ORIGINAL CONTRIBUTION
decision analysis;
medical education

# Medical Problem Solving and Uncertainty in the Emergency Department

David P Sklar, MD, FACEP
Mark Hauswald, MD
David R Johnson, MD
Albuquerque, New Mexico

From the Division of Emergency Medicine, University of New Mexico School of Medicine, Albuquerque.

Received for publication October 19, 1990.
Revision received March 1, 1991.
Accepted for publication May 1, 1991.

Address for reprints: David P Sklar, MD, FACEP, The University of New Mexico, School of Medicine, Division of Emergency Medicine, Albuquerque, New Mexico 87131.

Study objective: *To compare the diagnostic processes of experienced emergency physicians with those of novices.*

Design: *Prospective, convenience sample of patients.*

Setting: *Emergency department of a county university medical center in a large southwestern urban community.*

Participants: *Experienced emergency physicians (attending and senior residents) and novice clinicians (junior residents and senior medical students).*

Interventions: *Participants developed initial diagnostic impressions after reviewing the chief complaint, nurse triage notes, and vital signs. Tests were then selected, and a final diagnostic impression was identified after results were known. Clinicians also marked a visual analog scale corresponding to their estimate that each diagnostic possibility was correct.*

Results: *Experienced physicians increased their certainty more than novices (P = .014). They deviated from a standard history-physical-laboratory sequence more often than novices (P = .008).*

Conclusion: *Expertise in medical decision making is characterized by a moderate initial level of certainty concerning a diagnosis that significantly increases as the experienced clinician follows a flexible strategy of testing to arrive at a final diagnosis. [Sklar DP, Hauswald M, Johnson DR: Medical problem solving and uncertainty in the emergency department. Ann Emerg Med September 1991;20:987-991.]*

## INTRODUCTION

Differences in diagnostic decision making between expert and novice physicians reflect adaptation and refinement of basic skills and knowledge by the expert through experience. Both expert and novice are faced with myriad diagnostic possibilities and inadequate data of questionable value. Knowledge of differences in how experts and novices approach clinical diagnostic problems and make decisions despite uncertainty might allow medical educators to improve or accelerate the development of expertise. Differences have been examined for novice and expert chess players,[1] physics problem solvers,[2] and radiologists,[3,4] and it has been suggested that experts are capable of greater pattern recognition because of their stored experience and that this explains much of the difference in performance between experts and novices.

Medical diagnostic decision making is characterized by early hypothesis development,[5] selection of tests to confirm hypotheses,[6] recognition of patterns,[3] refinement of hypotheses,[7,8] and eventual selection of a diagnosis.[9,10] Recent attempts to model the process have included the use of Bayes theorem[7] and decision analysis.[11,12] Much research on diagnostic decision making has focused on the verbal descriptions of expert clinicians in the process of diagnostic decision making[10,13] or on simulated cases given to both novice and expert clinicians.[5,14] However, clinicians do not behave as many decision rules would predict,[8,10] and little has been learned in the clinical setting about the actual diagnostic decision-making process.

We chose to analyze diagnostic decision making in our emergency department because this clinical setting has several characteristics that simplify the analysis. In particular, the limited time available for patient eval-

**FIGURE 1.** *Data collection form.*

Please fill out this form as you go through the process of caring for the patient. Its purpose is to help us understand medical decision making in an acute care setting. Fill this out only if you are the primary caregiver.

Name _____ Date _____

Mark Level: ___ Medical Student ___ HO1 ___ HO2 ___ HO3 ___ Attending

Name of Patient _____ Patient Number _____

After review of the patient's chief complaint and vital signs if available on the chart your diagnosis is: (mark probability along line)

1. _____ No Chance 0 _____ 1 Certain
2. _____ No Chance 0 _____ 1 Certain
3. _____ No Chance 0 _____ 1 Certain

The first test chosen (a test can be history-physical-laboratory including ECG or radiograph) was _____

The second test (a test can be history-physical-laboratory including ECG or radiograph) was _____.

