Available online at www.sciencedirect.com

ScienceDirect

Health Professions Education 3 (2017) 15–25

www.elsevier.com/locate/hpe

Reflection in Medical Diagnosis: A Literature Review

Silvia Mamede*, Henk G. Schmidt

Institute of Medical Education Research, Erasmus Medical Center and Department of Psychology, Erasmus University Rotterdam, The Netherlands

Received 15 January 2017; accepted 16 January 2017
Available online 20 January 2017

Abstract

Purpose: Reflection in medical diagnosis has been said to prevent errors by minimizing flaws in clinical reasoning. This claim, however, has been much disputed. While some studies show reflective reasoning to improve diagnostic performance, others find it to add nothing. This paper presents a narrative review of the literature on reflection in medical diagnosis aimed at addressing two questions: (1) how reflective reasoning has been conceived in this literature; and (2) what is the effect of different forms of reflective reasoning on diagnostic performance.
Method: We searched PubMed and Web of Science for papers published until June 2016 and identified additional literature through the list of references from the initial publications. By building upon dual-processing theories of reasoning, we classified the empirical studies according to two dimensions: (1) the phase of the diagnostic process in which reflection was applied, and (2) the type of reasoning instructions provided to participants.
Results: We identified 46 papers for full review, 31 of them reporting on empirical studies. Different conceptualizations of reflective reasoning exist in the literature. In 16 studies, reflective reasoning was triggered to verify previously generated diagnosis, usually (13/16 studies) by following specific reasoning instructions. Participants were requested to reflect for generating diagnostic hypothesis in 4 studies, all using specific instructions. In 8 studies, 2 of them employing specific instructions, reflection was assumed as taking place throughout the diagnostic process. Reflective reasoning positively affected diagnostic performance when conceived as a process of examining the grounds of initial diagnoses generated through intuitive judgment. The benefits of reflection were particularly substantial when physicians were provided with specific reasoning instructions that led them to be confronted with evidence from the case. Studies using other forms of reflection led to contradictory findings.
Discussion: Reflective reasoning can be a powerful tool to reduce diagnostic errors and increase diagnostic performance. For this to happen, reflection should be triggered for diagnosis verification and needs to interfere with initial diagnostic reasoning, which requires confrontation with evidence from the case.

Keywords: Reflection; Reflective reasoning; Analytic reasoning; Medical diagnosis; Diagnostic error

*Correspondence to: Institute of Medical Education Research Rotterdam, Erasmus MC, Ae building, room 242, Wytemaweg 80, Rotterdam 3015, CN, The Netherlands. Fax: +31 10 7044752.

Peer review under responsibility of AMEEMR: the Association for Medical Education in the Eastern Mediterranean Region

http://dx.doi.org/10.1016/j.hpe.2017.01.003

Contents

Introduction

Anecdotes about clinicians who diagnose a patient's problem in an instant are common in clinical settings and have always enchanted students and medical staff alike. No matter how much fascination diagnosis in the blink of an eye can hold, physicians are, however, usually recommended to stay away from it. Clinical teachers advise medical students to "do not jump to conclusions" before completing a comprehensive process of gathering information from the patient. And the recommendation does not apply only to novice diagnosticians. Practicing physicians have been warned as well about the perils of relying on fast, intuitive diagnostic reasoning, which has often been appointed as a source of diagnostic errors.[1,2] By reflecting upon the case and carefully considering all available information, it is said, physicians would avoid reasoning flaws that underlie most mistakes. Such a claim, however, is far from consensual.[3,4] Indeed, while some studies have shown reflection to improve diagnoses,[5,6] others have found no benefits of a more analytic relative to an intuitive reasoning approach.[7,8] These studies differ not only in their findings. They also seem to be referring to different sorts of reflection. What reflection in medical diagnosis entails and how it affects the quality of physicians' decisions are still to be determined. We will explore these questions by reviewing the existing literature in particular on empirical research that has employed different forms of reflective diagnostic reasoning to study their impact on diagnostic performance.

It has long been known that physicians can – and usually do – generate diagnoses in the first minutes of a clinical encounter. They do so thanks to "pattern-recognition", a fast, largely unconscious, recognition of similarities between the case at hand and illness scripts that the physician has stored in memory, either in the form of prototypes of diseases or examples of previously seen patients.[9] Illness scripts, when activated by cues in the patient's history, lead to generation of one or a few diagnostic hypotheses and guide the subsequent search for additional information to either confirm or refute the hypotheses.[10] Pattern-recognition tends to take place largely unconsciously, and physicians are only aware of its outcome, that is, the diagnostic hypothesis. Verifying this diagnostic hypothesis requires analytic reasoning to match the elements of the illness script with findings of the case at hand.

The two modes of reasoning tend therefore to be involved in diagnosis making. Nevertheless, it is the intuitive reasoning that has been considered the hallmark of expertise and has traditionally caught much of researchers' attention. Only recently, the role of reflection in the diagnostic process has started to attract interest, which has possibly been nurtured by two factors. First, an increasing awareness of the problem of diagnostic error and its adverse consequences for patients. The magnitude of the problem has been shown to be high, and the literature has associated errors with relying on intuitive judgments that are not sufficiently examined through analytic reasoning.[11,12] Second, the prominence achieved by dual-processing theories of reasoning and decision-making in the psychology literature.[13] Dual-processing theories have been frequently applied to clinical reasoning and diagnostic error in the medical literature[1,14] and these models, as we discuss below, tend to value the role of reflection.

