Teaching Diagnostic Skills: Clinical Vignettes or Chief Complaints?

Nendaz, Mathieu; Raetzo, Marc-André Eugene; Junod, Alain; Vu, Nu Viet

Advances in Health Sciences Education **5:** 3–10, 2000.

Teaching Diagnostic Skills: Clinical Vignettes or Chief Complaints?

MATHIEU R. NENDAZ[1,2,4], MARC A. RAETZO[3], ALAIN F. JUNOD[1] and NU V. VU[2]

[1]*Department of internal medicine, Clinique de Médecine 1, Hôpital Universitaire, Geneva, Switzerland;*
[2]*Unité de Recherche et Développement en Education Médicale, Geneva, Switzerland;*
[3]*Groupe Médical, Onex, Switzerland;*
[4]*Current address: University of Illinois at Chicago, Department of Medical Education (MC 591), College of Medicine, 808 South Wood Street, Chicago, IL 60612-7309, USA, Phone: +1-708-358-0865; Fax: +1-312-413-2048; E-mail: mnenda1@uic.edu*

Abstract. Two formats of case presentation are traditionally used for teaching problem-solving skills: clinical vignette or chief complaint formats. While the first one is more commonly used, it does not completely reflect the actual problem-solving process during a real encounter, which may hamper the learners to integrate separately acquired data gathering skills into their reasoning process and affect their diagnostic performance in practice. The present study compared diagnostic accuracy when the reasoning stimulus was a case vignette containing all diagnostic information versus the patient's chief complaint only. Forty-two medical students, 53 residents and 60 general internists participated in the study. Diagnostic accuracy was significantly lower for the chief complaint format at the student, resident, and practitioner levels. Analysis of the data gathered in the chief-complaint format revealed that faulty diagnostic decisions resulted from a failure to gather critical data. The results suggest that data gathering techniques, semiology, and medical reasoning should be trained in association and that this effort should be pursued beyond medical school.

Key words: case-based learning, clinical reasoning, clinical skills, higher education, learning strategies, medical education, problem solving, science education, teaching methods

Introduction

Two formats are traditionally used to present medical problems for training diagnostic reasoning. The first one consists of exposing a medical case containing the patient's essential features through a vignette and asking the learners for the diagnosis. The second one consists of making only the patients' chief complaint available and asking the learners to acquire the additional information needed to make the diagnosis. The clinical vignette format is more practical, time-efficient, and therefore broadly used, especially in undergraduate courses and tutorials, graduate teaching rounds and conferences, and in most textbooks on medical diagnosis. However, this format does not completely reflect the actual clinical

process where the clinician has to determine the type of additional information needed from the initial complaint to make the appropriate diagnosis (Barrows et al., 1982; Elstein et al., 1978). Additionally, research has shown positive but weak correlations between the processes of data inquiry, data interpretation and integration, and diagnosis elaboration, raising the issue of interdependence between these skills (Vu and Steward, 1990). Therefore, relying predominantly on a vignette format to train clinical reasoning may not enable the learners to integrate their data acquisition skills, acquired previously and separately, into the reasoning process. This may hamper the acquisition of necessary skills required in practice and may affect diagnostic performance. The present study tested this hypothesis through the following questions: a) In a sample of subjects at diverse levels of medical experience, is diagnostic accuracy affected when the reasoning stimulus is a clinical vignette versus the patient's chief complaint only? b) If diagnostic accuracy is affected, what are the characteristics of the data gathered, in terms of relevance?

Methods

Two clinical problems were selected and presented, either in a clinical vignette format or in a chief complaint format. In the clinical vignette format, each participant read a short summary of a case containing clinical information and the results of initial ancillary exams and was asked to write down and rank three diagnostic hypotheses and the findings supporting each diagnosis. In the chief complaint format, the participants received a written statement of the patient's chief complaint and were asked to write on personal forms their requests for additional information on history, physical examination, ancillary exams, or therapeutic effects, one request at a time. Along with each request, each participant also mentioned which diagnostic hypothesis was tested by the question asked. For each individual request, an investigator wrote the corresponding answer from the actual patient's chart, before the participant decided on the next information to require. This process went on until the participants chose to stop, either because they believed to have reached the final diagnosis, or because they were not able to continue the diagnostic process.

