Effect of a Diagnostic Suggestion on Diagnostic Accuracy and Identification of Clinical Features

VICKI R. LEBLANC, GEOFFREY R. NORMAN, and LEE R. BROOKS

Downloaded from http://journals.lww.com/academicmedicine by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdgGj2MwlZLeI= on 01/07/2025

The diagnoses initially entertained by physicians strongly influence identification of clinical features. When the diagnostic hypothesis is correct, it increases the likelihood that the correct features will be identified from the patient.[1,2] The diagnosis seems to suggest what to look for and where to look for it. In addition, the diagnosis can change the interpretation of physical characteristics. When considering an incorrect diagnosis, clinicians not only fail to identify the correct features, but they identify features that are not present.[3] Thus, the accuracy of a diagnostic hypothesis will influence the physicians' accuracy in identifying clinical signs.

This has serious implications for diagnostic efficacy, as errors in initial diagnostic hypotheses are not necessarily going to be corrected when physicians gather clinical information about the patient. Rather, the identification of clinical information will be biased in the direction of the erroneous diagnosis. This poses a challenge to medical education: How do we get clinicians, whether they are novices or experts, to correct themselves when they begin a clinical encounter with an incorrect diagnostic hypothesis? One proposed solution has been to instruct clinicians to avoid generating diagnostic hypotheses until all of the clinical information has been gathered, in hopes that they will be more thorough in their gathering and identification of features. However, this strategy is unlikely to be effective, as researchers have demonstrated that early hypothesis generation appears to be an ingrained way to proceed for novices and experts.[4] In situations where instructions to avoid diagnostic hypotheses were successful, with extreme novices, such instructions had detrimental effects on diagnostic accuracy and feature identification.[5]

Another proposed solution is to get clinicians to consider alternatives. In social psychology, it has also been demonstrated that individuals engage in selective hypothesis testing, that is, searching for evidence in support of the favored hypotheses while neglecting evidence supportive of alternate hypotheses.[6] When instructed to consider alternatives, individuals undertake a more exhaustive search, looking for evidence that will effectively test the likelihood of the various alternatives.[6] Thus, getting physicians to consider alternatives to the favored hypothesis might reduce clinical errors. If, when initially biased towards an incorrect diagnosis, they are reminded of alternative diagnoses by either a supervisor or a computer decision-aid system, they may reevaluate and reinterpret the clinical features in light of the alternative diagnoses. For such an intervention to be effective, clinicians need to be sensitive to information presented later in a clinical case. The two studies that have investigated whether clinicians attend to later information have found conflicting results. Bergus et al.[7] found that clinicians gave more weight to a prior medical history when it was presented last in a clinical case than when presented first. Conversely, Cunnington et al.[8] found that participants favored the diagnosis for which information was presented first in the clinical case.

The present study investigated whether the point at which a diagnosis is presented will influence identification of features and diagnostic accuracy. Medical students were shown head-and-shoulders photographs accompanied by short case histories and tentative diagnoses that biased them towards either the correct diagnoses or incorrect alternative diagnoses. Prior to reporting the clinical features or making any diagnostic decision in each case, they received the suggestion of a differential diagnosis (the correct one if initially biased towards the alternate one, and vice versa). If medical students attend to the early diagnosis only and ignore subsequent diagnostic suggestions, their diagnostic decisions and identification of features will favor the initially suggested diagnosis. However, if the students attend equally to diagnostic suggestions presented early and late in a clinical encounter, their diagnostic decisions and identification of features should not be influenced by the diagnosis that is suggested.

Methods

Participants. The entire second-year class of medical students at McMaster University's Medical Center (approximately 100 students) was solicited by e-mail toward the end of the academic year. A convenience sample of 20 students volunteered through informed consent, and they were paid for their participation; student academic measurements were not determined for this study. The students were 17–27 months into a three-year program, with six of the 20 having just started their clinical clerkships.

