

Evaluation of Standards of Practice for Primary Care Physicians Using 12 Hypothetical Cases

Frank H. Lawler, MD, and Michael C. Hosokawa, EdD
Greenville, North Carolina, and Columbia, Missouri

Concerns about the quality of medical care are increasing. As quality of medical care is difficult to measure, physicians have relied on standards based on research, consensus, or personal experience. The authors surveyed primary care physicians in Missouri to determine the presence and extent of standards of care for 12 hypothetical cases. The results demonstrate that within and between diagnoses there is both consensus and disagreement in case management. A standard of practice can be inferred for those management options achieving consensus. Further research is indicated for those options generating considerable disagreement to resolve the discrepancies in standards of care.

Practice standards for medical care differ in their origins and durability. These standards originate in three ways. The most precise standards are based on published research—childhood immunizations are examples of such practice standards. The second standard is that established by group norms or consensus. An example is the use of antibiotics for bronchitis. This standard is less rigorous and harder to measure. Finally, there are personal standards based on an individual physician's personal background and experience. Such standards are subject to extreme bias and are highly variable among physicians.

Practice standards are in a state of flux, changing with the discovery of new knowledge. The irradiation of acne, once a consensus standard, was replaced by other therapy when research demonstrated unacceptable risks. Some research-based standards, such as immunotherapy for renal transplantation, are replaced by improved procedures or therapies as new knowledge is disseminated.

Recent research has focused on assessing quality of care using practice standards as indicators of quality.¹⁻⁶ In the absence of rigorous measurements of quality, increased knowledge of practice standards is of obvious importance.^{7,8} In primary care research the practice standards of most interest are perhaps the personal standards, because of their origins, durability, and fallibility.

This study was designed to survey primary care physicians in Missouri to determine their responses to various management options presented for 12 hypothetical cases and to ascertain whether there were consensus standards for these cases. Consensus and personal standards, not research standards, were studied.

METHODS

In October 1985 a questionnaire was mailed to 1,300 Missouri primary care physicians whose names were derived from a master list of physicians maintained by the American Medical Association (AMA). The random sample included 650 family and general physicians, 325 pediatricians, and 325 internists.

Each physician received a questionnaire that presented six hypothetical cases and associated management options. Pediatricians received six cases involving children, and internists received six cases dealing with adults. Family physicians received three cases of each type.

Twelve cases were used in the questionnaire (Table 1). These cases, which represented problems seen in primary care, were chosen arbitrarily, but corresponded somewhat to diagnoses used by Greenwald et al.⁹ In addition, demographic information was elicited. A single reminder letter was sent to those who failed to respond within two weeks of the initial mailing.

In Table 2 a sample case is presented to illustrate the case history and management options. Each case used a vignette consisting of a short case history and eight to 12

Submitted, revised, December 8, 1986.

From the Robert Wood Johnson Fellowship Program, Department of Family and Community Medicine, University of Missouri-Columbia, Columbia, Missouri. Requests for reprints should be addressed to Dr. Frank H. Lawler, Department of Family Medicine, P.O. Box 1846, Greenville, NC 27835-1846.

TABLE 1. CASES USED IN THE MANAGEMENT SURVEY

Case No.	Pediatric Cases	Case No.	Adult Cases
1	Asthma	1	Angina pectoris
2	Pharyngitis	2	Hypertension
3	Bicycle accident	3	Asthma
4	Recurrent otitis media	4	Gastrointestinal hemorrhage
5	Febrile seizures	5	Thrombophlebitis
6	Dog bite	6	Systemic lupus erythematosus

management choices in three categories—referral, diagnosis, and treatment. Each management choice was listed with a five-point Likert scale. Means and standard deviations were calculated for each option, with a mean score approaching 5 signifying high utilization and a score toward 1 implying low utilization.

Management options were classified into four types of standards based on the mean and standard deviation for that option. The mean was used as a measure of the relative popularity (prevalence) of each item. The standard deviation was used as a measure of the amount of agreement between respondents for a given option; thus the higher the value of the standard deviation, the more variation between responses. Consensus standards were those options with high prevalence and low variance. Personal standards were those management options with moderate popularity and wide variance (high standard deviation). Options were defined as consensus standards if the mean was greater than 4.0, regardless of the standard deviation. An option was also considered to be a consensus standard if the mean was greater than 3.5 with a standard deviation less than 1.0. An option was classified as a personal standard if the mean was between 3.5 and 4.0, and the standard deviation was equal to or greater than 1.0. Coun-

terstandards were those options with means less than 2.5. Equivocal standards included all other values of means.