The cases were derived from two real patient files and the vignettes contained only selected information relevant to the chief complaint. They were tested independently by five academic internists not participating in the study to ensure that only one unequivocal accurate diagnosis could be reached. These physicians also identified the critical findings of each vignette.

> *Case vignette one*: "A 26-year-old man comes to your office complaining of coughing up blood. On history, the patient indicates that breathing has become progressively difficult. He traveled through the United States four months ago and admits to using heroin regularly. He mentions having lost 10 kilograms during the past two months and denies any fever or chills. Except for a moderate asymmetry in the testicles, the physical examination reveals no fever and

no palpable lymph nodes. Cardiac, respiratory, and abdominal auscultation is normal. A CBC is normal and a chest X-ray reveals multiple 1- to 4-cm nodules in both lung fields."

Chief complaint one: "A 26 year-old man comes to your office complaining of coughing up blood."

Case vignette two: "A 22 year-old woman comes in, complaining of coughing that started four weeks ago after she had "a bout of flu"; this cough is persistent and dry. She also has a mild frontal headache and a moderate morning sputum production. She denies any fever or sinus pain. Her physical examination is normal."

Chief complaint two: "A 22 year-old woman comes in, complaining of coughing."

The correct diagnosis was a metastatic testicular cancer in case one, and chronic sinusitis in case two. Data collection took place during scheduled student-resident's teaching rounds of the department of internal medicine and continuing medical education conferences at a private outpatient clinic. The participants in these conferences were told at the beginning of the session that the results of the exercise would be used for a study on clinical reasoning. All 155 physicians or students who attended the teaching rounds and conferences agreed to participate. Based on this convenience sampling, the participants regrouped 42 medical students of a six-year curriculum (12 fourth-year and 30 sixth-year), 53 internal medicine residents, and 60 general internists. Sixty-four subjects completed the vignette format (34 with case one, 30 with case two) and 91 the chief complaint format (39 with case one, 52 with case two).

The first author, an internist, tallied the written answers about diagnoses, according to a pre-established classification by five internists not participating in the study. For the vignette format, the diagnostic was considered correct if the first of the three diagnostic hypotheses matched the patient's diagnosis. For the chief complaint format, the diagnosis proposed at the end of the written protocol was classified as correct if it matched the patient's diagnosis and as not correct if it did not. The absence of any final diagnosis was considered as incorrect diagnosis. In addition for this format, the first author and another physician not participating in the study tallied the number of critical and less critical findings, which were available in the vignettes and requested by the subjects. To determine the reliability of their evaluation, 15% of the protocols were evaluated independently by both raters. Concordance in the raters' evaluation was found for more than 95% of the protocols. There was no systematic error and the few discrepancies were due to the omission of less critical findings by one of the raters.

Pearson Chi-square test was used to analyze the proportions of subjects who listed the correct diagnosis by case, format, and level of experience; t-tests with Bonferroni's correction for multiple tests and analysis of variance were used to analyze the number of findings. An alpha value of less than 0.05 was considered significant.

Table I. Ratio and percentage (in parentheses) of subjects having listed the correct diagnosis by case, case format, and levels of experience

	Vignette format n = 64			Chief complaint format n = 91			Total n = 155
	Case 1	Case 2	Total	Case 1	Case 2	Total	
Students	4/13 (31)	4/9 (44)	8/22 (36)	2/9 (22)	0/11 (0)	2/20 (10)	10/42 (24)
Residents	9/10 (90)	8/11 (73)	17/21 (81)	6/13 (46)	9/19 (47)	15/32 (47)	32/53 (60)
General Internists	11/11 (100)	10/10 (100)	21/21 (100)	11/17 (65)	12/22 (54)	23/39 (59)	44/60 (73)
Total (n = 155)	24 (71)	22 (73)	46 (72)	19 (49)	21 (40)	40 (44)	86 (55)

Results

Overall, 55% of the participants made a correct diagnosis (see Table I). There were no significant differences for diagnostic accuracy between cases one and two within the vignette format (71% and 73%, Fisher's Exact Test, $p \leq 1.00$) and the chief complaint format (49% and 40%, Fisher's Exact Test, $p = 0.523$). Consequently, the rest of the analyses were based on the combined performances on the two cases.