Materials. Ten head-and-shoulders photographs were selected from medical textbooks or clinicians' slide libraries and were considered by experts to be classic representations of the given diagnoses. Eight photographs served as the test stimuli and the other two were used in practice scenarios. The latter two scenarios were used to create the sense that the suggested diagnoses were plausible, and they had been correctly diagnosed by 90% of medical students in a previous study.[2]

Case histories biasing towards the correct diagnosis and a plausible but incorrect alternate diagnosis were generated for each test scenario. The correct case histories had been used in previous studies with these photographs.[2,3] The alternate case histories were generated by choosing a feature in the picture from which a plausible diagnosis could be constructed. In some scenarios, these features were caused by normal variation (i.e., tanned skin misinterpreted as jaundice, despite white sclerae of the eyes). In other scenarios the feature resulted from the misinterpretation of one of the cardinal features of the correct diagnosis (i.e., a moon-shaped face reinterpreted as facial edema). Case histories were generated by an expert in internal medicine and independently verified by another expert in internal medicine to ensure that they were plausible given the information present in the photographs.

For each diagnosis, a list of features was generated from an authoritative internal medicine textbook.[9] Two experts in internal medicine were asked to indicate which features of the correct diagnoses were present in the photograph and which features of the alternate diagnoses could be reported if participants misidentified either a correct feature or a characteristic caused by normal variation. Only the features for which both experts agreed were included in the analyses, resulting in a sum of 21 features of the correct diagnoses and 25 features of the alternate diagnoses over all the scenarios. There were averages of 2.6 correct features and 3.1 incorrect features per scenario.

Procedure. The students individually worked through the ten scenarios. Practice scenarios preceded the test scenarios. For each scenario, the procedure was the same. On a first page, the students were shown a head-and-shoulders photograph of a patient accompanied by either the correct or the alternate case history that included the tentative diagnosis. On a second page, they read the suggestion of a differential diagnosis (the alternate diagnosis when

the tentative one had been the correct diagnosis, and vice versa). They were then asked to write down any and all clinically important features observed in the photograph and to give likelihood ratings (on a scale of 0 to 100%) of the tentative diagnosis, of the differential diagnosis, of any self-generated diagnosis, and of a residual category. They were instructed that these represented all possible options and that their ratings should sum to 100%.

The students were divided equally into two counterbalancing orders; each order had four scenarios in which the initial bias was towards the correct diagnosis and four scenarios in which the initial bias was towards the alternate diagnosis. The order of presentation of the scenarios was randomized and held constant across students.

Measures. To assess the diagnostic decisions of the students, there were two measures of interest. (1) How often they concluded for the correct diagnosis; for each scenario, they were given a 1 if they gave the highest likelihood rating to the correct diagnosis and a 0 if they gave the highest likelihood rating to another diagnosis. (2) How often they concluded for the alternate diagnosis; for each scenario they were scored a 1 if they gave the highest likelihood rating to the alternate diagnosis and a 0 if they gave the highest likelihood rating to another diagnosis. There were also two measures of feature identification; (3) the percentage of the correct features and (4) the percentage of the alternate features reported by the students. These four measures were submitted to separate 2 × 8 mixed-design analyses of variance (ANOVA) with the condition (correct diagnosis suggested first; and alternate diagnosis suggested first) as the between-subjects variable and the scenarios as a repeated-measures variable. The scenarios were anticipated to be variable in difficulty and were included in the analyses to account for some of the variance in the measures.

Results

The hypotheses that the students would conclude more strongly in favor of the diagnosis that was presented first was confirmed; students concluded for the correct diagnoses more often when initially biased towards the correct diagnoses (78% of the time) than when initially biased towards the alternate diagnoses (35% of the time) ($F[1,20] = 73.67$, MSE = .11, $p < .01$). Alternatively, students concluded for the alternate diagnoses more often when initially biased towards the alternate diagnoses (60% of the time) than when initially biased towards the correct diagnoses (7% of the time) ($F[1,20] = 66.53$, MSE = .185, $p < .01$). There was some variability in the credibility of the alternate diagnoses, as indicated by a bias by case interaction for concluding for the alternate diagnosis ($F[7, 140] = 2.45$, MSE = .105, $p < .05$). In all the scenarios, the students concluded for the alternate diagnosis most often when it was suggested to them, but the sizes of the biasing effect differed across the scenarios.