RESULTS

The overall response rate for the survey was 34 percent. Information on physician age, sex, and board-certification status was provided by the AMA. Comparisons show that respondents were not significantly different from nonrespondents in age or sex (47.5 vs 48.8 years, and 13.1 percent male vs 13.2 percent female, respectively). Respondents, however, were more likely to be board certified (64.5 percent vs 50.1 percent not board certified, $P < .05$). The response rate by specialty was not significantly different ($\chi^2 = 4.88$), although there was a trend toward a lower response rate by internal medicine specialists.

In Tables 3 and 4 the case histories are delineated with summaries of the standards of practice.

Cases Involving Children

For case 1, primary care physicians preferred to manage the teenager with asthma rather than refer the patient for consultation or management. Subcutaneous epinephrine, outpatient theophylline, and prompt follow-up were the consensus treatment options. The use of outpatient inhalation bronchodilator treatment appears to be a personal standard for some physicians.

In case 2, responding physicians would manage the girl with pharyngitis without referral. The three diagnostic options most frequently used were a throat culture for streptococcus, a 10-minute screening test for streptococcus, and the Monospot test. Because of the high variance, none of these procedures would be considered as a consensus practice standard. The variance, however, was highest for the 10-minute screening test, which is a relatively new diagnostic capability. The high variance most likely results from the adoption rate of a new procedure.

TABLE 2. SAMPLE OF CASE HISTORY AND MANAGEMENT OPTIONS

A 25-year-old white man comes in with severe abdominal pain, initially periumbilical but now localized to the right lower quadrant. His appetite is poor. Please specify the frequency with which you would perform the following:

	Always	Usually Would	Sometimes	Usually Not	Never
Complete blood count	5	4	3	2	1
Barium enema	5	4	3	2	1
Observation	5	4	3	2	1
Referral to psychiatry	5	4	3	2	1
Referral to surgery	5	4	3	2	1
Upper gastrointestinal series	5	4	3	2	1
Paracentesis	5	4	3	2	1

TABLE 3. RESULTS OF QUESTIONNAIRE FOR PEDIATRIC CASES

Case	Consensus Standards	Counterstandards	Personal Standards	Equivocal Options
1. A 17-year-old girl with known chronic asthma comes in following three days of wheezing, which is getting worse. She is short of breath and somewhat anxious, using accessory muscles of respiration. She is not on any medications. How would you manage this case?	Subcutaneous epinephrine Follow-up with office visit or telephone call Outpatient theophylline	Refer for consultation Refer for management	Outpatient treatment with bronchodilator Chest roentgenogram	Arterial blood gas Admit to hospital Spirometry Sputum culture
2. A 13-year-old girl comes in with seven-day history of "sore throat." She says her little brother had strep throat nine days ago but did not see a physician. She has no known allergies. Examination shows pharynx to be red with whitish exudate on tonsils, and presence of tender cervical nodes. Right upper quadrant is mildly tender. What is your management?	Antibiotic if culture is positive Antipyretics and analgesics Rest at home	Refer for consultation Lymph node aspiration or biopsy Refer for management	Streptococcus culture Monospot test 10-min streptococcus screening test Prescribe antibiotic now	Complete blood count Return visit
3. An 8-year-old boy is in the emergency department following a bicycle accident in which he hit a tree. He is alert but has left chest pain and left upper quadrant pain. Blood pressure is 90/60 mmHg, pulse 100 beats/min, respirations 24/min. Ribs are tender in mid-axillary line. Abdomen shows some guarding in the left upper quadrant but no rigidity. What would your approach be?	Chest x-ray examination Serial hematocrits Abdominal x-ray series Admit for observation Refer for surgical consultation	Peritoneoscopy Peritoneal lavage	Liver/spleen scan	Refer for surgical management
4. A 3-year-old boy comes in with his third episode in five months of ear pain, irritability, and low-grade fever. Examination shows right tympanic membrane to be red and bulging with an effusion present. What is your approach?	Course of amoxicillin, trimethoprim/sulfamethoxazole (eg, Septra, Bactrim or Cefaclor)	Complete blood count Diagnostic myringotomy	Return visit Decongestants	Refer for polyethylene tube placement Prophylactic antibiotic therapy after acute episode
5. A 2-year-old girl with no previous history of seizures is brought in because of a temperature of 105°F and a 5-min episode of generalized convulsions. She is now postictal, with a temperature of 102.5°F. What would your management be?	Complete blood count	Refer to a neurologist for consultation Refer to a neurologist for management Given intravenous diazepam Computed tomography of head	Measure electrolytes and blood glucose immediately Give intravenous or intramuscular phenobarbital	Sponge bath Electroencephalogram After acute care give maintenance antiepileptics and return every 3 mo for 2 yr Lumbar puncture Blood cultures
6. A 5-year-old boy comes to your office, having been bitten by a neighbor's dog four hours previously. He has a 3-cm laceration on his arm which is 3- to 4-mm deep. How would you manage this case?	Confine and observe dog Return office visit or telephone call Check rabies status of dog	Culture wound Initiate rabies vaccine	Prescribe antibiotic	Close wound primarily Refer for wound management

