Evidence on Use of Clinical Reasoning Checklists for Diagnostic Error Reduction

Issue Brief

Evidence on Use of Clinical Reasoning Checklists for Diagnostic Error Reduction

Introduction

The use of checklists as a tool to improve performance has proven successful in a variety of healthcare settings. For instance, checklists have been successful in preventing hospital-acquired infections[1] and preventing errors in the surgical process.[2] The use of checklists has also been recommended as a tool to reduce diagnostic errors.[3] Diagnostic errors are frequent and often have severe consequences[4] but have received little attention in the field of patient safety.

Checklists are considered a promising intervention for the area of diagnosis because they can support clinicians in their diagnostic decision making by helping them take correct diagnostic steps and ensuring that possible diagnoses are not overlooked. In this issue brief, we summarize current evidence on using checklists to improve diagnostic reasoning.

Rationale for Use

Diagnosis is complex because clinicians need not only a large fund of available baseline knowledge, but also the ability to apply this knowledge in a challenging work environment. In addition, they need to consider alternative explanations, including rare diseases and diseases that require urgent treatment.

Given the success of checklists in other domains of patient safety, experts in diagnostic safety have advocated for the development of checklists to support clinicians in their diagnostic reasoning process.[3,5] However, only a few studies have evaluated the effect of checklists on diagnostic accuracy,[6] and the results are mixed.

In this brief, we will focus on checklists used for cognitive support (i.e., those that remind clinicians of the correct diagnostic steps with the goal of increasing diagnostic accuracy). We will discuss the evidence for the effectiveness of checklists for diagnostic error reduction and factors that influence their effectiveness. Finally, we will discuss next steps for research.

Content-Specific Versus Process-Focused Checklists

Checklists used in the diagnostic reasoning process can be divided into two categories. *Content-specific* checklists provide clinicians with relevant knowledge during the diagnostic process or trigger them to activate their knowledge. For example, these may list specific diagnostic steps or suggestions of possible diagnoses that should be considered for a specific patient.

One example is a checklist to interpret electrocardiograms (ECGs) that included content-specific features, such as calculation of heart rate. Sibbald and colleagues found in several studies that the use of this checklist reduced diagnostic error based on interpretation of ECGs.[7–9] Other studies also showed an effect of content-specific checklists, but the effects are often small or apply only to a subgroup of clinicians.[10,11]

Process-focused checklists aim to trigger more deliberate thinking when diagnosing a patient. An example is a "debiasing" checklist that aims to reduce errors that occur due to shortcuts in the diagnostic reasoning process (i.e., cognitive biases).[12] These checklists often contain items such as "what else can it be?"

A recent study by O'Sullivan and Shofield evaluated the use of a cognitive forcing mnemonic, called "SLOW", on diagnostic accuracy. "SLOW" is an acronym for four questions: (1) "Sure about that? Why?" (2) "Look at the data, what is lacking? Does it all link together?" (3) "Opposite: what if the opposite is true?" and (4) Worst case scenario; "What else could this be?" A randomized trial found no effect of the SLOW intervention (compared with no intervention) on diagnostic accuracy based on clinical vignettes.[13] Similarly, most studies that evaluated process-focused checklists found no significant effects on accuracy.[14,15]

Two studies have directly compared content-specific checklists and process-focused checklists. In a study by Sibbald and colleagues on ECG interpretation, the content-specific (knowledge-based) checklist as described before was compared with a process-focused (debiasing) checklist and a control group. The overall results did not show a significant improved performance on ECG interpretation with either checklist.[14] This finding is in contrast to several earlier studies by Sibbald, et al., in which the content-specific checklist showed an effect.[7,8]

A study by Shimizu and colleagues compared medical students' intuitive process (i.e., list the three most likely diagnoses) with one of two checklists: (1) a content-specific checklist that suggested differential diagnosis for the case at hand or (2) a process-focused checklist, i.e., a general debiasing checklist developed by Ely, et al.,[5] with checklist items such as "Did I obtain a complete medical history?" and taking a "diagnostic time out."

