn by experience and to adapt when new demands emerge. By using these technologies, it is possible to avoid some of the most serious cognitive traps and biases that undermine the daily work of health personnel, such as premature closure and overconfidence (Lucchiari and Pravettoni 2013). AI also makes it possible to keep track of each step, to analyze the entire process backwards when things go wrong (e.g., due to a violation) and to find the origin of possible adverse events.

However, these strategies are not suitable for all medical settings. Also, in many health organizations, including national health systems, the use

of cognitive-led AI is poor or, at least, is scarcely integrated into daily activities. Thus, a more substantial effort by organizations in this direction would be desirable, with health personnel education and training devoted to developing the smarter use of decision technologies, even if the issue is not only a matter of the technology itself. Indeed, clinical decision-making tools are rarely used by physicians to make real decisions about specific patients, because the task they support does not match the physician's mental task. In previous works (Lucchiari and Pravettoni 2012), a cognitive balanced model (CBM) was suggested to describe how the clinical decision setting should be represented by a functional balance between analysis and intuition, guidelines and mindlines. CBM underlines the need for a doctor to develop both intuitive and analytical skills, but also the need for a support system that will help physicians find the balance needed case by case, adapting their thinking style to fit with the actual demands of the problem.

Consequently, physicians, nurses and all health personnel should all develop higher decision awareness. Although medical decision making is under in-depth scientific investigation all over the world, and results are diffused and shared, many physicians still lack necessary knowledge about basic decision theory, especially when we shift attention from probabilities to psychological issues. Heuristics and biases are now well-known lemmas, but few decision makers appraise their role during a decision process – for instance, when making a diagnosis. Many errors are due to cognitive factors, and most of them should be prevented by developing awareness of this fact. We strongly believe that cognitive-led AI research could soon play a significant role in addressing this issue and in aiding physicians in improving their decision-making skills.

References

- Baxt, William G. (1992) Analysis of the clinical variables driving decision in an artificial neural network trained to identify the presence of myocardial infarction. *Annals of Emergency Medicine* 21 (12): 1439–1444. [https://doi.org/10.1016/S0196-0644\(05\)80056-3](https://doi.org/10.1016/S0196-0644(05)80056-3)
- Bertoni, Alberto, Raffaella Folgieri and Giorgio Valentini (2005) Bio-molecular cancer prediction with random subspace ensembles of support vector machines. *Neurocomputing* 63: 535–539. <https://doi.org/10.1016/j.neucom.2004.07.007>
- Berwick, Donald M., Thomas W. Nolan and John Whittington (2008) The triple aim: Care, health, and cost. *Health affairs* 27 (3): 759–769. <https://doi.org/10.1377/hlthaff.27.3.759>
- Bornstein, Brian H. and A. Christine Emler (2001) Rationality in medical decision making: A review of the literature on doctors' decision-making biases *Journal of Evaluation in Clinical Practice* 7 (2): 97–107. <https://doi.org/10.1046/j.1365-2753.2001.00284.x>
- Böttiger, Ylva, Kari Laine, Marine L Andersson, Tuomas Korhonen, Björn Molin, Marie-Louise Ovesjö *et al.* (2009) SFINX – A drug-drug interaction database designed for clinical decision support systems. *European Journal of Clinical Pharmacology* 65 (6): 627–633. <https://doi.org/10.1007/s00228-008-0612-5>
- Croskerry, Pat, Geeta Singhal and Sílvia Mamede (2013) Cognitive debiasing 1: Origins of bias and theory of debiasing. *BMJ Quality Safety* 22 (Suppl. 2): ii58–ii64. <https://doi.org/10.1136/bmjqs-2012-001712>
- Demuth, Howard and Mark Beale (1993) *Neural Network Toolbox For Use with Matlab – User's Guide* (Version 3.0). Natick, MA: The MathWorks.
- Ely, John W. and Mark A. Graber (2015) Checklists to prevent diagnostic errors: A pilot randomized controlled trial. *Diagnosis* 2 (3): 163–169. <https://doi.org/10.1515/dx-2015-0008>
- Ferrah, Noha, Janaka J. Lovell and Joseph E. Ibrahim (2017) Systematic review of the prevalence of medication errors resulting in hospitalization and death of nursing home residents. *Journal of the American Geriatrics Society* 65 (2): 433–442. <https://doi.org/10.1111/jgs.14683>
- Fischhoff, Baruch (1975) Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance* 1 (3): 288–299. <https://doi.org/10.1037/0096-1523.1.3.288>
- Gabbay, John and Andrée le May (2008) Practice made perfect: Discovering the roles of a community of general practice. In Andrée le May (ed.) *Communities of Practice in Health and Social Care*, 49–65. Oxford: Blackwell.