TABLE 4. RESULTS OF QUESTIONNAIRE FOR ADULT CASES

Case	Consensus Standards	Counterstandards	Personal Standards	Equivocal Options
1. A 60-year-old man with known angina comes to your office. Episodes have been once a week and fairly well controlled with nitroglycerin and a beta-blocker, but have recently occurred three times per week, and require 2 nitroglycerin tablets to alleviate the pain. What would be your approach to this problem?	Return visit Electrocardiogram	Alter medication regimen		Refer for consultation Chest roentgenogram Coronary angiography Multipanel blood chemistry (SMAC, etc) Refer for management Exercise stress test Admit to hospital
2. A 75-year-old woman with benign hypertension comes in for a routine visit. Her blood pressure has been stable on a thiazide diuretic. Blood pressure today is 160/100 mmHg. She is asymptomatic. What is your management?	Recheck blood pressure in one week Return visit	Refer for consultation Chest roentgenogram Hypertension intravenous pyelogram 24-hour urine for creatinine, vanillylmandelic acid, metanephrins, catecholamines Refer for management Urine electrolytes	Modify diet	Serum electrolytes Alter medication regimen Electrocardiogram
3. A 27-year-old woman with known chronic asthma comes in following three days of wheezing, which is getting worse. She is short of breath and somewhat anxious, using accessory muscles of respiration. She is not on any medications. How would you manage this case?	Chest roentgenogram Telephone call or office visit Outpatient oral theophylline Outpatient inhalation treatment with bronchodilator	Refer for consultation Refer for management	Subcutaneous epinephrine	Arterial blood gas Spirometry Sputum culture Admit to hospital

Rest at home, antipyretics and analgesics, and antibiotics if cultures were positive were the most common treatment options. These three options had high prevalence and low variance, and thus are considered to be consensus standards.

Unlike the management of the other cases, most respondents would obtain a consultation for the boy in the bicycle accident (case 3). The consensus diagnostic choices were chest roentgenogram, abdominal films, and serial hematocrits. Counterstandards appear to be peritoneal lavage and peritoneoscopy. Admission for observation was the consensus treatment standard.

For the child with recurrent otitis media (case 4), there was a consensus that diagnostic myringotomy was not indicated (low prevalence and low variance). Antibiotics were the treatment of choice, or consensus standard, for treatment. Use of decongestants was a personal treatment standard rather than a consensus standard. This therapy had relatively high prevalence and variance, indicating divergent opinions about the value of these agents. The high variance probably reflects the tension between an

established traditional practice standard and research that fails to support their use.

Case 5 was a child with a febrile seizure. Primary care physicians would manage this case. The diagnostic procedure considered to be a consensus practice standard was a complete blood count. Procedures such as a computerized axial tomographic scan and the use of diazepam have low use and low variance (counterstandards). There was little agreement regarding electrolytes and blood glucose determination, blood cultures, lumbar puncture, or sponge bath for treatment.

For the five-year-old boy with the dog bite (case 6), the consensus standards were related to the dog's rabies status. Primary closure of the wound generated the most variance, and culturing the wound or initiating rabies vaccine were considered not to be indicated for this case.