Subsequent tests were _____. My final diagnosis was: (mark probability along line)

_____ No Chance 0 _____ 1 Certain

**1**

**FIGURE 2.** *Final diagnoses of cases.*

Myocardial infarction/angina (10)
Renal stone (8)
Flu/upper respiratory infection (7)
Pneumonia (6)
Ulcer (5)
Appendicitis (4)
Gastritis/gastroenteritis (4)
Diabetic ketoacidosis (4)
Congestive heart failure (3)
Pleuritis (3)
Asthma (3)
Strep throat/pharyngitis (3)
Urinary tract infection/pyelonephritis (3)
Costochondritis (3)
Cellulitis (2)
Diabetes/nonketotic (2)
Deep venous thrombosis (2)
Reflux esophagitis (2)
Pelvic inflammatory disease (2)
Dysmenorrhea (2)
Corneal abrasion/foreign body (2)
Facial fracture (2)
Cerebrovascular accident (2)
Rib fractures (2)
Gallstones (2)
Other diagnoses (1 each) (41)

**2**

uation and the need to complete the diagnostic process before discharge compress the process into a relatively short period of time. We looked at differences in the generation of initial diagnostic impression, differences in the certainty of the leading initial diagnostic impression, and the certainty of the final diagnosis as well as differences in deviation from a standard history-physical-laboratory sequence between experienced and novice clinicians in the ED. We hypothesized that experienced clinicians would develop better initial diagnostic impressions, be more certain about their leading and final diagnoses, and deviate from a standard history-physical-laboratory sequence more often than novice clinicians.

## MATERIALS AND METHODS

A data collection sheet was prepared (Figure) to facilitate the identification of diagnostic impressions and the physician's degree of certainty. Participants were told that they were part of a study of clinical decision making and were asked to fill in their initial three leading diagnostic impressions immediately after examining the patient's chart. Compliance with the protocol was assessed by researchers when data sheets were submitted, and only those in which the protocol was followed were tabulated. The chart contained only a chief complaint, brief triage notes, and vital signs obtained by an ED nurse. This was used as the starting point because it corresponds to the way in which emergency physicians actually get their initial information about patients.

After filling in the leading three diagnostic impressions on the data sheet, the participants estimated their level of certainty that each diagnosis would turn out to be the correct final diagnosis by marking a visual analog scale[15] between the extremes of "no chance" and "absolute certainty." The distance from the beginning of the line to the mark was measured, and the ratio of the measured length to the total length of the line was calculated to give a certainty score between 0 and 1. Participants then sequentially selected three sets of tests in any order they thought appropriate. When the results of one test became available, they then selected another test. Tests included history and physical as well as ECG, radiography, and laboratory analysis. After the results of the tests were available, the participants formulated a final diagnostic impression and estimated a level of certainty by marking along a visual analog scale as noted above. The final diagnostic impression was made before the patient left the ED. Participants were grouped into experienced clinicians (attending physicians and senior residents, who had at least ten months' and an average of several years' ED experience) and novices (junior residents and senior medical students, who had less than three months' emergency medicine experience).

Although increased experience tends to imply expertise within the context of a university emergency medicine residency training program, as demonstrated by increasing level of responsibility for critically ill or injured patients, we refer to our attendings and senior residents as experienced clinicians rather than experts throughout this article to avoid any confusion that the term "expert" might imply.

The study took place between July 1989 and February 1990. Participation was voluntary, and participants were allowed to select any patient in whom the diagnosis was not clear from the triage note. Fifty-seven percent of the eligible experienced clinicians participated, and 65% of eligi-