- Hagan, Martin T., Howard B. Demuth and Mark H. Beale (1996). *Neural Network Design*. Boston: PWS Publishing.
- Holyoak, Keith J. (2012) Analogy and relational reasoning. In Keith J. Holyoak and Robert G. Morrison (eds) *The Oxford Handbook of Thinking and Reasoning*, 234–259. New York: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199734689.001.0001>
- Kahneman, Daniel and Amos Tversky (1973) On the psychology of prediction. *Psychological Review* 80 (4): 237–251. <https://doi.org/10.1037/h0034747>
- Klein, Jill G. (2005) Five pitfalls in decisions about diagnosis and prescribing. *British Medical Journal* 330 (7494): 781–783. <https://doi.org/10.1136/bmj.330.7494.781>
- Kononenko, Igor (2001). Machine learning for medical diagnosis: History, state of the art and perspective. *Artificial Intelligence in Medicine* 23 (1): 89–109. [https://doi.org/10.1016/S0933-3657\(01\)00077-X](https://doi.org/10.1016/S0933-3657(01)00077-X)
- Lisboa, Paulo J. and Azzam F. G. Taktak (2006). The use of artificial neural networks in decision support in cancer: A systematic review. *Neural Networks* 19 (4): 408–415. <https://doi.org/10.1016/j.neunet.2005.10.007>
- Lohr, Steve (2012) The age of big data. *The New York Times*, 11 February. Available online: <https://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html>
- Lucchiari, Claudio, Raffaella Folgieri and Gabriella Pravettoni (2014) Fuzzy cognitive maps: A tool to improve diagnostic decisions. *Diagnosis* 1 (4): 289–293. <https://doi.org/10.1515/dx-2014-0026>
- Lucchiari, Claudio, Fabio Macchi and Gabriella Pravettoni (2015) Patient empowerment and quality of life: A qualitative study on the doctors' perspective. *Abstracts of the 2015 International MASCC/ISOO Symposium*. Special issue of *Supportive Care in Cancer* 23 (Suppl. 1): S252. <https://doi.org/10.1007/s00520-015-2712-y>
- Lucchiari, Claudio and Gabriella Pravettoni (2012) Cognitive balanced model: a conceptual scheme of diagnostic decision making. *Journal of Evaluation in Clinical Practice* 18 (1): 82–88. <https://doi.org/10.1111/j.1365-2753.2011.01771.x>
- Lucchiari, Claudio and Gabriella Pravettoni (2013) The role of patient involvement in the diagnostic process in internal medicine: A cognitive approach. *European Journal of Internal Medicine* 24 (5): 411–415. <https://doi.org/10.1016/j.ejim.2013.01.022>
- Marcum, James A. (2012) An integrated model of clinical reasoning: Dual-process theory of cognition and metacognition. *Journal of Evaluation in Clinical Practice* 18 (5): 954–961. <https://doi.org/10.1111/j.1365-2753.2012.01900.x>
- Memon, Tasneem, Jie Lu and Farookh Khadeer Hussain (2014) Human-centric cognitive decision support system for ill-structured problems. In Peijun Guo and Witold Pedrycz (eds) *Human-Centric Decision-Making Models for Social Sciences*, 289–313. Berlin: Springer. https://doi.org/10.1007/978-3-642-39307-5_12
- Norman, Geoffrey R., Sandra D Monteiro, Jonathan Sherbino, Jonathan S Ilgen, Henk G Schmidt and Silvia Mamede (2017) The causes of errors in clinical reasoning: Cognitive biases, knowledge deficits, and dual process thinking. *Academic Medicine* 92 (1): 23–30. <https://doi.org/10.1097/ACM.0000000000001421>
- Olczak, Jakub, Niklas Fahlberg, Atsuto Maki, Ali Sharif Razavian, Anthony Jilert, André Stark, Olof Sköldenberg and Max Gordon (2017) Artificial intelligence for analyzing orthopedic trauma radiographs: Deep learning algorithms – Are they on par with humans for diagnosing fractures? *Acta Orthopaedica* 88 (6): 581–586. <https://doi.org/10.1080/17453674.2017.1344459>
- Randolph, Adrienne G., Mary B. Zollo, Marlene J. Egger, Gordon H. Guyatt, Robert M. Nelson and Gregory L. Stidham (1999) Variability in physician opinion on limiting pediatric life support. *Pediatrics* 103 (4): e46–e46. <https://doi.org/10.1542/peds.103.4.e46>
- Reason, James (1990) *Human Error*. Cambridge: Cambridge University Press.
- Sharpe, Peter K. and Praminda Caleb (1994) Artificial neural networks within medical decision support systems. *Scandinavian Journal of Clinical and Laboratory Investigation* 54 (Suppl. 219): 3–11. <https://doi.org/10.3109/00365519409088571>
- Sherbino, Jonathan, Kulamakan Kulasegaram, Elizabeth Howey and Geoffrey Norman (2014) Ineffectiveness of cognitive forcing strategies to reduce biases in diagnostic reasoning: A controlled trial. *Canadian Journal of Emergency Medicine* 16 (1): 34–40. <https://doi.org/10.2310/8000.2013.130860>
- Shojania, Kaveh G. and Mary Dixon-Woods (2017) Estimating deaths due to medical error: The ongoing controversy and why it matters. *BMJ Quality & Safety* 26 (5): 423–428. <https://doi.org/10.1136/bmjqs-2016-006144>