Cases Involving Adults

In the first adult case, most respondents would manage the patient with increasing angina themselves, but there

TABLE 4, continued

Case	Consensus Standards	Counterstandards	Personal Standards	Equivocal Options
4. A 65-year-old retired man with a two-week history of epigastric pain, belching, and heartburn comes into your office. He has no nausea or vomiting, but has noted darker stools for two to three days. Physical examination is normal except for slight epigastric tenderness. Rectal examination is negative except for positive stool guaiac. Blood pressure and pulse are normal, with minimal postural changes. What is your management?	Upper gastrointestinal series Complete blood count	CT scan of abdomen Refer for consultation	Multipanel blood chemistry (SMAC, etc)	Electrocardiogram Chest roentgenogram Barium enema Admit to hospital Serum amylase Abdominal x-ray series Gastroscopy
5. A 33-year-old woman, nonsmoker on birth control pills comes in with pain in her right calf. Homans' sign is equivocal. Examination also shows superficial varicosities, and palpation over mid-calf region shows moderate tenderness. She has no breathing problems and her lungs are clear. What would your management be?		Refer for consultation Ventilation/perfusion scan Refer for management		Chest roentgenogram Doppler study of legs Admit for heparin, heat, and elevation Outpatient treatment with anti-inflammatory medication, heat, and elevation After acute episode refer for tubal ligation, or switch to barrier form of birth control
6. A 38-year-old woman comes into your office with a three-month history of fever, malaise, and pain in the hand and wrist joints. She had a facial rash two weeks ago. She is under considerable stress at home and at work. What would your management be?	Complete blood count Multipanel blood chemistry (SMAC) Sedimentation rate Rheumatoid factor Return visit Antinuclear antibody titer	Refer to psychiatrist	Lupus erythematosus preparation	Refer to rheumatologist 24-hr urine for protein and creatinine Serologic test for syphilis Admit to hospital

was not a clear indication whether the primary care physicians would seek a consultation. Among the diagnostic choices, only the electrocardiogram was a consensus practice standard. There was apparent consensus not to alter the therapeutic regimen, suggesting that respondents agreed with the medications being used.

In case 2, the elderly woman with benign hypertension would have been managed by the primary care physicians without a consultation. Most of the diagnostic options had low means and standard deviations, indicating that the physicians would not perform tests. The consensus treatment standard was to check the patient's blood pressure in one week.

In case 3 (the 27-year-old woman with known chronic asthma) the respondents elected to manage this patient

without consultation. A chest roentgenogram was the consensus diagnostic standard. Outpatient oral theophylline, outpatient inhalation treatment with a bronchodilator, and follow-up were the consensus treatment standards. Unlike the previous case in which the respondents would do little diagnosis or treatment, there were diverse approaches to this case, which were reflected by equivocal standards. The use of arterial blood gases and subcutaneous epinephrine were controversial management options.

In case 4, the elderly man with guaiac-positive stools would have been managed by the primary care physicians. A complete blood count and upper gastrointestinal series were the consensus diagnostic standards. Other management options, such as a multipanel blood chemistry,

serum amylase, chest films, and electrocardiogram, were commonly selected, but the variance was relatively high, indicating lack of consensus.

The woman with possible thrombophlebitis (case 5) would have been managed by the responding physicians. There were no consensus standards for diagnosis or treatment. Venogram, admission for heparin, heat and elevation, chest films, and outpatient treatment with anti-inflammatory medication were common management options, but relatively high standard deviations indicated a lack of consensus.

The woman in case 6 with probable systemic lupus erythematosus would probably be managed by most primary care physicians; however, there were some respondents who would have referred her to a rheumatologist. Complete blood count, blood chemistry, rheumatoid factor, sedimentation rate, and antinuclear antibody titer were consensus management standards. The lupus erythematosus preparation test was the most controversial, probably because it has recently been deemphasized as a diagnostic test, but some physicians continue to order it.

DISCUSSION

The management of the cases themselves may not be so important as the patterns of options chosen. Consensus was determined by the prevalence of a choice and the variance. Management choices considered as consensus standards of practice were those with high prevalence and low variance. Management choices with low prevalence and low variance, procedures the respondents would not utilize, also indicated standards of care, ie, those actions that were not considered appropriate.

Personal practice standards were those choices with relatively high prevalence and high variance, indicating a large number of physicians selected this option, but there were varied opinions on the appropriate action. Practice counterstandards were those choices with low prevalence, indicating options that are not considered appropriate to the care required for that case. Lack of consensus may reflect a transition period in which a standard encompassing a new technology is replacing an established standard. Lack of adherence to the established standard may derive from a lack of knowledge or availability of the new technology or skepticism about its accuracy or value.

These findings have implications for health policy in regard to the issues of quality and cost of care and litigation. Federal and state governments are taking steps to control the cost of medical care, assure quality, and resolve the problems of malpractice litigation. The role of legislation in the context of these issues remains unclear.