**TABLE.** *Analysis of differences between novices and experienced clinicians*

| | Novice Clinicians' Cases (N = 81) | Experienced Clinicians' Cases (N = 48) | Statistical Significance of Difference | Level of Significance After Bonferroni Adjustment |
|---|---|---|---|---|
| Final diagnosis listed as leading initial diagnosis | 30 (37%) | 22 (46%) | $\chi^2$ = .638<br>$P$ = .428 | .012 |
| Final diagnosis included in list of three initial diagnostic impressions | 47 (58%) | 28 (58%) | $\chi^2$ = .000<br>$P$ = 1.0 | |
| Deviation from "conventional" history-physical-laboratory sequence | 26 (32%) | 28 (58%) | $\chi^2$ = 7.47<br>$P$ = .0065 | |
| Severity (admission rate) | 36 (44%) | 21 (44%) | $\chi^2$ = .01<br>$P$ = 1.0 | |
| **Certainty (0 to 1 ± SEM)** | | | | |
| Mean score for each provider used as a sampling unit | | | | |
| | **Novices (N = 40)** | **Experienced Clinicians (N = 8)** | | |
| Initial certainty of leading diagnostic impression | .66 ± .02 | .58 ± .03 | $P$ = .09 (Wilcoxon) | .017 |
| Change in certainty | .17 ± .03 | .33 ± .02 | $P$ = .014 (Wilcoxon) | |
| Certainty of final diagnostic impression | .83 ± .03 | .91 ± .03 | $P$ = .49 (Wilcoxon) | |

ble inexperienced clinicians participated. Data sheets were collected daily. Changes in certainty between initial and final diagnostic impression within groups were analyzed by a paired $t$ test. Differences in certainty between groups were compared with a Wilcoxon test. Deviation from standard workup, differences in initial and final diagnostic impression, and admission rate (as a gauge of severity of cases) were assessed with a $\chi^2$ test. A significance level of $P < .05$ was used, and a Bonferroni correction was made for multiple comparisons.

## RESULTS

One hundred thirty-nine patient cases were collected. Ten of these were excluded because protocol violations occurred or vital information was missing, leaving 129 patients on whom the analysis was performed. Certainty analysis was completed on 128 of these because certainty data were missing from one encounter in which all other data were present. There were 48 encounters by eight experienced clinicians (mean, six; range, two to 11 cases per clinician) and 81 encounters by 40 novices (mean, two; range, one to five). Final diagnoses in these cases are shown (Figure 2).

Novice clinicians were essentially as good as experienced clinicians at developing an initial differential diagnosis. There was no significant difference between novice and experienced clinicians in either including the final diagnosis among their initial three diagnostic impressions or initially listing it first (Table).

Although both experienced and novice clinicians demonstrated significant increases in certainty ($P < .001$) between initial and final diagnosis, the difference in certainty change between groups was also significant ($P = .014$) (ie, experienced clinicians increased their diagnostic certainty by a greater margin than did novices). There was also a tendency of inexperienced clinicians to be more certain of their initial diagnosis than the experienced clinicians ($P = .08$ Wilcoxon [Table]). Mean scores for each provider were used as the sampling unit to adjust for the different number of cases per provider.

The choice of tests was also very different between groups; experienced clinicians deviated from the "conventional" history-physical-laboratory sequence almost twice as often as did novices (Table). Severity as measured by admission rate was similar in the groups (Table).

## DISCUSSION

In this study, clinicians were able to develop an initial set of diagnostic impressions from information derived from a triage note, chief complaint, and vital signs that included the final diagnosis more than half of the time. In 40% of cases, the initial leading diagnosis became the leading final one. This occurred before the physician actually saw the patient. Previous studies have also called attention to the rapidness of hypothesis development after presentation of a problem.[3,5]

We could not detect a difference between experienced and novice clinicians in their initial diagnostic impressions. With our sample size, we had 82% power to detect a 26% difference between experienced and novice clinicians. The 9% difference found in our small study was neither statistically nor clinically significant. However, the failure to detect a difference also may have been a result of the influence of experienced ED nurses, technicians, or physicians on the initial diagnostic impression of novices. Such influence could be subtle, consisting of nonverbal cues or prompting for orders for particular tests that might suggest certain diagnostic possibilities. We did not specifically forbid such interaction because just like the triage notes and vital signs, it is part of the environment of the emergency physician and forbidding contact would produce an artificial situation. Others have also suggested no difference in development of differential diagnoses between novice and experts.[5]

It is also possible that experts and novices do not differ in their ability to develop an accurate initial impression because the list of possibilities based on a presenting complaint is finite and is stressed during basic medical education. The difference may come in the ability of the expert to assess accurately the probabilities of the various diagnostic choices and then select tests that increase the probability of one diagnosis to a point at which the clinician is willing to settle on a diagnosis and begin treatment.

