

Published in final edited form as:

Acad Med. 2010 July ; 85(7): 1112–1117. doi:10.1097/ACM.0b013e3181e16103.

Swapping Horses Midstream: Factors Related to Physicians' Changing Their Minds About a Diagnosis

Kevin W. Eva, PhD [associate professor],

Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton Ontario, Canada.

Carol L. Link, PhD [senior research scientist],

New England Research Institutes (NERI), Watertown, Massachusetts.

Karen E. Lutfey, PhD [senior research scientist and director], and

Center on Patient-Provider Relationship, New England Research Institutes (NERI), Watertown, Massachusetts.

John B. McKinlay, PhD [founding director, senior vice president, and chief scientist]

New England Research Institutes (NERI), Watertown, Massachusetts.

Abstract

Purpose—Premature closure has been identified as the single most common cause of diagnostic error. The authors conducted a factorial experiment to explore which variables exert an unconfounded influence on physicians' diagnostic flexibility (changing their minds about the most likely diagnosis during a clinical case presentation).

Methods—In 2007–2008, 256 practicing physicians viewed a clinically authentic vignette simulating a patient presenting with possible coronary heart disease (CHD), provided their initial impression midway through the case, answered questions about the case, indicated how they would continue their clinical investigation, and made a final diagnosis. The authors used general linear models to determine which patient factors (age, gender, socioeconomic status, race), physician factors (gender, age/experience), and process variables were related to the likelihood of physicians' changing their minds about the most likely diagnosis.

Results—Physicians who had less experience, those who named a non-CHD diagnosis as their initial impression, and those who did not ask for information about the patient's prior cardiac disease history were the most likely to change their minds. Participants' certainty in their initial diagnosis, the additional information desired, the diagnostic hypotheses generated, and the follow-up intended were not related to the likelihood of change in diagnostic hypotheses.

Discussion—While efforts encouraging physicians to avoid cognitive biases and to reason in a more analytic manner may yield some benefit, this study suggests that experience is a more important determinant of diagnostic flexibility than is the consideration of additional diagnoses or the amount of additional information collected.

Multidisciplinary research and policy work that focuses on identifying and mitigating sources of medical errors has proliferated in recent years. System-level safeguards to protect against human error now range from increasing use of information technology to avoid

Correspondence should be addressed to Dr. Eva, Department of Clinical Epidemiology and Biostatistics, Program for Educational Research and Development, MDCL 3522, McMaster University, 1200 Main Street West, Hamilton, Ontario, Canada L8N 3Z5; telephone: (905) 525-9140, x27241; fax: (905) 572-7099; evakw@mcmaster.ca.

Other disclosures: None.

medication errors to having patients sign their own surgical sites as part of the consent process to avoid mistakes during operations. Despite this breadth of strategy, medical errors persist and research suggests a substantial human component. Diagnostic errors, a subset of all medical errors, have been reported to occur in 10–15% of all patient cases,¹ and 74% of all diagnostic errors have been identified as having some cognitive basis.²

Understanding physicians' cognitive processes is therefore a linchpin to error reduction, but there are multiple challenges that make it an especially difficult topic to research. Researchers have identified more than 30 cognitive biases,³ but most are believed to operate without conscious awareness, making it unrealistic to ask physicians which biases influence their judgment, how often, with which patients, and whether those biases lead them toward suboptimal decisions. Instead, labels are typically applied to diagnostic errors through post hoc explanations provided by case reviewers, a process that is itself susceptible to bias.⁴

The recent book *How Doctors Think*⁵ popularized the refrain that diagnosticians (and patients) must become more aware of decision-making biases to avoid falling prey to nonanalytic heuristics (cognitive shortcuts) like confirmation bias (preferentially seeking out information that confirms a hypothesis) and premature closure (concluding in favor of a particular diagnosis before sufficient information has been gathered). The latter issue has been identified as the single most prevalent cause of diagnostic error.¹ Some researchers believe that making clinicians aware of the influence of nonanalytic processes (also known as “System 1” processes), such as pattern recognition, will help them overcome these misleading influences and the resulting certainty in their initial diagnostic impressions.⁶ Others argue that such biases are difficult to extinguish naturally because they represent fast and frugal ways of dealing with challenging problems that, in general, yield more benefit than harm.⁷

While this debate highlights the complexity of the issue and the amount of attention it has received, researchers still know relatively little about how the various biases identified affect physicians' actual decision making and whether cognitive de-biasing strategies have potential to reduce errors in diagnosticians' thinking. Therefore, implementing targeted and effective educational and policy interventions to reduce error remains an elusive goal.

