Clinicians Are From Mars and Pathologists Are From Venus

Clinician Interpretation of Pathology Reports

Seth M. Powsner, MD; José Costa, MD; Robert J. Homer, MD, PhD

• ***Context.*—Text reports convey critical medical information from pathologists, radiologists, and subspecialty consultants. These reports must be clear and comprehensible to avoid medical errors. Pathologists have paid much attention to report completeness but have ignored the corresponding issue of report comprehension. This situation presents an increasingly serious potential problem. As laboratories are consolidated and as reports are disseminated in new ways (eg, via the World Wide Web), the target audience becomes more diverse and less likely to have any contact with pathologists beyond the written reports themselves.**

***Objective.*—To compare clinician comprehension with pathologist intent in written pathology reports.**

***Methods.*—Typical surgical pathology reports relevant to surgeons and covering a range of specimen complexity were taken from our hospital files. Questionnaires based on these cases were administered open-book-examination style to surgical attending physicians and trainees during surgical conferences at an academic medical center.**

***Main Outcome Measures.*—Scores from questionnaires.**

***Results.*—Surgeons misunderstood pathologists' reports 30% of the time. Surgical experience reduced but did not eliminate the problem. Streamlined report formatting exacerbated the problem.**

***Conclusions.*—A communication gap exists between pathologists and surgeons. Familiarity with report format and clinical experience help reduce this gap. Paradoxically, stylistic improvements to report formatting can interfere with comprehension and increase the number of misunderstandings. Further investigation is required to reduce the number of misunderstandings and, thus, medical errors.**

(*Arch Pathol Lab Med.* 2000;124:1040–1046)

Just as medical language may be hard for lay people to understand, medical specialists may use language that is obscure to practitioners outside their specialty. Among specialists, the language of diagnostic anatomic pathologists is arguably furthest from daily medical discourse. There has been an implicit recognition of this fact in efforts to standardize pathology reporting. These efforts have lead to the adoption of the so-called Bethesda system for Papanicolaou smear reporting[1] and similar attempts to standardize reports in other areas, including tumor resections, transplant biopsies, and liver biopsies.[2–4] However, little explicit attention has been focused on the ability of clinicians to understand pathology reports, despite the role of reports in guiding treatment.

For editorial comment, see p 943.

Advances in anatomic pathology information systems offer an opportunity to change pathology reports, for example, by reporting via the World Wide Web and digital imaging. We hoped to improve report comprehension by using new forms of presentation,[5–9] beginning with a streamlined text report format, and we hoped to show empirically that our new formats increased comprehension or reduced errors. Contrary to our original hypothesis, our modified anatomic pathology report formats reduced comprehension when compared to our hospital system's standard report. There were a surprising number of misunderstandings of what pathologists intended to convey, even when using our system's original report format. More problems are likely, given current trends to present reports solely on computer terminals as part of a computer-based patient record. New report formats may decrease comprehension, presumably due to reduced familiarity. The problem of report comprehension warrants exploration.

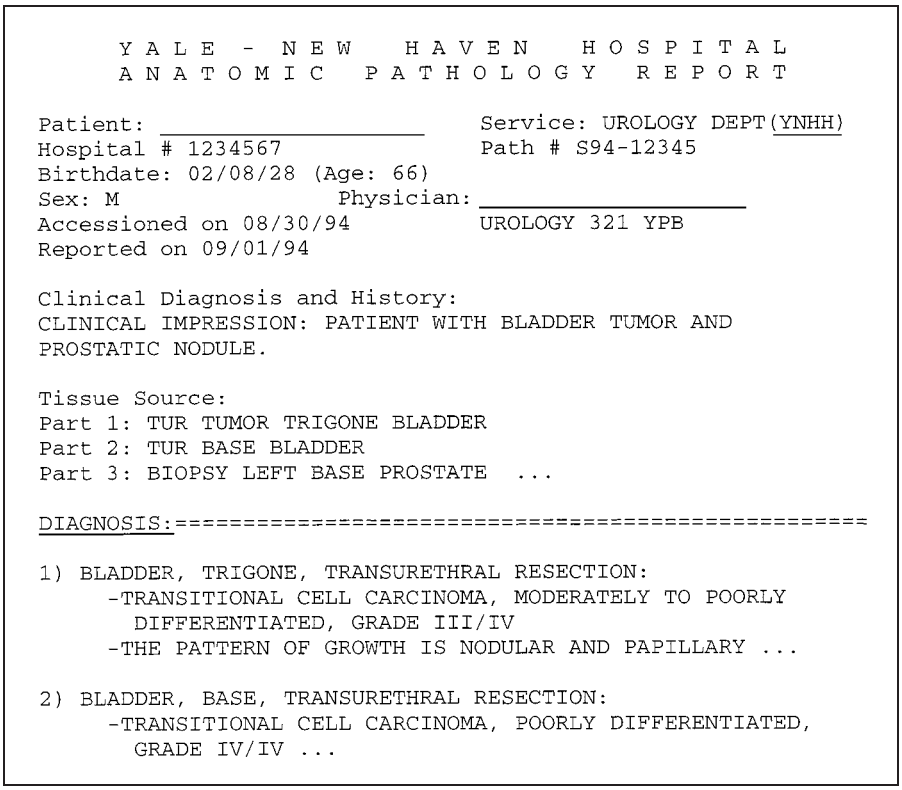
METHODS

We adopted an experimental design similar to Fletcher's[10] approach for testing problem-oriented medical record comprehension. Six anatomic pathology reports were selected to cover a range of specimens and complexity. They included 1 report each of a transurethral resection of a bladder tumor, a medical liver biopsy, a transplant kidney biopsy (using the Banff[11] reporting scheme), a parathyroid resection, a transurethral resection of prostate, and a gastric biopsy. Report text came from our hospital's online anatomical pathology system (CoPath, Dynamic

Accepted for publication December 12, 1999.