Briefly, dual-processing theories distinguish between two principal types of reasoning. System 1 processes (sometimes labeled 'Type 1', 'intuitive', 'implicit' or 'heuristic') are unconscious, fast, automatic, and do not

suffer from limitations of working-memory, whereas System 2 processes (also named 'Type 2', 'analytic', 'explicit' or 'reflective') are conscious, slow, deliberative, and restricted by working memory capacity.[13] System 1, largely based on prior experience, operates through holistic recognition of a situation as of a kind encountered previously, which retrieves from memory a schema that brings a 'solution'. System 2, on the other hand, depends on application of rules (for example, the rules of diagnosis that associate certain symptoms with the likelihood of a particular disease) and is therefore the type of reasoning that allows for hypothetical thinking.

One of the most influential dual-process theories (for reviews, see 13,15) the 'heuristic-analytic' theory, advocates that the two reasoning modes are interdependent and sequential: when we are confronted with a problem, preconscious heuristic processes provide default responses that may or may not be altered by analytic reasoning.[16] Very roughly, what happens is that System 1 selects relevant aspects of presented information, cueing a mental model of the problem that leads to a response (for example, a diagnostic hypothesis in a clinical problem). System 2 may or may not intervene to revise or replace the mental model of the problem and the response that comes with it (in a clinical problem, System 2 intervention would for example lead to recognition of contradictory findings, bringing an alternative diagnosis under consideration). As our processing capacity is limited, the theory assumes that we tend to generate only one mental model of the problem at a time. Moreover, we have a universal tendency to satisfice with a plausible enough model, unless we have good reasons to discard it. That is, reflective reasoning is not our basic mode of thought and in many situations it does not intervene, which means that we simply go with responses triggered by heuristic processes. Whether and to what extent System 2 comes into action depends on several factors, but two of them have emerged from psychological research as critical: availability of time for this more effortful mode of reasoning, and strong, elaborate, instructions requiring deductive reasoning.[16,17]

Dual-processing theories have been traditionally investigated in experiments with reasoning problems very different from medical diagnosis, but it should not come as a surprise that they have become so popular in the literature on diagnostic error. Their account seems to match the diagnostic process quite well: pattern-recognition (System 1) triggered by a patient's cues leads to generation of a hypothesis, and reflective reasoning (System 2) enters into action to verify it by checking whether the patient's findings indeed fit with what would be expected were the initial hypothesis correct. Failure to engage in reflection for an appropriate verification of initial hypotheses would make physicians prone to fall prey of cognitive biases and premature closure.[1,2,14] This account of diagnostic error abounds in the literature but has been questioned by authors who point to its lack of empirical support.[3,4] Indeed, while additional thought on to-be-diagnosed problems led to substantial improvements in some studies[5,6] it did not bring any benefit in others.[8,18] Whether reasoning modes affect the quality of diagnoses has therefore been a source of much discussion, and it may be difficult to make sense of the existing research.

In the present review of the literature, we intended to contribute to clarify this controversy by addressing two questions. First, how reflective diagnostic reasoning has been conceived in the literature? Different studies seem to have employed very different instructions to induce their participants to reflect upon to-be-solved problems, an indication that distinct views of the nature of reflective reasoning exist. Second, what is the effect of reflection during the reasoning process on its outcomes? At first glance, the literature shows contradictory findings, leaving it unclear if the accuracy of diagnoses is affected by which reasoning mode produced them. As different constructs of reflective diagnostic reasoning seem to exist, it may be worth exploring whether and how its effects and its conceived nature are associated.

We approached these two questions from the perspective of the dual-processing theories of reasoning applied to medical diagnosis. In the next section, we will describe how we used this theoretical framework to categorize the papers identified in our search and to summarize their findings.

Methods

We searched the PubMed and the Web of Science databases for papers published until June 2016 which contained the terms "reflection"; "diagnostic reasoning"; "reflective reasoning"; "critical thinking"; "analytic reasoning"; "non-analytic reasoning"; "pattern-recognition"; and "medical diagnosis" in the title or in the abstract. The search query for the Web of Science search was the following: (("reflective reasoning" OR "analytic reasoning" OR "non-analytic reasoning" OR ("pattern recognition" NEAR reasoning) OR "critical thinking" OR (reflection NEAR diagnosis)) AND ("diagnostic reasoning" OR (medic* AND

diagnos*))), and the search strategy for the PubMed was modeled on it. We limited our search to English papers indexed as articles, proceeding papers or reviews.

This database search led to 144 papers. Our review questions focused on the nature of physicians' reflective diagnostic reasoning and its effect on their diagnostic performance. The two authors first read the abstracts of the papers to identify the potentially relevant ones. Papers dealing with topics beyond the scope of our review questions, such as the development of assessment tools or the effectiveness of educational interventions to teach diagnostic reasoning, were excluded. Thirty-five papers remained, and their full text was obtained for full review. The database search was then supplemented by manually reviewing the bibliographies of the 35 articles and the authors' personal archives. Eleven additional papers were identified, leading to a total of 46 papers to be fully reviewed.