To partially address this knowledge gap, we conducted a study to examine diagnostic flexibility—that is, the extent to which physicians change their minds about the most likely diagnosis during a clinical encounter. By examining which characteristics of patients, diagnosticians, or physicians' approaches to a particular case are predictive of physicians' diagnostic flexibility, we may be able to more precisely guide continuing education efforts, facilitate self-monitoring habits, and better understand when cognitive biases yield diagnostic errors.

Examining Cognitive Flexibility

Studies of dual process models of cognition, such as those alluded to above, yield consistent findings in the context of both everyday problem solving⁸ and medical diagnosis.⁹ The most extensive way in which researchers have studied the type of cognitive flexibility we explore here is by using a psychological tool called the Wisconsin Card Sorting Test (WCST). Participants' task in the WCST is to sort cards into meaningful categories that the person administering the test defines as right or wrong. On each sort, participants are told simply that they are correct or incorrect, but they are not told the basis of their (in)accuracy. This is because the key manipulation of the WCST is that the underlying categorization scheme deemed “correct” is altered throughout the test to determine the rate at which participants notice and adapt their problem solving categorizations to the new information being provided.

The WCST is most often used as a clinical test of executive mental functioning, but recent meta-analyses have suggested that healthy older adults are more likely to make perseveration errors than are their younger counterparts, with years of education playing a moderating role.¹⁰ This finding is consistent with other research that has suggested that age and/or experience (which are typically confounded in medical practice) are related to the tendency to rely on nonanalytic processes (and the first impressions they create),¹¹ as well as with recent work suggesting that more experienced (older) physicians are less likely to be influenced by the presentation of clinical features that are inconsistent with their initial hypothesis.¹² More work is needed, however, because the WCST and similar tests that have been used to study the psychology of decision making do not tend to encompass the real world richness of diagnostic decision making. As a result, it prevents researchers from examining cognitive flexibility across people who have considerable and idiosyncratic experiences through which various biases might arise and with which many variables might interact.

In this article, we take a different approach to examining cognitive flexibility and present findings from a factorial experiment concerning physician decision making. We presented videotaped simulated patients with identical signs and symptoms to participants, whom we asked to answer questions about how they would work up and manage the case before them. In addition to the seven design factors (described below), we considered the questions physicians asked in response to the case, the initial and final diagnoses they generated, and the tests they wanted to order. Our goal was to examine which patient attributes, physician characteristics, and process variables were related to diagnostic flexibility.

Method

Ethics approval was obtained from the Institutional Review Board at the New England Research Institutes.

Design factors

This study had seven design factors: four patient factors (age, gender, race, and socioeconomic status); two physician factors (gender and experience); and one experimental factor (half of the physicians were primed to consider a CHD diagnosis). Five of the factors were experimentally manipulated (physicians were randomly assigned patient vignettes and were randomly assigned to be primed or not). The other two design factors were stratification factors (physicians were recruited to fill four strata defined by gender and experience).

Participants

In 2007–2008, we recruited 256 primary care physicians who worked at least half time in North Carolina or South Carolina to participate. This sample size enabled two replications of the full factorial design and was determined to provide 80% power to detect differences with an effect size of 0.2. We mailed a letter of introduction to potential participants and followed up with a screening call to confirm eligibility. To be eligible, physicians had to have completed medical school between 1960 and 1987 or between 1996 and 2001. We chose these date ranges so we could make clean distinctions between age groups (i.e., physicians with more or less clinical experience).

We mailed letters to 1,930 physicians, inviting them to participate in a study of primary care physicians' decision making. Of these, 1,131 were deemed ineligible due to inaccurate contact information, illness, death, or the proxy decision of the letter recipient. We continued recruitment of the remaining 799 until 64 male and 64 female physicians within

each experience level agreed to participate. We contacted 606 physicians to achieve the total sample of 256, thus yielding a participation rate of 42.2% among the contacted eligible physicians.