The proportion of participants who elicited the correct diagnosis was higher in the clinical vignette format than in the chief complaint format (72% versus 44%; $\chi^2(1) = 11.858$, $p = 0.001$). This difference was present for the students (36% versus 10%; $\chi^2(1) = 4.014$, $p = 0.04$), the residents (81% versus 47%; $\chi^2(1) = 6.155$, $p = 0.013$) and general internists (100% versus 59%; $\chi^2(1) = 11.748$, $p < 0.001$). Regardless of formats, the subjects' diagnostic accuracy increased significantly with clinical experience (from 24% among students to 73% among general internists; $\chi^2(2) = 25.313$, $p < 0.001$).

In the chief-complaint format, the subjects who found the accurate diagnosis obtained significantly more total and critical data available in the vignettes than the subjects who did not get the correct diagnosis. This was true for both cases, except for the number of total findings in case one (see Table II).

When compared to the number of findings contained in the vignettes, subjects in the chief-complaint condition gathered overall a larger amount of data for case one (19 versus 12, $t = 8.58$, $df = 38$, $p < 0.001$) and case two (22 versus 11, $t = 11.77$, $df = 51$, $p < 0.001$). No differences were found across the levels of experience ($F(2,88) = 0.05$, $p = 0.95$). However, consistently, the total and critical findings of the vignettes were only partially obtained by the subjects in the chief-complaint format, as shown in Table III.

Discussion

Based on two clinical cases, subjects in a condition where only the chief complaint was available had a lower diagnostic accuracy than those in a condition where all

Table II. Mean (standard deviation) of critical and total findings which were available in vignettes and were requested by subjects in the chief-complaint condition, by diagnostic accuracy

		Diagnosis found		t	df	p*
		Yes	No			
Case 1	Critical	2.89 (0.32)	1.75 (0.44)	9.32	37	<0.001
	Total	6.95 (1.43)	6.35 (1.6)	1.23	37	≤1.00
Case 2	Critical	2.86 (0.85)	2.10 (0.98)	2.89	50	0.03
	Total	6.33 (1.46)	5.06 (1.53)	2.99	50	0.02
All cases	Critical	2.88 (0.65)	1.96 (0.82)	5.76	89	<0.001
	Total	6.63 (1.46)	5.57 (1.66)	3.17	89	0.01

*Corrected for multiple comparisons (Bonferroni).

Table III. Mean (standard deviation) of critical and total findings which were available in vignettes and were requested by subjects in the chief-complaint condition

		Vignette	Chief complaint	t*	df	p†
Case 1	Critical	3	2.31 (0.69)	6.22	38	<0.001
	Total	12	6.64 (1.53)	21.87	38	<0.001
Case 2	Critical	5	2.40 (1.00)	18.81	51	<0.001
	Total	11	5.58 (1.61)	24.24	51	<0.001

* One-sample t-test.
† Corrected for multiple comparisons (Bonferroni).

the pertinent patient's findings were provided in a case scenario. To further explain these results, whether and under which conditions the findings available in the vignettes were gathered in the chief-complaint format was analyzed. In order for the diagnosis to be derived unequivocally, all the findings in the vignettes were chosen to be reasonably relevant to the chief complaint. Hence, subjects in the chief-complaint format and with the correct diagnosis requested more critical data contained in the vignettes than their counterparts who did not derive the correct diagnosis. Results of Table II are consistent with the nature of the vignettes and confirmed that all the data presented were relevant at various degrees to the chief complaint, but that not all the information was needed to find the diagnosis (for example, case one contained some distractors).