The authors exposed the participants to both simple and difficult clinical case vignettes based on actual patient experiences. Overall, they found that the use of a checklist did not improve accuracy in the easy cases; on the contrary, diagnostic accuracy was reduced by the use of checklists. For difficult cases, the content-specific checklist improved diagnostic accuracy, but the debiasing checklist was not effective in either simple or difficult cases.[16]

Taking all this research into account, content-specific checklists seem to be more promising than process-focused checklists, but the evidence is relatively thin with few studies.

Factors That Influence Effectiveness of Checklists

Several factors seem to be associated with the effectiveness of checklists for diagnostic safety. First, some studies have shown that checklists are more effective when used by novices compared with experts.[8] Thus, checklists may work differently for clinicians with different levels of experience. This finding may be related to the second factor of influence: the level of difficulty of a case.

Checklists seem to help more in complex cases than in simple cases,[16] which is similar to studies on the effects of reflection.[17] However, the evidence is not conclusive, as one study showed a positive effect on simple cases as well.[18] Checklists may be more effective in difficult cases because there is more room for error and for improvement. However, in clinical practice it is often hard to distinguish a simple case from a difficult case.

Most studies that have examined the effect of checklists on diagnostic accuracy were conducted in experimental settings.[7,8,10,13,14,16] In such settings, potentially confounding factors such as case mix and complexity of the cases can be controlled. Past studies have also typically recruited medical students and residents, who have lower levels of experience. Furthermore, studies in experimental settings often use complex cases, which reflect a different sample than most clinicians encounter in clinical practice. Lastly, while in experimental studies participants are required to use the checklists on all cases they see, in clinical

practice checklists may be used inconsistently.[19] Thus, past study designs may have overestimated the effects of checklists on diagnostic performance.

Most studies have not taken into account the potentially negative effects of implementing clinical reasoning checklists in clinical practice. Specifically, the use of checklists can be time consuming[10,19] and can result in ordering more laboratory tests and imaging.[19]

Why Checklists for Diagnosis Are Not Performing as Expected

The evidence that checklists can improve diagnostic safety is thin, which is surprising because the face validity of checklist use is high and several experts have promoted checklist use to reduce diagnostic errors. Furthermore, checklists have been very successful for addressing other threats to patient safety.[1,2] Why is this not the case in diagnostic safety?

Successful checklists for preventing other error types list very specific tasks. For example, the first step in the well-known checklist to reduce central-line infections is "Wash your hands with soap,"[1] and the widely adopted surgical checklist starts with confirming the patient's identity.[2] These checklists are meant to prevent errors of execution, so-called "slips" (attention failures) or "lapses" (memory failures).[20] Typical for these errors is that the clinician had the right plan for the task but erred only in the execution (e.g., forgetting a step in the preoperative process, marking the wrong limb). These types of errors are easily prevented by a checklist that prevents clinicians from skipping steps in the process.

Conversely, checklists used for diagnostic safety seem to focus on errors of planning. These errors occur when the plan of an action was incorrect (e.g., due to lack of knowledge). A frequently used item on checklists in the diagnostic process is "What else can it be?",[13] which prompts the clinician to reconsider the diagnostic process and reflect on possible alternatives. In other words, clinicians are asked to evaluate the task they have just performed without any suggestion of what they might have missed. An important and unanswered question for diagnostic safety is whether checklists can prevent such errors. Even the current content-specific checklists for diagnostic safety may not be specific enough.