Upon confirming eligibility and willingness to participate, we assigned each participant randomly to a priming condition and one of 16 patient vignettes (described below). A research assistant arranged a one-hour appointment to visit the physician in his or her practice setting during the course of a normal work day to conduct the patient presentation and interview.

The diagnostic task

Participants viewed a videotaped vignette of a patient presenting with signs and symptoms suggestive of coronary heart disease (CHD) and answered a series of questions about how they would examine, question, and manage the presenting patient. Physicians were asked to view the patient as one of their own and to respond in the way they would respond within the context of their own practice. We chose CHD as the diagnosis because it is a common and costly problem that primary care providers regularly encounter and because it represents a well-defined and extensively studied health condition.

Vignette content and creation

We created 16 vignettes using one script and professional actors to ensure consistency of presentation. One vignette was filmed for each possible combination of patient age (55 or 75), gender (male or female), race (black or white), and socioeconomic status (higher or lower; indicated by current/past employment as a school teacher or janitor). These variations were based on past work indicating that these patient characteristics affect the rate at which physicians assign a diagnosis of CHD.¹³ In general, inquiry into the social determinants of disease has shown that patients who are male, older, Caucasian, or of higher socioeconomic status are more likely than their counterparts to be diagnosed with CHD, even when all of the diagnostically relevant features are held constant across these variables. We based the script on several video-recorded role-playing sessions we conducted with experienced clinical advisors. A physician consultant was present during filming to ensure clinical authenticity and continuity in the case presentation across actors.

The vignettes portrayed several indications of CHD (chest pain worsening with exertion, pain between the shoulder blades, stress, and increased blood pressure), but for the sake of authenticity they also contained various misleading indications of other disorders. Some of these were gastrointestinal (GI) in nature (indigestion, feeling worse after a spicy or large meal, having heartburn-like pain that was unresponsive to antacids); others were psychological (feeling irritated, being lethargic, and a spouse's report of being difficult to be around).

We used videos so nonverbal cues, such as the "Levine fist" (indicating chest pain), could be incorporated and because past work has shown that such vignettes provide valid indicators of outpatient care.¹⁴ We asked participants to rate the extent to which the patient in the vignette resembled patients they encounter in their everyday practice, and 230/256 (89.8%) considered the taped presentations typical or very typical.

Patient presentation and interview

Half the participants, randomly chosen upon recruitment within each gender and experience stratum, were primed (explicitly directed) to consider CHD as a diagnosis. The interviewer told them: "The patient in the video was recently on vacation and sought medical advice for her/his symptoms. The physician mentioned the possibility of coronary heart disease and

suggested s/he see her/his primary care physician upon returning home.” This was done in an attempt to examine the influence of an external bias on participants’ decision making.

Approximately halfway through the vignette, the interviewer stopped the video and asked the participants to name the primary diagnosis they were considering to that point, to rate their certainty of that diagnosis (using a 0–100 rating scale, with 0 indicating no certainty and 100 indicating complete certainty), and to identify the “most important piece of information you still hope to obtain from the remaining portion of the vignette.” The video was then played to completion, at which point the interviewer instructed the participants: “We recognize that you might be considering several possible diagnoses for this patient. Which do you think is the most likely condition?” After the participants responded to this question, the interviewer asked them to list additional diagnoses they were considering, to rate their level of certainty for each diagnosis named, and to respond to a series of structured interview questions regarding what they would do with this patient (additional questions they would ask, physical examinations they would perform, tests they would order, medications they would prescribe, lifestyle advice they would provide, and other physicians to whom they would refer).

Questions were open-ended; responses were recorded verbatim and coded in-house (after the interview was completed) as being relevant to CHD, GI disorder, or another diagnosis. This coding was completed using a consensus model. Pilot interviews were conducted in early 2007 and a coding rubric created based on the responses through extensive consultation with two clinical colleagues. After each interview the research assistant who conducted the interview applied the coding rubric to the physician’s responses and reviewed the coding decisions with one principal investigator (KL). Responses for which the appropriate code remained unclear were discussed with the other principal investigator (KE) and reviewed by our clinical consultants.