A major difference between experi-

enced and novice clinicians in this study was in the selection of tests. Experienced clinicians did not follow the order of history-physical-laboratory taught to medical students. Such divergence is comparable to the scanning patterns of expert radiologists. Expert radiologists either rapidly recognize an abnormality or follow a slower, more complex scan compared with the more systematic but slower reading of novices.[3] In our study, experienced clinicians obtained an ECG first in seven cases. Early use of an ECG in the diagnosis of chest pain is consistent with the decision rule of Goldman et al[16] and follows from the mathematical models of decision making based on Bayes theorem. Because a positive ECG makes a myocardial infarction extremely likely, no other tests are necessary before deciding to admit the patient to the hospital.[17] In another seven cases, experienced clinicians chose radiographs, urinalyses, or urine pregnancy tests as their initial tests, presumably for the same reason – a positive test would greatly decrease the possible diagnoses.

Despite most medical school teaching, Bayes theorem suggests that if physicians wish to maximize their use of tests, they should select tests with high likelihood ratios to increase the odds of disease past a threshold of uncertainty.[18] History[16] and physical examination[7,19] are not always helpful in this regard. A flexible strategy that optimizes test benefit for each case would therefore be appropriate. Laboratory tests may be particularly likely to be chosen by emergency physicians because their major disadvantage, increased cost, is not borne by the physician but rather the patient and/or institution. The major limiting factor is a busy ED (ie, physician time is conserved if the relatively time-consuming history and physical are deferred). This strategy will work as long as the laboratory tests chosen are sufficiently sensitive to eliminate the need for a detailed history and physical. Poor selection of laboratory tests may have the opposite effect and actually increase the need for physician time. Experienced clinicians may prefer laboratory testing in certain cases, both because their time is at more of a premium and because their experience leads them to a more effective choice of tests compared with novice clinicians.

The certainty surrounding a diagnosis in the ED is a vexing but rarely discussed issue. Experienced clinicians tended toward less certainty of their initial diagnoses than novices. After the initial assessment, the increase in certainty by experienced clinicians was significantly greater than for less experienced clinicians. The novice clinicians appeared to overestimate the certainty of their initial diagnostic impressions. This may be because medical education stresses "classic" descriptions of clinical entities and makes little reference to their prevalence within a typical outpatient population.[14] Experienced physicians are probably more aware that unusual presentations of common diseases are more common than the classic presentation of exceedingly rare diseases. Emphasis on characteristic features of a problem rather than on the prevalence in a population has been identified as a major cause of errors in judgment.[20] Novices may not be able to balance the emphasis on characteristic descriptions of disease with experience of the frequencies of occurrence of the diseases in practice.

Dolan et al have shown that clinicians' subjective estimates of the prior probability of disease are quite variable and inaccurate.[21] Our study suggests that less experienced clinicians differ from more experienced clinicians in this regard. Implications for medical education are that students and residents should be taught an accurate estimation of the initial probability of disease and a flexible approach to testing for disease in which history and physical examination data can be critically compared with other laboratory data and the time and sequence of performing the tests have some relation to the usefulness that can be expected from them. As students and residents progress through training, they should refine their understanding of initial probability of disease based on their experience of the various contexts of a patient's presentation. As they develop expertise, they should also be able to select and analyze tests to increase the certainty of their diagnostic impressions.

There are several potential sources of bias in our study. First, the case selection in the ED by the subjects meant that a nonrandom set of cases was evaluated. This was necessary to ensure that cases with diagnostic questions were used. Typically, patients with obvious lacerations or fractures were excluded. However, the possible lack of uniformity in case selection between groups must be considered when interpreting the results of our study. Future studies that could more narrowly select cases might allow for a more specific comparison of diagnostic behaviors between groups that were not possible in our study.

It is also possible that subtle hints were made to medical students or interns whose initial evaluation differed strongly from that of the nurses. Such hints may have modified testing behavior. Also, because the nurse's triage notes were often based on a short history and physical examination, it would be inaccurate to say that tests were ordered without any history or physical examination. Although the physician may not have performed the history or physical examination before ordering the radiograph or laboratory test, the triage nurse did a rudimentary evaluation that included some history and physical examination.