In the chief-complaint format, subjects of different levels of experience gathered more findings than those provided in the vignettes and had therefore no problems, real or related to the study conditions, to obtain information. However, as displayed in Table III, they all had difficulty gathering the relevant data presented in the vignettes. As these data were either critical or relevant to the diagnosis, this

suggests that the subjects in the chief-complaint format had a failure to collect critical and relevant information and that this lack of effective data collection, in turn, decreased diagnostic accuracy. Subjects of the present study, regardless of medical experience, had more difficulty in determining which useful findings to collect in order to make the diagnosis, rather than problems in interpreting available information, as previously described by others (Gruppen et al., 1991).

What could be the reasons why relevant information was not sought in the chief-complaint format? The subjects in the vignette condition were able to recognize and interpret critical information to derive the correct diagnosis when all the data were available. Therefore, a lack of knowledge is not sufficient to account for the between-format differences in diagnostic accuracy, even though it may explain the differences due to medical experience within the vignette format. In the chief complaint condition, subjects acted as if they were missing a trigger enabling them to ask pertinent questions. Bordage and colleagues (Bordage, 1994; Chang et al., 1998) showed that diagnostic accuracy improved with early and thorough problem representation, defined by the degree of abstraction from the actual findings. The difficulty in gathering appropriate data found in the present study may have been due to a failure in building the initial mental representation of the patient's problem, before generating and testing multiple hypotheses. Having the correct hypothesis in mind early during the clinical work-up has been shown, not only to increase final diagnostic accuracy (Barrows et al., 1982), but also to improve the identification of visual features of a case (Norman et al., 1996). In the present study, it is possible that the failure to have the correct hypothesis in mind while gathering information may have also contributed to the lack of obtaining pertinent data. More research is clearly needed to define whether this phenomenon, observed with visual information, would also occur with more abstract verbal information.

The decrease in diagnostic accuracy related to the format was the strongest at the student level, when expressed in relative values, but was larger in absolute difference with increasing medical experience (Table I). While the very low values of the student proportion correct in the chief-complaint format may explain the relative figures, the increase in absolute difference with medical experience needs further explanation. First, as the recognition of the diagnoses illustrated by the vignettes depended largely on the knowledge base related to the cases, this format might have favored physicians with greater experience, rather than students with less knowledge. For this reason, the diagnostic accuracy in the vignette condition was very high for residents and practitioners. Second, while the main outcome of this study was diagnostic accuracy, there is evidence that the general practitioners might have solved the cases with a slightly different perspective than a pure diagnostic one. The protocols revealed that the practitioners were more considering the practical aspects of the management, rather than pursuing a pure intellectual diagnostic exercise. For example in case one, as soon as a systemic condition with massive pulmonary involvement was discovered, the patient was often "admitted" to the hospital for further management, without additional diagnostic effort. For

case two, once serious conditions, such as infection or asthma, had been ruled out, the patient was "referred" to a specialist after a symptomatic treatment had no effect. Third, although the chief complaint of the second case was representative of the problems a generalist meets in practice (chronic cough), its diagnosis (chronic sinusitis) may have been too specialized, even if non-specialists can generally manage uncomplicated cases. Fourth, the chief complaint condition used materials that simulated encounters without providing visual and perceptual information. These approaches have, however, been recognized as important aspects of the diagnostic process of physicians (Cox, 1996; Friedman et al., 1994; Norman et al., 1996). For all these reasons, the diagnostic ability of general internists might have been underestimated in the chief-complaint format. This underestimation, along with a high performance in the vignette condition may explain the higher than expected difference between both formats in this group.

The present study used only two cases, which prevents generalization of results. A prospective study with 10 or more cases using standardized patients to test the open-inquiry approach and including analyzes of thought processes would be enlightening to better understand the role of early problem representation and hypothesis generation on the ability to gather pertinent data. Assessing the impact of teaching efforts aiming at enhancing problem representation on the effectiveness of data gathering and on diagnostic performance are other issues to be considered in future studies.