We built upon dual-process theories as applied to medical diagnosis to categorize these papers according to two dimensions. The first was the phase of the diagnostic process to which reflection was applied. According to the heuristic-analytic theory, implicit reasoning generates hypothesis and it is the work of analytic reasoning to exam them. Based on this theory, we distinguished between studies that requested participants to engage in reflective reasoning either (i) during generation of diagnostic hypotheses, (ii) during verification of diagnostic hypotheses, or (ii) throughout the diagnostic process. The second dimension that we took into consideration was the type of reasoning instructions used in the studies, if any, as this factor has been shown to influence System 2 intervention in psychological research.[16,17] We distinguished between studies that provided participants with (i) specific reasoning instructions, (ii) general instructions (e.g. "be careful and consider all the data"), or (iii) no reasoning instructions.

All empirical papers were categorized according to these categories and their findings were summarized. The first 10 papers in the list of references were independently categorized by the two authors, and as they agreed in all of them, the categorization proceed with one single author.

Results

Thirty-one out of the 46 papers reviewed reported on empirical studies, 25 of which investigated the influence of different reasoning approaches on diagnostic performance either of students or physicians.[5–8,18–36] The remaining 15 articles consisted of theoretical papers that reviewed the literature on reasoning modes employed by physicians during the diagnostic process and their influence on the quality of diagnoses.[1,3,9,14,37–46] Interesting to notice, 12 out of these 15 papers addressed diagnostic reasoning in the light of dual-processing theories, which testifies for the influence that these theories have acquired in research on clinical reasoning.

Applying the two dimensions adopted to categorize the papers – the phase of the diagnostic process in which reflection was triggered and the type of instruction provided to participants – led to the matrix displayed in Table 1. To summarize the results of the review, we will group the papers according to the phase in which reflection was requested. Within each group of papers, we will then distinguish between the different types of instructions employed and summarize the existing empirical findings.

Reflection for the generation of diagnostic hypothesis

Four[23,29,35,36] out of the 31 empirical papers required participants to engage in reflective reasoning for the generation of initial diagnostic hypotheses. All studies provided participants with specific reasoning instructions on how to do this.

The typical approach adopted in these studies may be exemplified by the procedure used by Kulatunga–Moruzi and collaborators to compare the effects of intuitive and analytic reasoning strategies on novice students' diagnostic performance on dermatological cases.[23] Participants were first exposed, during a training phase, to a set of diseases and their key features, usually provided in a response sheet that lists the features relevant for the diagnoses of the cases used in the study. The training phase aimed at generating familiarity with the diseases, allowing for subsequent diagnosis to be made on the basis of similarity to the cases previously seen. Immediately after the training, students were requested to diagnose new cases either by providing a quick diagnosis based on similarity or by following instructions for analytic reasoning. The latter asked students to first identify, with the aid of the list of possible relevant clinical features, which features were present in the to-be-diagnosed case and only after that give a diagnosis. A similar approach was employed in subsequent studies on novice students' performance while diagnosing ECGs, but the comparisons then involved not only this 'reflection / features-first' approach and the similarity-based diagnosis.[35,36] A 'combined approach' was included, which requested

Table 1
Studies according to the phase of the diagnostic process in which reflection was triggered and the type of instructions employed to activate reflection.

	Specific reasoning instructions provided	Non-specific ('Be careful. Consider all the data')	NO instructions provided	Undefined or non-applicable
Reflection for the generation of diagnostic hypothesis	Ark et al. 35* Ark et al. 36* Kulatunga-Moruzi et al. 23 Regehr et al. 29			
Reflection for the verification of diagnostic hypothesis	Ark et al. 35*	Monteiro et al. 28	Jaimes et al. 22	
	Ark et al. 36* Mamede et al. 5 Mamede et al. 26 Mamede et al. 24 Schmidt et al. 6 Mamede et al. 24 Sibbald et al. 32 Sibbald et al. 33 Groves et al. 47 Groves et al. 48 Shimizu et al. 30 Sibbald et al. 34	Sibbald et al. 31		
Reflection throughout the diagnostic process	Ilgen et al. 7	Sherbino et al. 8	Coderre et al. 19	
	Ilgen et al. 21	Norman et al. 18	Mamede et al. 50 Hess et al. 20 Mamede et al. 25	
Undefined or non-applicable				Dong et al. 51 Durning et al. 52 Durning et al. 53 Michelsen et al. 54 Stolper et al. 55

*The studies compared two reasoning modes and they are therefore included in both categories.

students to first generate an initial diagnosis based on similarity and subsequently check if it was correct by going through the provided list of features. In the 'combined approach', reflection was therefore triggered only for the verification of initial diagnoses by instructions that compelled participants to deductive reasoning. These papers are therefore also included in the subsequent section.

What have these studies shown about the effects of reflection on diagnostic performance? The 'reflection / features-first' reasoning approach did not increase accuracy relative to simply request students to diagnose dermatological cases or ECGs based on similarity to previously seen ones (i.e., by pattern-recognition).[23,35,36] This finding replicated what had been observed with medical residents in a previous study.[29] However, the 'combined approach' proved more effective than both 'reflection/features-first' and pattern-recognition in the two studies that compared the influence of these four approaches on students' performance while diagnosing ECGs.[35,36]

These studies suggest that no benefit is to be expected from reflecting for generating initial diagnosis, at least not if this means firstly identifying all features that are possibly relevant in the case. While explaining the poor results of the reflection/features-first approach, the authors argued that trying to list features unguided by a diagnostic hypothesis exposes the diagnostician to the risk of being led astray by an incoherent list of features that cannot be combined into

a diagnosis.[35] Their argument is supported by their finding that participants required to first list features identified more features indicative of incorrect diagnoses than did participants in the other conditions.