Conclusions and Next Steps

Checklists to improve diagnostic reasoning are not ready for use in clinical practice. Evidence that checklists improve diagnostic accuracy is mixed (see Table 1), and positive effects of checklists on diagnostic accuracy are mainly found in subgroups of cases (difficult cases) or clinicians (junior clinicians). Furthermore, checklists have potentially negative effects, such as time pressure and overdiagnosis, which have been insufficiently studied. Finally, most studies that measure effects on accuracy are performed in controlled settings that do not resemble typical clinical practice, and even the modest benefits of checklists may therefore be overestimated. The fact that most studies already show a limited effect of checklists on diagnostic accuracy in experimental settings does not bode well for the use of checklists in clinical practice.

Conceptually determining whether checklists can be useful for improving diagnostic safety requires answering some critical questions. Can we develop checklists for diagnostic error reduction that focus on errors of execution rather than errors of planning? Are checklists effective when tested in a diverse population (including experienced clinicians) under realistic circumstances and with a realistic case mix? Subsequently, pilot testing of potentially effective checklists in clinical settings is crucial.

While research on the use of checklists in diagnostic safety is still in its infancy, more indepth evaluation, including a focus on implementation factors and the contexts for use, will help answer these and other critical questions and demonstrate if and how checklists can be a viable tool for diagnostic error reduction.

Table 1. Overview of studies on the effectiveness of checklists

Content-Specific Checklists

Study	Description of the Checklist	Participants	Setting	Outcome
Sibbald, et al., 2013[7]	ECG interpretation checklist	Cardiology fellows	Experimental	Checklist was effective. Average of 0.39 errors in checklist condition versus 1.04 in non-checklist condition.
Sibbald, et al., 2013[9]	Cardiac exam checklist	Internal medicine residents	Experimental	Checklist was effective. Accuracy improved from 46% pre-checklist condition to 51% post-checklist condition. The checklist was only effective if the residents could access more information while using the checklist.
Ely and Graber, 2015[11]	Differential diagnosis checklist	Primary care physicians	Clinical	No significant difference between checklist condition and control condition.
Kok, et al., 2017[10]	Chest radiograph interpretation checklist	Medical students	Experimental	More abnormalities were correctly detected when the checklist was used (41.9% accuracy without checklist and 50.1% with checklist).

Process-Focused Checklists

O'Sullivan and Schofield, 2019[13]	Mnemonic tool (SLOW) focused on slowing down reasoning and countering specific biases	Medical professionals (ranging from medical students to attending physicians)	Experimental	No significant difference between checklist condition and control condition.