These results suggest important implications for teaching and learning diagnostic reasoning and improving data gathering skills. In the majority of medical schools, interviewing and physical examination techniques, the semiology of diseases, and medical reasoning are taught separately, at different times during the curriculum and in different settings. Therefore, students may be able technically to perform interview and physical examination and to recite or recognize lists of signs and symptoms related to medical conditions, but prove unable to find and integrate this information in practice. As the quality of data gathering seems tightly linked to specific steps of the reasoning process, such as initial problem representation and hypothesis generation, how to obtain and organize the information should therefore not be trained in dissociation from its diagnostic application if one wants to favor the elaboration of usable "illness scripts" (Schmidt et al., 1990) in the learners' minds and to improve their skills in patient work-ups (Evans et al., 1996). Additionally, courses, conferences, and most so-called "practice-oriented" diagnostic textbooks should go beyond simply drawing static pictures of different illnesses or listing lengthy differential diagnosis about each finding. They would probably be more effective if they strove to offer for each chief complaint an inquiry strategy based on how critical each piece of information is to the related diagnostic hypotheses. These efforts should aim not only undergraduate training, but also graduate and continuing medical education.

Acknowledgements

We thank the students, faculty members, and practitioners who participated in the study and the faculty members who kindly accepted to review the cases. We also acknowledge Professors Georges Bordage, MD, Ph.D. and Reed G. Williams, Ph.D., University of Illinois at Chicago, Department of Medical Education, for their valuable comments on earlier versions of this manuscript. This paper was presented in part at the 8th Ottawa International Conference, Philadelphia, July 1998.

References

Barrows, H.S., Norman, G.R., Neufeld, V.R. & Feightner, J.W. (1982). The clinical reasoning process of randomly selected physicians in general medical practice. *Clinical and Investigative Medicine* **5**: 49–56.

Bordage, G. (1994). Elaborated knowledge: A key to successful diagnostic thinking. *Academic Medicine* **69**: 883–885.

Chang, R.W., Bordage, G. & Connell, K.J. (1998). The importance of early problem representation during case presentations. *Academic Medicine* **73**(10 Suppl): S109–S111.

Cox, K. (1996). Teaching and learning clinical perception. *Medical Education* **30**: 90–96.

Elstein, A.S., Shulman, L.S. & Sprafka, S.A. (1978). *Medical Problem Solving: An Analysis of Clinical Reasoning*. Cambridge, MA: Harvard University Press.

Evans, B.J., Coman, G.J. & Goss, B. (1996). Consulting skills training and medical students' interviewing efficiency. *Medical Education* **30**: 121–128.

Friedman, C.P., McNutt, R.A., de Bliek, R., Evans, A.T., Wallsten, T.S. & Martz, J.M. (1994). Visual information and the diagnosis of chest pain. *Academic Medicine* **69**(10 Suppl): S28–S30.

Gruppen, L.D., Wolf, F.M. & Billi, J.E. (1991). Information gathering and integration as sources of error in diagnostic decision making. *Medical Decision Making* **11**: 233–239.

Norman, G.R., Brooks, L.R., Cunnington, J.P.W., Shali, V., Marriott, M. & Regehr, G. (1996). Expert-novice differences in the use of history and visual information from patients. *Academic Medicine* **71**(1 Suppl): S62–S64.

Schmidt, H.G., Norman, G.R. & Boshuizen, H.P. (1990). A cognitive perspective on medical expertise: Theory and implication. *Academic Medicine* **65**: 611–621.

Vu, N.V. & Steward, D.E. (1990). Variant accuracy in methods of assessing clinical performance. In W. Bender, R.J. Hiemstra, A.J.J.A. Scherpbier & R.P. Zwierstra (eds.), *Teaching and Assessing Clinical Competence*, pp. 190–194. Groningen: BoekWerk Publications.