Notice that the instructions used for the reflection/features-first approach in these studies suggest that reflective reasoning has been conceptualized as a sort of categorization task based on prototype theories, requiring going through lists of features to match existing features with those belonging to alternative categories. Differently, the instructions for the 'combined-approach' seem consistent with a view of reflection shared by the studies discussed in the following section.

Reflection for the verification of diagnostic hypothesis

Out of the 31 empirical studies, 16 induced participants to reflect primarily for the verification of their initial diagnostic hypotheses. Participants were requested to follow specific reasoning instructions in 13 of these studies,[5,6,24,26,27,30,32–36,47,48] whereas 2 studies provided general instructions[28,31] and in 1 study no instruction was given.[22]

When participants were provided with specific reasoning instructions, two main approaches have been adopted: a step-by-step procedure to check predictions of initial diagnostic hypothesis against data from the case and search for contradictory evidence, and the use of checklists for diagnosis verification. The former can be exemplified by a procedure used in several studies with internal medicine residents.[5,6,24,26,27] Participants were requested to (1) read the case and write down the most likely diagnosis as fast as possible but without compromising accuracy; (2) list the findings in the case description that support this diagnosis, the findings that speak against it, and the findings that would be expected to be present if this diagnosis were true but that were not described in the case; (3) list alternative diagnoses if the initial diagnosis generated for the case had proved to be incorrect and proceed with the same analysis for each alternative diagnosis; and (4) draw a conclusion on their diagnosis for the case.

This procedure has been employed not only to compare the effect of intuitive and reflective reasoning on diagnostic performance[5,24,26] but also to explore whether reflective reasoning can counteract availability bias, which causes physicians to misdiagnose a similar-looking, but different, case as the disease seen before.[6,27] A similar procedure was also adopted in two studies that intended to explore the characteristics of experienced physicians' reasoning and the types of errors made at different levels of clinical expertise by distinguishing between failures in hypothesis generation or in identification or interpretation of relevant information.[47,48]

The second approach adopted when specific reasoning instructions were provided requested participants to reflect upon their first diagnosis for a clinical problem guided by a checklist.[30,32–36] For example, a checklist containing key variables in ECG interpretation (e.g., chamber hypertrophy, abnormal intervals) was used in a study that compared performance when participants from different levels of expertise interpreted sets of ECGs either as they normally would do or by using the checklist for the verification of their initial diagnoses.[34] Checklists have been also used in studies with non-visual materials. Sibbald et al. requested internal medicine residents to exam a high-fidelity cardiopulmonary simulator, provide a diagnosis, and subsequently verify their initial diagnoses by using a checklist with the major aspects of a cardiac exam.[32] To determine whether the verification alone improved diagnostic accuracy or whether this benefit depended on the possibility of collecting additional information (or re-assessing accuracy of initially collected one), they manipulated permission to re-examine the simulator.

A distinction between two types of checklist is noteworthy. The studies by Sibbald et al. have employed checklists specific to the problem domain. Their items direct the user to re-check features of the case (for example, in the study of physical cardiac exam, 'murmur radiation?'; 'extra sounds?'). However, general 'debiasing checklists' have also been suggested to prevent errors.[30,49] They consist of questions about the user's reasoning process ('did I consider the inherent flaws of heuristic thinking'? or 'was my judgment affected by any other bias?'). A study by Shimizu et al. compared the effect of these two types of checklists on medical students' diagnostic performance.[30]

In the aforementioned studies, participants followed a procedure that specified how to reason through the problem to verify their initial diagnosis, but specific instructions such as those have not always been provided. In three studies,[22,28,31] participants were requested to diagnose cases and, subsequently, diagnose the case again. In this second pass, they were simply asked to go through all the cases again or provided with a general instruction such as 'carefully reconsider every diagnosis'.[28]

The studies within this category seem to be built upon a view of reflection as a process of scrutinizing the grounds of a previously made intuitive judgment to either confirm or discard it, in line therefore with the

heuristic-analytic dual-process theory. Nevertheless, the studies differ in their assumption about how reflection can be induced. Whereas some studies employed strong deductive reasoning instructions, others seem to have assumed that reflective reasoning can be put into action simply by requesting participants to review their initial diagnoses.

The findings of these studies suggest this distinction between reflection triggers to be worth making. When reflection upon initial diagnosis was activated by specific reasoning instructions, it led to a substantial increase in diagnostic accuracy relative to diagnosis made through intuitive reasoning. In two studies with internal medicine residents, for instance, the gain in accuracy after participants reflected upon their initial diagnoses for complex cases was close to 40% or even higher.[5,24] Deliberate reflection also counteracted the negative consequences of availability bias. When physicians made wrong initial diagnoses because they confounded the to-be-diagnosed case with a disease seen before, reflection restored accuracy to the levels observed on cases that were not subject to bias.[6,27] The performance of novice students[35] and expert physicians[33] in interpreting ECGs also improved substantially when they revised their initial diagnoses by using a checklist. A specific diagnostic checklist was also useful for medical students to revise their initial diagnoses for complex internal medicine cases,[30] leading to gains in accuracy similar to those observed when residents engaged in deliberate reflection in other studies.[5,24]