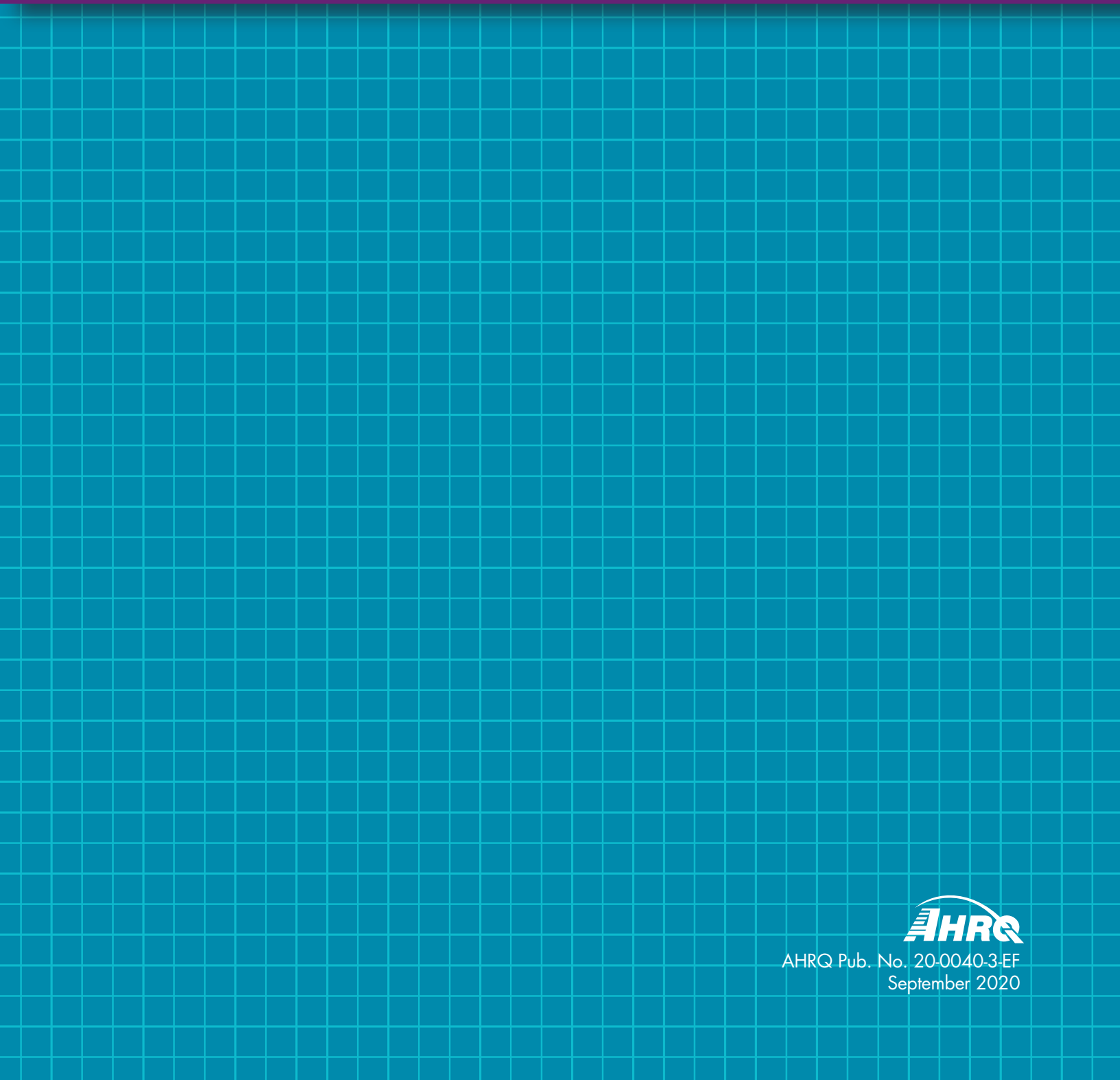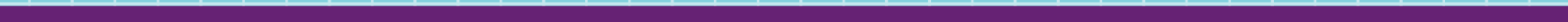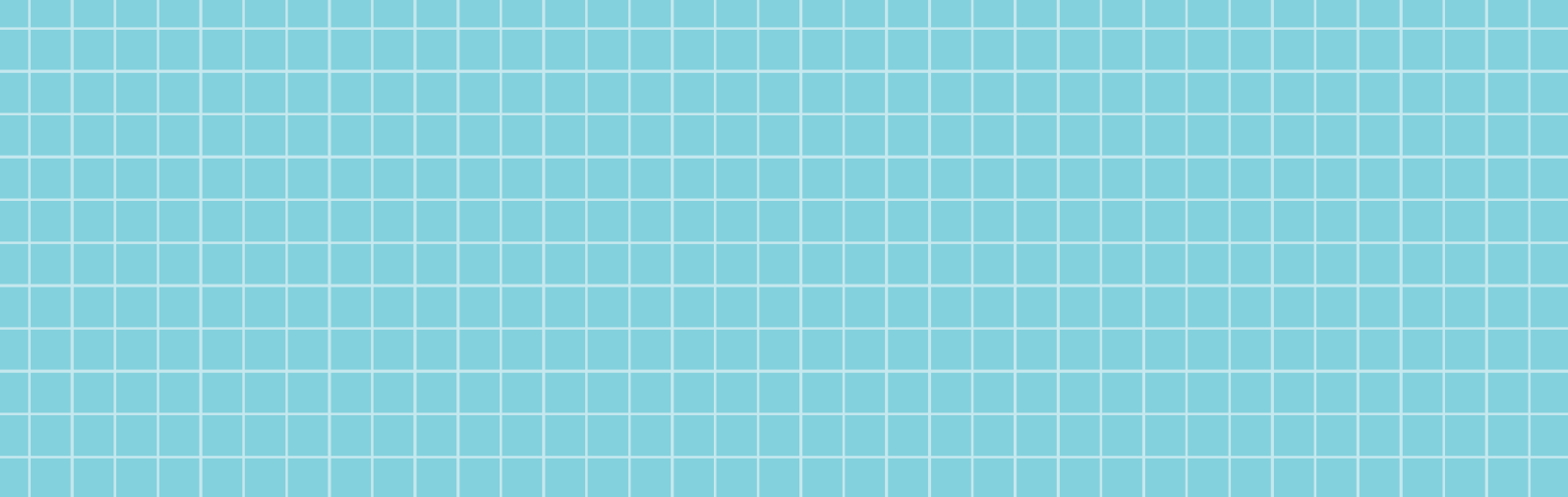**Compared Content-Specific and Process-Focused Checklists**