The identification of the alternate features was biased towards the initial suggested diagnosis ($F[1,20] = 5.93$, MSE = 0.43, $p < .05$). Students identified more alternate features when initially biased towards the alternate diagnoses (21.4% or 5.4 features) than when initially biased towards the correct diagnoses (13.9% or 3.5 features). There was some variability in this effect, as indicated by a case-by-scenario interaction, $F(7, 140) = 2.19$, MSE = .044, $p < .01$. On some of the scenarios, the suggestion of an alternate diagnosis did not lead to an increase in the reporting of the incorrect features.

Such a preference for the initial bias was not found in the identification of correct features. The students identified as many of the correct features when initially biased towards the correct diagnoses (43.7% or 9.2 features) as when initially biased towards the alternate diagnoses (40.6% or 8.5 features) ($F[1,20] = .687$, MSE = .059, $p = .42$).

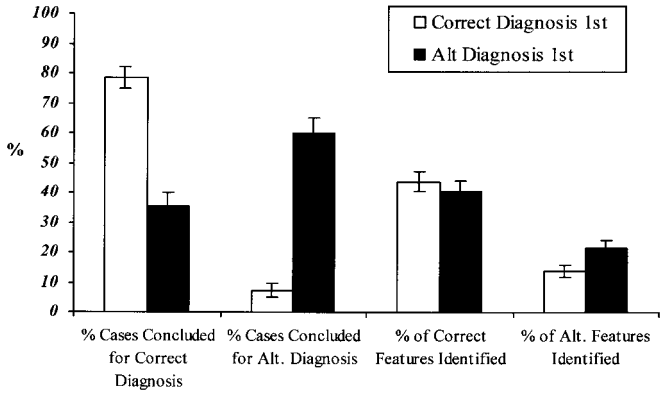

Figure 1. **Diagnostic conclusions and identification of features by medical students after reading diagnostic suggestions.**

Conclusion

The diagnostic decisions of the students were strongly in favor of the initial diagnostic suggestions. This is not surprising, given that the case histories presented in the scenarios were strongly suggestive of the initial diagnoses. Informally, these results were compared with those of a study by LeBlanc et al.,[3] which was conducted with a previous cohort of medical students from the same medical school and with the same materials. The instructions in the LeBlanc et al. study were similar to those of the present study, with the exception that students did not receive a subsequent suggestion of a different diagnosis after having been biased toward either the correct or the alternate diagnosis. In the previous study, students concluded for the correct diagnosis in 80% of the cases when biased towards the correct diagnosis and in 12% of the cases when biased towards the alternate diagnosis. In the present study, the subsequent suggestion of the alternate diagnosis, once students had initially been biased towards the correct diagnosis, did not lead to a decrease in diagnostic accuracy (78.4% vs. 80% in the previous study). The subsequent suggestion of the correct diagnosis, once students had been initially biased towards the alternate diagnosis, led to an increase in diagnostic accuracy (35.5% vs. 6.9% in the previous study). The increase in diagnostic accuracy resulting from reminding students of the correct diagnosis appears substantial. However, the accuracy level in this situation fails to be as high as when the correct diagnosis is considered early in the clinical case (35.5% vs. 80%).

Students reported as many correct features after the suggestion of the correct diagnoses, regardless of whether they were initially biased towards the correct or the alternate diagnoses. However, the time at which they considered the correct diagnoses had an impact on their identification of alternate features for some of the scenarios. A possible mechanism for this pattern of results is that a diagnosis's influence on feature identification is twofold. First, it serves as a "focus of attention," reminding clinicians of what features to look for and where. For example, the suggestion of stomach cancer will remind physicians to look for a lymph node in the superclavicular area. This focus of attention can take place either early in the case or after clinicians are already considering another diagnosis. Second, the diagnosis can influence the interpretation of the physical characteristics of the patient. For example, the tanned skin of a patient is more likely to be interpreted as jaundice if the diagnosis considered is liver cancer but more likely to be interpreted as normally tanned skin if the diagnosis being considered is one that does not involve the liver. This interpretation of the physical characteristics occurs early in the case and is not easily reinterpreted. In the present study, the correct features were features that were actually depicted in the photographs, while the alternate fea-

Downloaded from http://journals.lww.com/academicmedicine by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdgGj2MwlZLeI= on 01/07/2025

tures resulted from the misinterpretation of the physical characteristics. Thus, the suggestion of the correct diagnosis would focus the students' attention to those depicted correct features, and this influence can occur at any time in a clinical encounter. However, a later suggestion of a diagnosis would have a smaller impact on the alternate features, as the physical characteristics would already have been interpreted and they would be robust to reinterpretation.