Reflection seemed to be also beneficial when participants were simply requested to review their initial diagnoses without the provision of reasoning instructions, but the findings were not so clear-cut, and gains, when present, were less impressive. Dermatologists, for instance, improved their initial diagnoses on average in around 10% when requested to review each of them.[22] And when physicians were left to decide by themselves whether they would review their initial diagnoses or not, this choice was made for only 8% of the diagnoses despite the fact that the accuracy was around 60%.[28] The scores for these revised diagnoses increased substantially (from 0.64 to 0.90), but as the option for reconsidering was so rarely made, reflection led to a minimal (from 1.20/2 to 1.22/2), though significant, increase in overall accuracy scores relative to intuitive diagnosis. That elaborate reasoning instructions may play a crucial role has been also shown by a study on the use of checklists in ECGs interpretation. Verification of initial diagnoses only led to gains in accuracy when guided by a checklist that directed physicians' attention to key features of ECGs; verification alone did not make a difference.[33]

The findings of the studies within this category bring also some insight on how reflection helps improve initial diagnoses. Two studies on types of errors in clinical reasoning suggest that most of experienced physicians' mistakes derive from failures in identification and/or interpretation of relevant clinical findings.[47,48] In Sibbald's study with the cardiopulmonary simulator, when physicians reflected upon their initial diagnoses without being allowed to go back to the case to review its features, reflection did not help.[32] However, physicians compelled to go through the case again were able to correct initial mistakes,[5,24,32] even when this required overcoming the influence of contextual information that had originally biased reasoning.[6,27] Taken together, these studies suggest that reflective reasoning helps by leading physicians to identify relevant features that had previously passed unnoticed or to re-interpret case findings. Recognizing critical features that were initially overlooked seems to be only possible when the diagnostician is confronted with evidence from the case by looking at the problem rather than at one's reasoning,[30] and for those who possess relevant knowledge, as suggested by the students' failure to benefit from reflection on cases that were far beyond their reach.[24]

Reflection throughout the diagnostic process

In 8 out of the 31 empirical studies, participants were requested to reflect (or considered to have reflected) throughout the whole diagnostic process, i.e., both during generation of diagnostic hypothesis and during its verification.[7,8,18–21,25,50] This category comprises studies that explicitly requested participants to reflect upon to-be-solved cases as well as studies that simply examined the relationship between time on diagnosis and performance. Despite different in their design, these studies seem to share a conceptualization of reflection as taking place throughout the whole diagnostic process, deviating therefore from the heuristic-analytic dual process theory.

In one of these studies, reflection was triggered by simply requesting participants to "consider all the data" to make a diagnosis, not in a second pass that would review previously made diagnoses but from the start of the diagnostic process.[18] Two other studies provided specific reasoning strategies[7,21] that led participants through a step-by-step process: first providing a summary of the problem representation, then listing all the diagnoses considered for the case and indicating

findings in favour and against each diagnosis before making a decision. (Notice that the first two steps already involve considerable reflection) Finally, four papers that referred to reflection throughout the whole diagnostic process did not provide any reasoning instruction.[19,20,25,50] This happened, for instance, in an observational study of response data of a large group of residents who took the American Board of Internal Medicine certification exam.[20] Time spent on initial responses to questions that consisted of diagnosing clinical vignettes and changes in initial responses were taken as indication of reflection, and relationship with diagnostic accuracy was evaluated. No reasoning instructions were provided also in experimental studies that aimed at investigating factors that induce physicians to move from a more intuitive to reflective reasoning, such as case ambiguity or physicians' perception of the complexity of the to-be-solved problem.[25,50]

Among the studies within this category, those that explored the influence of reasoning mode on diagnostic accuracy led to contradictory findings. Ilgen and colleagues applied the same basic set up in two experiments and yet reached different conclusions: while reflection led to higher diagnostic accuracy relative to intuitive reasoning in the first experiment,[21] no difference between conditions was found in the second one.[7] The authors attributed this contradiction to slight methodological differences that may have led participants to apply reflection primarily for diagnosis verification in the first experiment, which would have allowed them to gain from it.[7]

When time spent in diagnosis was employed as a proxy for reflection, findings are again not consistent. Reflection triggered by providing information that the to-be-solved cases were extremely complex improved diagnoses in 27% relative to intuitive reasoning.[25] This finding conflicts with the results of two studies by the same research group, which showed diagnostic accuracy to be similar under conditions that encouraged residents to be fast or instructed them to take as much time as they wished to consider all the data.[8,18] The latter group indeed took longer on average to diagnose each clinical vignette though the time difference – around 20 s – seems too small to allow for the authors' consideration of this diagnostic process as "careful, thorough and reflective". Nevertheless, as cases diagnosed incorrectly took longer than those diagnosed correctly, the authors suggested that adopting a more analytic reasoning, consequently spending more time, is not beneficial. Indeed the relationship between time on diagnosis and accuracy seems to be far more complex, as suggested by a large study of residents' diagnostic responses in a high-stake exam.[20] Less time spent on diagnosis was also associated with more correct responses but only for low-complexity clinical vignettes and for low-ability residents. For complex cases, spending more time was associated with more correct responses, and residents with high cognitive ability benefitted from reflection in both easy and complex clinical cases. The study suggests that a certain level of competence is required for reflection to help, in line with previous findings showing students to fail to benefit from reflection when cases were extremely complex to them.[24]

Discussion

Interest in reflective reasoning in medical diagnosis has grown together with concerns about the problem of diagnostic error and its adverse consequences for patients. The increasing influence of dual-process theories in psychological research has nurtured this interest, but whether reflection has a role to play in minimizing diagnostic errors has been a source of much debate. We reviewed the literature to examine how reflective diagnostic reasoning has been conceived in the studies on medical diagnosis and whether it affects diagnostic performance. Research on the theme is in its first years, and we identified a limited number of studies, but a large fraction of them consists of empirical research. These empirical studies were categorized according to two dimensions: the phase of the diagnostic process in which reflection has been mobilized and the type of instruction (if any) employed to trigger reflection. Empirical findings were summarized, and a number of conclusions could be made.