Verghese, Abraham, Blake Charlton, Jerome P. Kassirer, Meghan Ramsey and John P. A. Ioannidis (2015) Inadequacies of physical examination as a cause of medical errors and adverse events: A collection of vignettes. *American Journal of Medicine* 128 (12): 1322–1324. <https://doi.org/10.1016/j.amjmed.2015.06.004>

Wilson, Ross M., Beth T. Harrison, Robert W. Gibberd and J. Devance Hamilton (1999) An analysis of the causes of adverse events from the Quality in Australian Health Care Study. *Medical Journal of Australia* 170 (9): 411-415. <https://doi.org/10.5694/j.1326-5377.1999.tb127814.x>

Claudio Lucchiari is Associate Professor at the Department of Philosophy of the University of Milan, where he teaches Cognitive Psychology and Psychology of Negotiation. He has a degree in Cognitive Psychology and a PhD in Communication Psychology. He has authored five books, more than 50 peer-reviewed articles and several contributions in conference proceedings and books. His research activities focus on cognitive and neurophysiological aspects of health management, decision making, negotiation and communication. He is a member of the European Society of Affective Neuroscience. Address for correspondence: Università degli Studi di Milano, Dipartimento di Filosofia, via Festa del Perdono 7, 20122 Milano. Email: claudio.lucchiari@unimi.it

Maria Elide Vanutelli has a master's degree in neuropsychology and a PhD in cognitive psychology. She currently has a post-doctoral position at the University of Milan. Her main research interests concern cognitive and social neuroscience, with a specific focus on learning, creativity and emotions. She is also working on a research activity aimed at developing technological tools to support special needs in education settings. She has authored more than 30 scientific papers. Address for correspondence: Università degli Studi di Milano, Dipartimento di Filosofia, via Festa del Perdono 7, 20122 Milano. Email: maria.vanutelli@unimi.it

Raffaella Folgieri has a PhD in computer science and is Assistant Professor and Researcher in Computer Skills at the Philosophy Department of the University of Milan. She teaches on artificial intelligence and intelligence systems as part of the university's postgraduate degrees in psychiatry and philosophy respectively, as well as virtual reality for the bachelor's degree in information technology and digital communication. She is a member of the Italian Society of Engineering and of SIREN (Italian Neural Networks Society). Her research focuses on artificial intelligence and cognitive science. Address for correspondence: Università degli Studi di Milano, Dipartimento di Filosofia, Via Festa del Perdono 7, 20122 Milano. Email: raffaella.folgieri@unimi.it