Study	Description of the Checklist	Participants	Setting	Outcome
Shimizu, et al., 2013[16]	Differential diagnosis checklist (content specific) compared with a general debiasing checklist (process focused) and a control group (intuitive diagnosis)	Medical students	Experimental	Significant effect for the use of the content-specific checklist compared with intuitive diagnosis. The checklist particularly improved accuracy in the difficult cases. No significant effect for the debiasing checklist.
Sibbald, et al., 2019[14]	ECG interpretation checklist (content specific) compared with a cognitive debiasing checklist (process focused) and a control group	Internal medicine residents and cardiology fellows	Experimental	No overall significant differences between the content-specific checklist, process-focused checklist, and control group.

References

1. Pronovost P, Needham D, Berenholtz S, Sinopoli D, Chu H, Cosgrove S, Sexton B, Hyzy R, Welsh R, Roth G, Bander J, Kepros J, Goeschel C. An intervention to decrease catheter-related bloodstream infections in the ICU. New Engl J Med. 2006;355:2725-2732. https://www.nejm.org/doi/10.1056/NEJMoa061115. Accessed September 4, 2020.

2. Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat AH, Dellinger EP, Herbosa T, Joseph S, Kibatala PL, Lapitan MC, Merry AF, Moorthy K, Reznick RK, Taylor B, Gawande AA; Safe Surgery Saves Lives Study Group. A surgical safety checklist to reduce morbidity and mortality in a global population. New Engl J Med. 2009;360:491-499. https://www.nejm.org/doi/10.1056/NEJMsa0810119. Accessed September 4, 2020.

3. Gupta A, Graber ML. Annals for Hospitalists Inpatient Notes - Just what the doctor ordered—checklists to improve diagnosis. Ann Intern Med 2019;170:HO2-HO3. https://www.acpjournals.org/doi/10.7326/M19-0829. Accessed September 4, 2020.

4. Zwaan L, de Bruijne M, Wagner C, Thijs A, Smits M, van der Wal G, Timmermans DR. Patient record review of the incidence, consequences, and causes of diagnostic adverse events. Arch Intern Med. 2010;170:1015-1021. https://pubmed.ncbi.nlm.nih.gov/20585065/. Accessed September 4, 2020.

5. Ely JW, Graber ML, Croskerry P. Checklists to reduce diagnostic errors. Acad Med. 2011;86:307-313. https://pubmed.ncbi.nlm.nih.gov/21248608/. Accessed September 4, 2020.

6. McDonald KM, Matesic B, Contopoulos-Ioannidis DG, Lonhart J, Schmidt E, Pineda N, Ioannidis JP. Patient safety strategies targeted at diagnostic errors: a systematic review. Ann Intern Med 2013;158:381-389. https://www.acpjournals.org/doi/10.7326/0003-4819-158-5-201303051-00004. Accessed September 4, 2020.