First, the moment of the diagnostic process in which reflection is activated, which reveals different conceptualizations of reflective reasoning, apparently determines whether reflection affects diagnostic performance. When reflection was brought into action in the phase of diagnosis verification, it substantially increased diagnostic accuracy relative to non-analytic diagnostic reasoning. In these studies, reflection is conceived as a deliberate consideration of a diagnosis previously generated through intuitive reasoning to examine the grounds that support it. This view is in line with the heuristic-analytic dual process theory, which advocates that System 2 intervenes to revise the mental representation of the problem and the solution for the problem previously generated by System 1. Reflection does not seem to help when it is mobilized only for the generation of a diagnostic hypothesis, at least if it consists of identifying all features in a case before

considering a diagnosis, and the effects of reflection are not so clear when it is requested to take place throughout the whole diagnostic process.

The literature allows for a second conclusion. Reflection tends to be much more effective with the aid of a procedure that compels physicians to search for flaws in their initial judgments. A close look into the findings of the studies that induced reflection for diagnosis verification supports this statement. Studies that employed elaborate instructions, either in the form of a step-by-step procedure[5,6,24,27] or of checklists,[30,32–34] showed a much larger effect of reflection on accuracy than the studies that provided only a general instruction[28] or no guidance for the verification.[22] Moreover, a study that compared the benefits of verifying initial diagnoses of ECGs with or without guidance found that verification only improved previous decisions when guided by instructions on what to check for.[33]

It seems clear that, in line with what has been found in psychological research, elaborate reasoning instructions optimize the potential of reflection to improve diagnosis. What is not clear is why this is so. It may be simply because the instructions increase the frequency with which physicians opt for checking the grounds of their initial judgments about the case. Physicians, as anybody else, do not tend to adopt reflective reasoning as their routine mode of thought. Indeed, when given the possibility of reviewing their initial diagnoses, they rarely opted to do so even if their diagnostic performance was far from optimal.[20,28] Provided with a list of items to be checked, they would tend to skip the verification less frequently. It may also be, however, that the nature of the instructions plays a crucial role by favouring restructuring of initial reasoning. We all have a natural tendency to satisfice with our initial explanation for a problem and stick to it unless a reason emerges which is strong enough to convince us that our explanation should be abandoned.[13,16] Physicians will also tend to go with their initial, intuitive judgments unless confrontation with sufficient conflicting evidence compels them to change their mind. When specific instructions were used to guide reflection upon initial diagnoses, they requested physicians to go through the case again and search for evidence that contradicted initial judgments or check specific, potentially relevant features. This process possibly redirects physicians' attention to features in the case that may have been initially overlooked or leads to re-interpretation of case features, activating other illness scripts and bringing other diagnostic hypothesis to mind. For reflection to help, it apparently has to interfere with the original diagnostic reasoning. Without being confronted with contradictory evidence, initial reasoning may remain untouched. If this is correct, it may help to explain the absence of any effect of the few extra seconds spent in diagnosing the case by participants under the 'analytic condition' in studies on the relationship between time and accuracy.[8,18] While spending more time is certainly a requirement for diagnostic reflective reasoning to take place, time per se may not be sufficient. It may also be a reason why a checklist to guide thinking about the problem itself improved students' diagnoses whereas one requiring thinking about one's own reasoning produced no gains.[30]

Finally, the findings of the studies suggest that reflection may be a powerful tool to improve diagnoses, but its potential is affected by an interaction between the to-be-solved problem and the diagnostician. Reflection helps when there is floor for mistakes. When cases were so straightforward that intuitive judgment led to accuracy scores close to the maximum, reflecting upon the initial diagnosis did not add much.[5,24,30] On the other hand, reflection can only act by mobilizing existing knowledge, and therefore nothing was gained from reflecting upon cases that were far beyond the diagnostician's expertise.[24]

To sum up, the literature is scarce but the existing empirical studies support the claim that reflective reasoning in medical diagnosis improves accuracy and can be a tool to minimize errors. This only seems to happen, however, when reflective reasoning is conceived as reflection upon initial diagnosis generated through more intuitive reasoning and involves scrutinizing its grounds through confrontation with evidence from the case. Although the literature provides insights on factors that influence the benefits of reflection for diagnostic reasoning, further research is required to clarify the circumstances under which reflection helps and how its potential can be optimized.

Ethical approval

Not applicable.

Funding

None.

Other disclosures

None.