Can practice standards be used as measures of quality

of care? These findings show that the consensus practice standards (high prevalence, low variance) and the counterstandards (low prevalence) can be most clearly identified. The personal practice standards are more elusive; however, even crude indicators (moderate prevalence and high variance) identified a lack of consensus, possibly indicating controversy, changing knowledge, new technology, or access to resources. Equivocal management choices need rigorous examination to demonstrate utility.

If the respondent's choices were valid measures of intent, studies such as this one would be of local or regional value in the context of available resources. Finding that most physicians agree on the management of a patient with suspected lupus erythematosus identifies the standards for care among physicians in an area. On the other hand, the moderate prevalence and high variance found in the use of decongestants for the treatment of otitis media reflect a preference for a traditional approach, although published research does not support the efficacy of this therapy.

Cost containment is a major topic for discussion today. Personal standards may not be cost effective or cost beneficial, such as a complete blood count for a child with recurrent otitis media. Some physicians routinely order this test, although the benefits are questionable. Surveys such as this one may be useful in identifying personal practice standards based on personal bias or preferences that cannot be justified. Identifying the practice standards of physicians has significant implications for continuing medical education.

Finally, practice may be carried beyond medical quality to excessive thoroughness as a defense against litigation. For a patient in a bicycle accident with a possible ruptured spleen, chest and abdominal roentgenograms and serial hematocrits were consensus practice standards, but peritoneal lavage or peritoneoscopy were personal standards for some respondents. Perhaps such management practices are adopted to decrease the risk of litigation. How much the potential for litigation influences personal practice standards, if any, is unknown.

One might wonder whether interspecialty differences are related to standards. While some small differences were found, the specialties are much more similar than dissimilar. For example, family physicians were more likely than pediatricians to suture a laceration. Also, family physicians were more likely than internists to refer adult patients. Regardless of these minor differences, specialty was not a major factor in determining the standards of care for a given case or option.

The findings of this study can be questioned on reliability and validity. Reliability coefficients (Cronbach's alpha) were greater than 0.9 for all cases. Internal and external validity are important facets of a survey. External validity deals with the generalizability of the findings, that

is, whether these results can be extrapolated to other populations. The modest response rate and the difference in board-certification status in this study will weaken the ability to apply the standards to the whole population surveyed and other populations of physicians. The generalizability is strengthened by the finding of nonrespondents being similar to respondents in both sex and age distributions.

Internal validity deals with the question, Do the physicians' selections really reflect the way they practice medicine? The validity of using hypothetical cases as a method of measuring physician actions has been challenged in the literature.¹⁰ Evidence favoring the validity of this method is that physicians who rated themselves as being more aggressive actually responded more aggressively on the hypothetical cases, giving a measure of construct validity. Also, physician preferences for management options were consistent with published standards. These preferences give a suggestion of criterion validity, but it is possible that the respondents were repeating their ideas of the standards rather than providing their own management approach.

While this study provides only incomplete data on practice standards, a series of coordinated surveys would provide data on the trends and changes of personal and

consensus standards. Such a series would also help to validate the use of this method.

References

1. Sanazaro PJ, Worth RM: Measuring clinical performance of individual internists in office and hospital practice. *Med Care* 1985; 23:1097-1114
2. Freiman MP: The rate of adoption of new procedures among physicians. *Med Care* 1985; 23:939-945
3. Palmer RH, Louis TA, Hsu LN, et al: A randomized controlled trial of quality assurance in sixteen ambulatory care practices. *Med Care* 1985; 23:751-770
4. Estes EH, Sullivan RJ: Standards for practice: Effectiveness and acceptance. *Ann Intern Med* 1978; 89(part 2):826-828
5. Baker R: Comparison of standards in training and nontraining practices. *J R Coll Gen Pract* 1985; 35:330-332
6. Campion PD: Setting standards in general practice. *Br Med J* 1985; 291:499
7. Sheps S, Robertson A: Evaluation of primary care in a community clinic by means of explicit process criteria. *Can Med Assoc J* 1984; 131:881-886
8. Payne BC, Lyons TF, Neuhaus E, et al: Method of evaluating and improving ambulatory medical care. *Health Serv Res* 1984; 19: 219-245
9. Greenwald HP, Peterson ML, Garrison LP, Hart LG, et al: Inter-specialty variation in office-based care. *Med Care* 1984; 22:14-29
10. Hartley RM, Charlton JR, Jarman B, Harris CM: Case history questionnaires in the study of doctor's use of resources: Are they measuring what we want? *Med Care* 1985; 23:1163-1170