Statistical analysis

The balanced factorial design allows the unconfounded estimation of main effects for each of the experimental variables included in the study—that is, the 4 patient factors \times 2 physician factors \times whether or not physicians were primed to think of CHD. We used descriptive statistics, chisquared tests, ANOVA, and logistic regression analyses to assess the relationship between these variables and the outcomes of interest. We used ANOVA to compare dichotomous variables (e.g., CHD named as the most likely diagnosis or not, coded as 0 or 1) when it was desirable to consider the effect of many variables and interactions in a single analysis. While logistic regression was an alternative, we chose ANOVA for three reasons: (1) it allows a complete model to be specified (main effects and all interactions) using the pure error due to replication (128 degrees of freedom) as the error term; (2) we have found the two types of analyses to yield comparable *P* values when no missing data exist, as in the current study; and (3) we have found the output of ANOVA to be more readily interpretable to others working in this field. To determine which physicians changed their minds during the course of the case presentation, we classified those physicians who provided a different final diagnosis relative to their initial impression as having changed their minds and those who provided the same diagnosis as the most probable at both response points as having consistent opinions.

Results

Initial impression

At the midway point, when the video was paused, 154/256 (60.2%) of participants named the primary diagnosis as CHD, 55/256 (21.5%) named a GI disorder, and 47/256 (18.4%)

named something else. Using ANOVA to consider which main effects (patient factors, physician factors, and priming) or two-way interactions were associated with participants' naming CHD as their initial diagnosis revealed that those primed to think of CHD were more likely (84/128; 65.6%) to name CHD than were those in the nonprimed cohort (70/128; 54.7%; $F(1,128)=3.50$, $P=.06$). This comparison did not interact with experience level of the participant ($F(1,128)=0.07$, $P=.79$), although older/more experienced physicians were more likely to name CHD as their initial diagnosis (65.6%) than were less experienced physicians (54.7%; $F(1,128)=3.50$, $P=.06$). Patient age was the only other variable that predicted generation of CHD as a diagnosis: older patients elicited the diagnosis more frequently (69.5%) than did younger patients (50.8%; $F(1,128)=10.29$, $P=.002$).

The average certainty rating assigned to this initial impression was 70.8 (standard deviation [SD] = 19.4) on a scale of 0–100. Physicians' certainty in their initial impression was not influenced by any of the experimental variables. Similar analyses performed on the coded responses to the question asking for “the most important information you still hope to obtain” revealed that those who named CHD as their initial diagnosis were more likely to indicate a desire for further information indicative of the diagnosis of CHD (72/154; 46.8%) than were individuals who did not name CHD as their primary diagnosis (35/102; 34.3%; $\chi^2=3.9$, $P=.048$). The rate at which such confirmatory information was requested was not influenced by any of the experimental variables.

Final diagnosis

At completion of the video, 155/256 (60.5%) of participants named CHD as the “most likely” diagnosis, 64/256 (25.0%) named a GI disorder, and 37/256 (14.5%) named something else. Despite the similarity in proportions at both the initial and final response points, 94 of the 256 participants (36.8%) changed their “most likely” diagnosis away from their initial impression, whereas the other 162 (63.2%) maintained their first opinion until the end of the study. The average certainty rating (of the “most likely” diagnosis) was 66.7 (SD = 17.3); male patients elicited greater certainty ratings (69.4) than did female patients (64.0; $F(1,128)=5.77$, $P=.018$). Certainty ratings were not related to any other experimental variable.

We used participant characteristics (gender, experience, and priming status), patient characteristics (age, gender, race, and socioeconomic status) and participant response variables (the questions asked about cardiac risk factors, number of examinations to be performed, number of diagnoses named, whether CHD was the initial diagnosis, certainty in the initial diagnosis, the number of tests to be ordered, and whether the additional piece of information desired at the stopping point was aimed at confirming the initial diagnosis) as independent variables in a logistic regression to determine which were predictive of physicians' changing their opinion. The only variables that were statistically predictive of change of opinion were as follows:

1. Experience level of the physician: less experienced physicians were more likely to change their minds relative to more experienced physicians (Odds Ratio [OR] =2.02; 95% confidence interval [CI] =1.14–3.60, $P=.016$);
2. Whether CHD was named as the participants' initial diagnosis: participants were less likely to change their opinions when the initial diagnosis was CHD relative to when it was not (OR=0.21; 95% CI=0.12–0.37, $P<.0001$); and,
3. Whether prior cardiac disease was queried: those physicians who asked about a prior history of cardiac disease were less likely to change their minds than were those who did not (OR=0.53; 95% CI=0.29–0.98, $P=.042$).