References

1. Croskerry P. From mindless to mindful practice–cognitive bias and clinical decision making. *N Engl J Med* 2013;368(26): 2445–2448.
2. Redelmeier DA. Improving patient care. The cognitive psychology of missed diagnoses. *Ann Intern Med* 2005;142(2):115–120.
3. Norman G. Dual processing and diagnostic errors. *Adv Health Sci Educ Theory Pract* 2009;14(Suppl 1):37–49.
4. Norman GR, Eva KW. Diagnostic error and clinical reasoning. *Med Educ* 2010;44(1):94–100.
5. Mamede S, Schmidt HG, Penaforte JC. Effects of reflective practice on the accuracy of medical diagnoses. *Med Educ* 2008;42(5):468–475.
6. Schmidt HG, Mamede S, van den Berge K, van Gog T, van Saase JLCM. Rikers RMJP. Exposure to media information about a disease can cause doctors to misdiagnose similar-looking clinical cases. *Acad Med* 2014;89(2):285–291.
7. Ilgen JS, Bowen JL, McIntyre LA, Banh KV, Barnes D, Coates, WC, et al. Comparing diagnostic performance and the utility of clinical vignette-based assessment under testing conditions designed to encourage either automatic or analytic thought. *Acad Med* 2013;88(10):1545–1551.
8. Sherbino J, Dore KL, Wood TJ, Young ME, Gaissmaier W, Kreuger, S, et al. The relationship between response time and diagnostic accuracy. *Acad Med* 2012;87(6):785–791.
9. Norman GR, Brooks LR. The non-analytical basis of clinical reasoning. *Adv Health Sci Educ Theory Pract* 1997;2(2): 173–184.
10. Schmidt HG, Norman GR, Boshuizen HPA. A cognitive perspective on medical expertise – theory and implications. *Acad Med* 1990;65(10):611–621.
11. Graber ML. The incidence of diagnostic error in medicine. *BMJ Qual Saf* 2013;22(Suppl 2):ii21–ii27.
12. Singh H, Meyer AN, Thomas EJ. The frequency of diagnostic errors in outpatient care: estimations from three large observational studies involving US adult populations. *BMJ Qual Saf* 2014;23(9):727–731.
13. Evans JSBT. Dual-processing accounts of reasoning, judgment, and social cognition. *Annu Rev Psychol* 2008;59:255–278.
14. Croskerry P, Singhal G, Mamede S. Cognitive debiasing 1: origins of bias and theory of debiasing. *BMJ Qual Saf* 2013;22 (Suppl 2):ii58–ii64.
15. Osman M. An evaluation of dual-process theories of reasoning. *Psychon Bull Rev* 2004;11(6):988–1010.
16. Evans JST. The heuristic-analytic theory of reasoning: extension and evaluation. *Psychon Bull Rev* 2006;13(3):378–395.
17. Evans JST. Logic and human reasoning: an assessment of the deduction paradigm. *Psychol Bull* 2002;128(6):978–996.
18. Norman G, Sherbino J, Dore K, Wood T, Young M, Gaissmaier, W, et al. The etiology of diagnostic errors: a controlled trial of system 1 versus system 2 reasoning. *Acad Med* 2014;89(2): 277–284.
19. Coderre S, Mandin H, Harasym PH, Fick GH. Diagnostic reasoning strategies and diagnostic success. *Med Educ* 2003;37 (8):695–703.
20. Hess BJ, Lipner RS, Thompson V, Holmboe ES, Graber ML. Blink or think: can further reflection improve initial diagnostic impressions? *Acad Med* 2015;90(1):112–118.
21. Ilgen JS, Bowen JL, Yarris LM, Fu R, Lowe RA, Eva K. Adjusting our lens: can developmental differences in diagnostic reasoning be harnessed to improve health professional and trainee assessment?. *Acad Emerg Med* 2011;18(Suppl 2): S79–S86.
22. Jaimes N, Dusza SW, Quigley EA, Braun RP, Puig S, Malvehy, J, et al. Influence of time on dermoscopic diagnosis and management. *Australas J Dermatol* 2013;54(2):96–104.
23. Kulatunga-Moruzi C, Brooks LR, Norman GR. Coordination of analytic and similarity-based processing strategies and expertise in dermatological diagnosis. *Teach Learn Med* 2001;13(2): 110–116.
24. Mamede S, Schmidt HG, Rikers RM, Custers EJ, Splinter TA, van Saase JL. Conscious thought beats deliberation without attention in diagnostic decision-making: at least when you are an expert. *Psychol Res* 2010;74(6):586–592.
25. Mamede S, Schmidt HG, Rikers RMJP, Penaforte JC, Coelho-Filho JM. Influence of perceived difficulty of cases on physicians' diagnostic reasoning. *Acad Med* 2008;83(12):1210–1216.
26. Mamede S, Splinter TA, van Gog T, Rikers RM, Schmidt HG. Exploring the role of salient distracting clinical features in the emergence of diagnostic errors and the mechanisms through which reflection counteracts mistakes. *BMJ Qual Saf* 2012;21(4): 295–300.
27. Mamede S, van Gog T, van den Berge K, Rikers RMJP, van Saase JLCM, van Guldener, C, et al. Effect of availability bias and reflective reasoning on diagnostic accuracy among internal medicine residents. *JAMA-J Am Med Assoc* 2010;304(11): 1198–1203.
28. Monteiro SD, Sherbino J, Patel A, Mazzetti I, Norman GR, Howey E. Reflecting on diagnostic errors: taking a second look is not enough. *J Gen Intern Med* 2015;30(9):1270–1274.
29. Regehr G, Cline J, Norman GR, Brooks L. Effect of processing strategy on diagnostic skill in dermatology. *Acad Med* 1994;69 (10 Suppl):S34–S36.
30. Shimizu T, Matsumoto K, Tokuda Y. Effects of the use of differential diagnosis checklist and general de-biasing checklist on diagnostic performance in comparison to intuitive diagnosis. *Med Teach* 2013;35(6):E1218–E1229.
31. Sibbald M, de Bruin AB. Feasibility of self-reflection as a tool to balance clinical reasoning strategies. *Adv Health Sci Educ Theory Pract* 2012;17(3):419–429.
32. Sibbald M, de Bruin AB, Cavalcanti RB, van Merrienboer JJ. Do you have to re-examine to reconsider your diagnosis? Checklists and cardiac exam. *BMJ Qual Saf* 2013;22(4):333–338.
33. Sibbald M, de Bruin AB, van Merrienboer JJ. Checklists improve experts' diagnostic decisions. *Med Educ* 2013;47(3):301–308.
34. Sibbald M, De Bruin AB, van Merrienboer JJ. Finding and fixing mistakes: do checklists work for clinicians with different levels of experience?. *Adv Health Sci Educ Theory Pract* 2014;19(1): 43–51.
35. Ark TK, Brooks LR, Eva KW. Giving learners the best of both worlds: do clinical teachers need to guard against teaching pattern recognition to novices? *Acad Med* 2006;81(4):405–409.
36. Ark TK, Brooks LR, Eva KW. The benefits of flexibility: the pedagogical value of instructions to adopt multifaceted diagnostic reasoning strategies. *Med Educ* 2007;41(3):281–287.
37. Croskerry P, Singhal G, Mamede S. Cognitive debiasing 2: impediments to and strategies for change. *BMJ Qual Saf* 2013;22 (Suppl 2):ii65–ii72.
38. Custers EJ. Medical education and cognitive continuum theory: an alternative perspective on medical problem solving and clinical reasoning. *Acad Med* 2013;88(8):1074–1080.