The small number of experienced clinicians participating meant that individual variations in numbers and types of cases selected could affect the means generated in these groups. This small number of participants was a function of the size of our educational program and could not be changed easily. We have attempted to limit some of the possible effects of this area of bias by analyzing our certainty data using the mean scores of each participant. The difference in initial certainty between experienced and novice clinicians did not reach statistical significance with this analysis, although it was highly significant when all 128 cases were used. This suggests that a few participants may have strongly influenced this parameter, and the results should be viewed with this in mind.

## CONCLUSION

Experienced clinicians in the ED are initially less certain of their leading diagnostic impressions than are novice clinicians, but through testing their certainty increases more than that of novices. Experienced clinicians deviate from the conventional history-physical-laboratory sequence

more than novice clinicians as they increase their diagnostic certainty.

## REFERENCES

1. de Groot AD: *Thought and Choice in Chess.* The Hague, Moulton, 1965.

2. Larkin J, McDerott J, Simon DP: Expert and novice performance in solving physics problems. *Science* 1980; 208:1335-1342.

3. Christensen EF, Murry RC, Hulland K, et al: Effect of search time on perception. *Radiology* 1981;138:361-365.

4. Kundel HL, La Follette PS: Visual search patterns and experience with radiological images. *Radiology* 1972;103:523-528.

5. Elstein AS, Shulman LS, Spafka S: *Medical Problem Solving: An Analysis of Clinical Reasoning.* Cambridge, Massachusetts, Harvard University Press, 1978.

6. Sox HC: Probability theory in the use of diagnostic tests. *Ann Intern Med* 1986;104:60-66.

7. Hampton JR, Harrison MJG, Mitchell JRA, et al: Relative contributions of history-taking, physical examination, and laboratory investigation to diagnosis and management of medical outpatients. *Br Med J* 1975;2: 486-489.

8. Wolf FM, Gruppen LD, Gilli JE: Differential diagnosis and competing – Hypotheses heuristic. *JAMA* 1985;253:2858-2862.

9. Pauker SG, Kassirer JP: The threshold approach to clinical decision making. *N Engl J Med* 1980;302:1109-1117.

10. Moskowitz AJ, Kuipers BJ, Kassirer JP: Dealing with uncertainty, risks and tradeoffs in clinical decisions. *Ann Intern Med* 1988;108:435-449.

11. Schwartz WB, Gerry GA, Kassirer JP: Decision analysis and clinical judgment. *Am J Med* 1973;55:459-472.

12. Kassirer JP, Moskowitz AJ, Lau J, et al: Decision analysis: A progress report. *Ann Intern Med* 1987;106: 275-291.

13. Kassirer JP, Gerry GA: Clinical problem solving: A behavioral analysis. *Ann Intern Med* 1978;89:245-255.

14. Benbassett J, Bachar-Bassan E: A comparison of initial diagnostic hypothesis of medical students and internists. *J Med Educ* 1984;59:951-956.

15. Aitken RCB: Measurement of feelings using visual analogue scales. *Proc Roy Soc Med* 1969;62:989-993.

16. Goldman L, Cook EF, Brand D, et al: A computer protocol to predict myocardial infarction in emergency department patients with chest pain. *N Engl J Med* 1988;318:797-803.

17. Fesmire FM, Percy RF, Wears RL: Initial EKG in Q wave and non-Q wave myocardial infarction. *Ann Emerg Med* 1989;18:741-746.

18. Sox HC, Blatt MA, Higgins MC, et al: *Medical Decision Making.* Boston, Butterworth, 1988.

19. Fitzgerald F: Physical diagnosis versus modern technology – A review. *West J Med* 1990;152:372-382.

20. Tversky A, Kahneman D: Judgment under uncertainty: Heuristics and biases. *Science* 1974;185: 1124-1131.

21. Dolan JG, Bardley DR, Mushlin AI: An evaluation of clinicians' subjective prior probability estimates. *Med Decis Making* 1986;6:216-223.