Table 1 illustrates these findings, revealing that when CHD was the initial diagnosis, less experienced physicians were more likely than experienced physicians to shift away from CHD, but when CHD was not the initial diagnosis, physicians in both experience groups changed their minds at equal rates.

Some of the nonsignificant relationships are worth noting explicitly because their lack of predictive capacity is equally as informative as knowing which variables predicted change of opinion. Physicians' changing their minds midstream was not related to the number of questions the physicians would ask the patient, the number of examinations the physicians wanted to perform, the number of tests the physicians would order, or the number of diagnoses generated within the physicians' differential. Desiring information confirmatory of their initial impression when asked for a tentative diagnosis midway through the case presentation did not influence the likelihood of changing their minds relative to desiring nonconfirmatory information. The certainty rating in the initial impression (i.e., expressed at the midway point) was equal in participants who subsequently changed their mind (70.9, SD = 17.2) and those who did not (70.7, SD = 22.9, $P=.95$), but the final certainty expressed was significantly lower in those who changed their mind (62.9, SD = 18.2) relative to those who did not (68.9, SD = 16.4, $P=.007$).

Discussion

The literature on diagnostic error is rife with examples of mistakes physicians make and accounts of cognitive processes that might lead to such errors, with the most common culprit identified as premature closure of one's diagnostic search.² Most explicit educational efforts attempt to help physicians and trainees overcome nonanalytic, heuristic-induced diagnostic errors (i.e., de-biasing themselves) through the use of careful, comprehensive, and analytic diagnostic strategies.⁶ We do not doubt that there are benefits to using analytic processes in decision making, but we would note that it is important to recognize that analytic and nonanalytic cognitive processes are not necessarily mutually exclusive and, indeed, may work best when used in conjunction with one another.¹⁵ The data we report in this article, however, lead us to question the strength of the relationship among the physician's diagnostic certainty, the amount of additional information the physician would gather to confirm the diagnosis, and the likelihood that the physician would change his or her mind over the course of a case presentation.

Older/more experienced clinicians in this sample were more likely than less experienced clinicians to name the diagnosis of CHD early in the case presentation. In medicine there is almost a perfect correlation between age and years of experience, thereby making it very difficult to tease apart the relative effect of these variables, but cognitive theories both of aging and of expertise would suggest similar influences of both variables in terms of increased reliance on nonanalytic/automatic processing and declining cognitive flexibility.^{8,11} Our finding of greater diagnostic accuracy on the part of the older/more experienced group runs counter to the general reported pattern of poorer performance in older physicians,¹⁶ but it is consistent with the notion that early hypotheses are more likely to be accurate when one has more experience on which to draw.¹¹ Being prompted to consider a diagnosis of CHD increased the likelihood that more and less experienced clinicians alike would name CHD as their primary diagnosis, as priming status and experience level did not interact with each another. Physicians' responses midway through the case presentation provided evidence of confirmation bias, as participants considering CHD as their initial diagnosis were more likely to desire information consistent with CHD than were those who thought a different diagnosis most likely.

Interestingly, however, fewer than half of the participants who initially favored CHD sought confirmatory information, and the confirmation bias exhibited appears to have had little influence on participants' eventual conclusions. Requesting a piece of confirmatory information was not related to the rate at which participants changed their minds about the most likely diagnosis. In fact, certainty in the initial diagnosis was also unrelated to the rate of physicians changing their minds, as were various process variables, including the number of diagnoses participants claimed to be considering in their differential diagnosis, the number of follow-up questions they would ask the patient, the number of physical maneuvers they would perform, and the number of tests they would order.