39. Elia F, Apra F, Verhovez A, Crupi V. "First, know thyself": cognition and error in medicine. *Acta Diabetol* 2016;53(2): 169–175.
40. Heiberg Engel PJ. Tacit knowledge and visual expertise in medical diagnostic reasoning: implications for medical education. *Med Teach* 2008;30(7):e184–e188.
41. Lambe KA, O'Reilly G, Kelly BD, Curristan S. Dual-process cognitive interventions to enhance diagnostic reasoning: a systematic review. *BMJ Qual Saf* 2016:12.
42. Pelaccia T, Tardif J, Triby E, Charlin B. An analysis of clinical reasoning through a recent and comprehensive approach: the dual-process theory. *Med Educ Online* 2011:16.
43. Phua DH, Tan NC. Cognitive aspect of diagnostic errors. *Ann Acad Med Singap* 2013;42(1):33–41.
44. Stolper E, Van de Wiel M, Van Royen P, Van Bokhoven M, Van der Weijden T, Dinant GJ. Gut feelings as a third track in general practitioners' diagnostic reasoning. *J Gen Intern Med* 2011;26(2): 197–203.
45. van den Berge K, Mamede S. Cognitive diagnostic error in internal medicine. *Eur J Intern Med* 2013;24(6):525–529.
46. Norman G, Young M, Brooks L. Non-analytical models of clinical reasoning: the role of experience. *Med Educ* 2007;41 (12):1140–1145.
47. Groves M, O'Rourke P, Alexander H. Clinical reasoning: the relative contribution of identification, interpretation and hypothesis errors to misdiagnosis. *Med Teach* 2003;25(6):621–625.
48. Groves M, O'Rourke P, Alexander H. The clinical reasoning characteristics of diagnostic experts. *Med Teach* 2003;25(3): 308–313.
49. Ely JW, Graber ML, Croskerry P. Checklists to reduce diagnostic errors. *Acad Med* 2011;86(3):307–313.
50. Mamede S, Schmidt HG, Rikers RM, Penaforte JC, Coelho-Filho JM. Breaking down automaticity: case ambiguity and the shift to reflective approaches in clinical reasoning. *Med Educ* 2007;41 (12):1185–1192.
51. Dong T, Durning SJ, Artino AR, van der Vleuten C, Holmboe E, Lipner R, et al. Dual process theory and intermediate effect: are faculty and residents' performance on multiple-choice, licensing exam questions different? *Mil Med* 2015;180(4):92–96.
52. Durning SJ, Graner J, Artino AR, Jr., Pangaro LN, Beckman T, Holmboe E, et al. Using functional neuroimaging combined with a think-aloud protocol to explore clinical reasoning expertise in internal medicine. *Mil Med* 2012;177(9 Suppl):72–78.
53. Durning SJ, Costanzo ME, Beckman TJ, Artino AR, Jr., Roy MJ, van der Vleuten C, et al. Functional neuroimaging correlates of thinking flexibility and knowledge structure in memory: Exploring the relationships between clinical reasoning and diagnostic thinking. *Med Teach* 2016;38(6):570–577.
54. Michelsen J, Malchow-Moller A, Charles P, Eika B. Young doctors' problem solving strategies on call may be improved. *Dan Med J* 2013;60(3):A4586.
55. Stolper CF, Van de Wiel MWJ, Hendriks RHM, Van Royen P, Van Bokhoven MA, Van der Weijden T, et al. How do gut feelings feature in tutorial dialogues on diagnostic reasoning in GP traineeship? *Adv Health Sci Educ* 2015;20(2):499–513.