The generalizability of these results is limited due to the use of highly visual stimuli. We predict, however, that diagnostic hypotheses will also have an impact on the identification of features from different types of medical stimuli, and plan such studies. For example, diagnoses considered by clinicians would have an impact on their decisions regarding what is important and relevant from a patient's verbal complaints.

While it is possible that our effects are limited to medical students with little medical experience, both residents and more experienced clinicians have been shown to be influenced by diagnostic suggestions in other studies,[1,3] suggesting that such effects are not restricted to novices.

The implication of this study for medical educators is that the suggestion of a correct diagnosis once students are initially biased towards an incorrect diagnosis will be limited in its effectiveness in reducing clinical errors. Students do attend to information that is presented later in a case, but they favor the early interpretation during the clinical encounter. Further studies should explore whether manipulations aimed at getting students to reconsider their initial clinical decisions will be more effective if done at the level of feature identification or at some intermediate level, such as the level of Bordage's semantic axes.[10] Alternatively, interventions aimed at changing the way that students are taught, by increasing exposure to real-world variations in case presentations, might increase the likelihood of the correct diagnoses' being considered early in a clinical encounter. Further research in this area should be conducted to guide educational changes.

Correspondence: Vicki R. LeBlanc, PhD, Center for Educational Research and Evaluation, Columbia University Health Sciences Division, 701 West 168th Street, HHSC LL12, New York, NY 10032. Reprints are not available.

References

1. Berbaum KS, Franken EA, Dorfman DD, et al. Tentative diagnoses facilitate the detection of diverse lesions in chest radiographs. Invest Radiol. 1995;21:532–9.
2. Brooks LR, LeBlanc VR, Norman GR. On the difficulty of noticing obvious features in patient appearance. Psychol Sci. 2000:11(2), 112–7.
3. LeBlanc VR, Brooks LR, Cunnington J, Norman GR. Diagnosis-specific observations are a source of errors in medical diagnostic tasks. Abstract presented at the annual meeting of the American Association of Medical Colleges, Washington, DC, May 1999.
4. Neufeld VR, Norman GR, Feightner JW, Barrows HS. Clinical problem-solving by medical students: a cross-sectional and longitudinal analysis. Med Educ. 1981; 15:315–22.
5. Norman GR, Brooks LR, Colle C, Hatala R. The benefit of diagnostic hypotheses in clinical reasoning: experimental study of an instructional intervention for forward and backward reasoning. Cogn Instruction. 1999;17:433–48.
6. Trope Y, Liberman A. Social hypothesis testing: cognitive and motivational mechanisms. In: Higgins ET, Kruglanski AW (eds). Social Psychology: Handbook of Basic Principles. New York: Guildford Press, 1996:239–70.
7. Bergus GR, Chapman GB, Gjerde C, Elstein AS. Clinical reasoning about symptoms despite preexisting disease: sources of errors and order effects. Fam Med. 1995;27:314–20.
8. Cunnington JPW, Turnbull JM, Regehr G, Marriott M, Norman GR. The effects of presentation order in clinical decision making. Acad Med. 1997;72(10, Suppl.): S40–S42.
9. Fauci AS, Braunwald E, Isselbacher KJ, et al (eds). Harrison's Principles of Internal Medicine. 14th ed. New York: McGraw–Hill, 1998.
10. Bordage G. Why did I miss the diagnosis? Some cognitive explanations and educational implications. Acad Med. 1999:74(10, Suppl):S138–S143.

Downloaded from http://journals.lww.com/academicmedicine by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdgGj2MwlZLeI= on 01/07/2025