These findings run counter to the notion that physicians should overcome initial diagnostic biases by simply being more deliberate, more analytic, or more tentative in their consideration of patient cases. Feelings of uncertainty or less of a desire for confirmatory information did not increase the rate at which physicians changed their minds. Rather, such change seems to have been driven predominantly by whether physicians generated the most parsimonious diagnosis (CHD in this case) early in a patient encounter, as was suggested by early explorations of clinical reasoning.¹⁷

Whether incorrect initial impressions can be overcome remains a difficult question to answer. The results of even this relatively constrained study yielded a complicated pattern of diagnostic flexibility. When participants' initial impression was CHD, they were less likely to change their diagnosis than when their initial impression was something other than CHD. Table 1 illustrates, however, that the results are not quite as straightforward as shifting to or from the most parsimonious diagnosis. Among physicians whose preliminary diagnosis was CHD, less experienced clinicians were more likely than more experienced clinicians to be swayed away from that diagnosis. Among physicians whose preliminary diagnosis was something other than CHD, roughly 40% of both more and less experienced physicians stayed with the original diagnosis and 25% switched to another non-CHD diagnosis. Taken together, these findings suggest a benefit to experience *and* a lessened diagnostic flexibility; the latter can be good or bad depending on the specific situation and the accuracy of that first impression.¹¹

Limitations and strengths

There are substantial methodological challenges to the study of cognitive processes, especially in real world contexts. Although this study overcame some of those challenges, it is not without limitations. By using videotaped vignettes and collecting physician responses in the context of their own practice, we strove for a compromise between ecological validity (i.e., a clinically authentic case presentation) and experimental control. Doing so limited us to the use of a single case presentation, which may call into question the generalizability of our findings. We put extensive effort into ensuring the case was realistic and that the tasks expected of the physician were ecologically valid. The success of that effort was supported both by participants' opinions that the case was relevant and typical of what they would see in their everyday practice and by the previously reported evidence that video vignettes, when designed carefully, can provide valid indicators of practice.¹⁴ Still, we would not promote the absolute numbers and percentages outlined here as being representative of all clinical situations (i.e., all patient scenarios or all practice contexts), though we believe that the relative influence of each variable on participants' diagnostic flexibility remains interesting and provides guidance for further exploration into this important issue. By fully exploring physicians' decision making around a particularly prevalent and diagnostically rich scenario, we hope to stimulate further exploration of the generalizability and limits of the findings reported here.

Further, researchers can never exert true experimental control over whether research participants change their minds within a case presentation, thereby making it possible that unidentified confounding variables might influence the self-selection of individuals into the groups of those who changed their minds and those who did not. The number and variety of variables considered here minimizes this inevitable limitation, however. In fact, the factorial combination of patient and physician factors utilized in this experimental design is a major strength that we believe to be unique within the medical education literature. The relatively large sample size and prospective nature of the examination into physician decision making are additional strengths of this study.

Conclusion

Medicine is practiced in a complex environment and making errors is inevitable. Although some errors are system-related and others are related to the knowledge/experience of the individual clinician, most are likely multi-faceted. Evidence suggesting that a large proportion of diagnostic errors are related to cognitive processes demands that researchers strive to better understand the influence of cognitive biases and how to best incorporate such knowledge into medical education practices. The findings we report in this article suggest that interventions that instruct diagnosticians to be more deliberate in their diagnostic decision making by prompting them to consider alternatives or to gather more information to avoid the errors that arise from nonanalytic, heuristic-induced cognitive biases may be ineffective—such process variables were unrelated to the diagnostic flexibility displayed. Rather, we would advocate that researchers further explore ways in which to coordinate and improve the strengths of both analytic and nonanalytic approaches to diagnostic decision making.

Acknowledgments

The authors would like to thank Dr. John Avanian and Dr. Richard Grant for providing their expertise as clinical consultants in support of this project

Funding/Support: Financial support for this project was provided by the National Institutes for Health (National Heart, Blood, and Lung Institute) HL079174.

Ethical approval: Ethics approval was obtained from the Institutional Review Board at the New England Research Institutes.