7. Sibbald M, de Bruin ABH, van Merrienboer JJG. Checklists improve experts' diagnostic decisions. Med Educ. 2013;47:301-308. https://pubmed.ncbi.nlm.nih.gov/23398016/. Accessed September 4, 2020.

8. Sibbald M, De Bruin ABH, van Merrienboer JJG. Finding and fixing mistakes: do checklists work for clinicians with different levels of experience? Adv Health Sci Educ Theory Pract. 2014;19:43-51. https://pubmed.ncbi.nlm.nih.gov/23625338/. Accessed September 4, 2020.

9. Sibbald M, de Bruin ABH, Cavalcanti RB, van Merrienboer JJ. Do you have to re-examine to reconsider your diagnosis? Checklists and cardiac exam. BMJ Qual Saf. 2013;22:333-338. https://pubmed.ncbi.nlm.nih.gov/23386730/. Accessed September 4, 2020.

10. Kok EM, Abed A, Robben SGF. Does the use of a checklist help medical students in the detection of abnormalities on a chest radiograph? J Digit Imaging. 2017;30:726-731. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5681473/. Accessed September 4, 2020.

11. Ely JW, Graber MA. Checklists to prevent diagnostic errors: a pilot randomized controlled trial. Diagnosis (Berl). 2015;2:163-169. https://pubmed.ncbi.nlm.nih.gov/29540029. Accessed September 4, 2020./

12. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. Science. 1974;185:1124-1131. https://pubmed.ncbi.nlm.nih.gov/17835457/. Accessed September 4, 2020.

13. O'Sullivan ED, Schofield SJ. A cognitive forcing tool to mitigate cognitive bias – a randomised control trial. BMC Med Educ. 2019;19:12. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6325867/. Accessed September 4, 2020.

14. Sibbald M, Sherbino J, Ilgen JS, Zwaan L, Blissett S, Monteiro S, Norman G. Debiasing versus knowledge retrieval checklists to reduce diagnostic error in ECG interpretation. Adv Health Sci Educ Theory Pract. 2019 Aug;24(3):427-440. Epub 2019 Jan 29. https://pubmed.ncbi.nlm.nih.gov/30694452. Accessed September 4, 2020.

15. Lambe KA, Hevey D, Kelly BD. Guided reflection interventions show no effect on diagnostic accuracy in medical students. Front Psychol. 2018 Nov 23;9:2297. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6265413/. Accessed September 4, 2020.

16. 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:e1218-1229. https://pubmed.ncbi.nlm.nih.gov/23228085/. Accessed September 4, 2020.

17. Mamede S, Hautz WE, Berendonk C, Hautz SC, Sauter TC, Rotgans J, Zwaan L, Schmidt HG. Think twice: effects on diagnostic accuracy of returning to the case to reflect upon the initial diagnosis. Acad Med. 2020 Aug;95(8):1223-1229. https://pubmed.ncbi.nlm.nih.gov/31972673/. Accessed September 4, 2020.

18. DiNardo D, Tilstra S, McNeil M, Follansbee W, Zimmer S, Farris C, Barnato AE. Identification of facilitators and barriers to residents' use of a clinical reasoning tool. Diagnosis (Berl). 2018;5:21-28. https://pubmed.ncbi.nlm.nih.gov/29601296/. Accessed September 4, 2020.

19. Graber ML, Sorensen AV, Biswas J, Modi V, Wackett A, Johnson S, Lenfestey N, Meyer AN, Singh H. Developing checklists to prevent diagnostic error in Emergency Room settings. Diagnosis (Berl). 2014 Sep;1(3):223231. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4799784/. Accessed September 4, 2020.

20. Reason J. Human Error. Cambridge, UK: Cambridge University Press; 1990.