References

1. Berner ES, Graber ML. Overconfidence as a cause of diagnostic error in medicine. *Amer J Med.* 2008; 121:S2–S33. [PubMed: 18440350]
2. Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. *Arch Int Med.* 2005; 165:1493–1499. [PubMed: 16009864]
3. Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. *Acad Med.* 2003; 78:775–780. [PubMed: 12915363]
4. Wears RL, Nemath CP. Replacing hindsight with insight: toward better understanding of diagnostic failures. *Ann Emerg Med.* 2007; 49:206–209. [PubMed: 17083994]
5. Groopman, JE. *How Doctors Think.* New York, New York: Houghton Mifflin Company; 2007.
6. Croskerry P. Achieving quality in clinical decision making: cognitive strategies and detection of bias. *Acad Emerg Med.* 2002; 9:1184–1204. [PubMed: 12414468]
7. Eva KW, Norman GR. Heuristics and biases – a biased perspective on clinical reasoning. *Med Educ.* 2005; 39:870–872. [PubMed: 16150023]
8. Pretz JE. Intuition versus analysis: Strategy and experience in complex everyday problem solving. *Mem Cognit.* 2008; 36:554–566.

9. Eva KW. What every teacher needs to know about clinical reasoning. *Med Educ.* 2005; 39:98–106. [PubMed: 15612906]
10. Rhodes MG. Age-related differences in performance on the Wisconsin Card Sorting Test: A meta-analytic review. *Psychol Aging.* 2004; 19:482–494. [PubMed: 15382998]
11. Eva KW. The aging physician: Changes in cognitive processing and their impact on medical practice. *Acad Med.* 2002; 77:S1–S6. [PubMed: 12377689]
12. Eva KW, Cunnington JP. The difficulty with experience: Does practice increase susceptibility to premature closure? *J Contin Educ Health Prof.* 2006; 26:192–198. [PubMed: 16986144]
13. McKinlay JB, Link CL, Freund KM, Marceau LD, O'Donnell AB, Lutfey KE. Sources of Variation in Physician Adherence with Clinical Guidelines: Results from a Factorial Experiment. *J Gen Int Med.* 2007; 22:289–296.
14. Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. Comparison of Vignettes, Standardized Patients, and Chart Abstraction: A Prospective Validation Study of 3 Methods for Measuring Quality. *JAMA.* 2000; 283:1715–1722. [PubMed: 10755498]
15. Eva KW, Hatala RM, LeBlanc VR, Brooks LR. Teaching from the clinical reasoning literature: Combined reasoning strategies help novice diagnosticians overcome misleading information. *Med Educ.* 2007; 41:1152–1158. [PubMed: 18045367]
16. Choudhry NK, Fletcher RH, Soumerai SB. The relationship between clinical experience and quality of health care. *Ann Intern Med.* 2005; 142:260–273. [PubMed: 15710959]
17. Barrows HS, Norman GR, Neufeld VR, Feightner JW. The clinical reasoning of randomly selected physicians in general medical practice. *Clin Invest Med.* 1982; 5:49–55. [PubMed: 7116714]

Table 1

Number, proportion, and direction of changes in physicians' diagnostic opinions across the case presentation as a function of experience level (N=256 primary care physicians practicing in North or South Carolina in 2007/2008).

Experience level	Initial impression*	Final diagnosis same as initial impression	Final diagnosis changed to CHD after mid-point of case	Final diagnosis changed to non-CHD diagnosis after mid-point
Less experienced (N=128)	CHD (N=70)	48 (68.6%)		22 (31.4%)
	Not CHD (N=58)	23 (39.7%)	21 (36.2%)	14 (24.1%)
More experienced (N=128)	CHD (N=84)	72 (85.7%)		12 (14.3%)
	Not CHD (N=44)	19 (43.2%)	14 (31.8%)	11 (25.0%)

* The initial impression corresponds to the diagnosis given at the mid-point of the case presentation at which time the video was stopped and participants were asked to provide their primary diagnosis, rate their certainty in that diagnosis, and indicate the most important piece of additional information they would like to receive. After that point the video was played to completion. Reading across rows indicates how many participants in each experience group indicated CHD (or not) as their initial impression and how many of those retained their initial impression or changed their diagnosis during the second half of the case presentation.