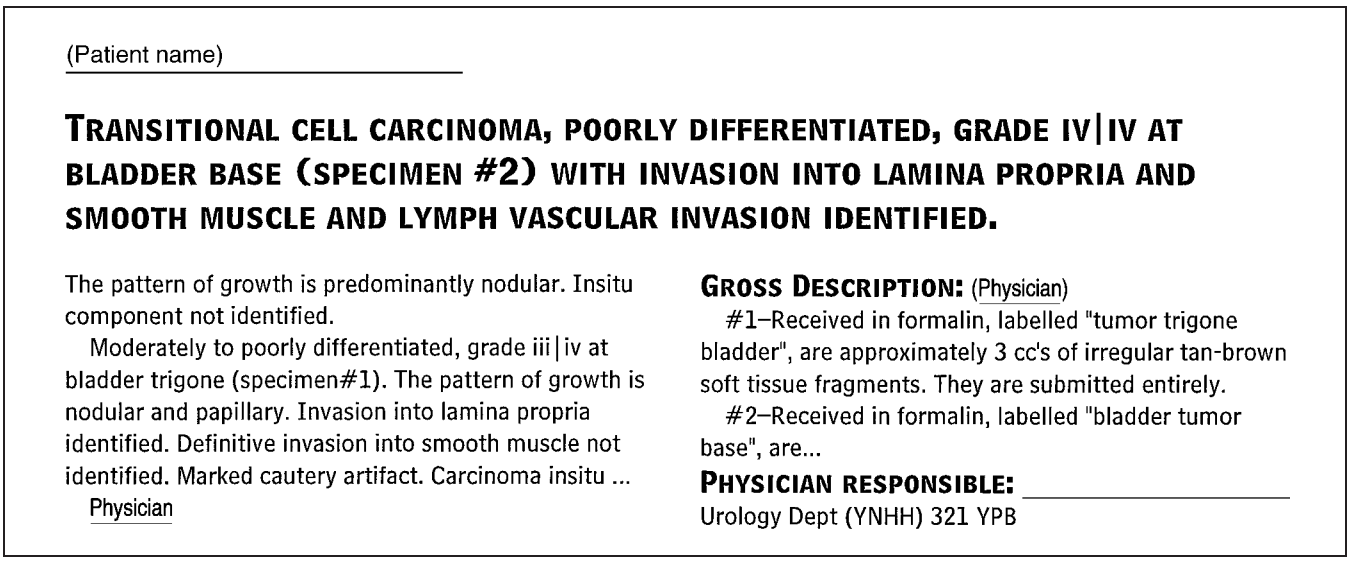
From the Department of Psychiatry and Center for Medical Informatics (Dr Powsner) and the Department of Pathology (Drs Costa and Homer), Yale University School of Medicine, New Haven, Conn; and Pathology and Laboratory Medicine Service, VA Connecticut Healthcare System, West Haven, Conn (Dr Homer).

Reprints: Seth M. Powsner, MD, Yale Psychiatry 2039cb, 20 York St, New Haven, CT 06504.



```
      Y A L E  -  N E W    H A V E N    H O S P I T A L
      A N A T O M I C    P A T H O L O G Y    R E P O R T

Patient: _____________________     Service: UROLOGY DEPT(YNHH)
Hospital # 1234567                 Path # S94-12345
Birthdate: 02/08/28 (Age: 66)
Sex: M                 Physician: _____________________
Accessioned on 08/30/94            UROLOGY 321 YPB
Reported on 09/01/94

Clinical Diagnosis and History:
CLINICAL IMPRESSION: PATIENT WITH BLADDER TUMOR AND
PROSTATIC NODULE.

Tissue Source:
Part 1: TUR TUMOR TRIGONE BLADDER
Part 2: TUR BASE BLADDER
Part 3: BIOPSY LEFT BASE PROSTATE  ...

DIAGNOSIS:=============================================

1) BLADDER, TRIGONE, TRANSURETHRAL RESECTION:
     -TRANSITIONAL CELL CARCINOMA, MODERATELY TO POORLY
       DIFFERENTIATED, GRADE III/IV
     -THE PATTERN OF GROWTH IS NODULAR AND PAPILLARY ...

2) BLADDER, BASE, TRANSURETHRAL RESECTION:
     -TRANSITIONAL CELL CARCINOMA, POORLY DIFFERENTIATED,
       GRADE IV/IV ...
```

Figure 1. *Original report format.*



(Patient name) _____

TRANSITIONAL CELL CARCINOMA, POORLY DIFFERENTIATED, GRADE IV|IV AT BLADDER BASE (SPECIMEN #2) WITH INVASION INTO LAMINA PROPRIA AND SMOOTH MUSCLE AND LYMPH VASCULAR INVASION IDENTIFIED.

The pattern of growth is predominantly nodular. Insitu component not identified.

Moderately to poorly differentiated, grade iii|iv at bladder trigone (specimen#1). The pattern of growth is nodular and papillary. Invasion into lamina propria identified. Definitive invasion into smooth muscle not identified. Marked cautery artifact. Carcinoma insitu ...

Physician

GROSS DESCRIPTION: (Physician)

#1–Received in formalin, labelled "tumor trigone bladder", are approximately 3 cc's of irregular tan-brown soft tissue fragments. They are submitted entirely.

#2–Received in formalin, labelled "bladder tumor base", are...

PHYSICIAN RESPONSIBLE: _____

Urology Dept (YNHH) 321 YPB

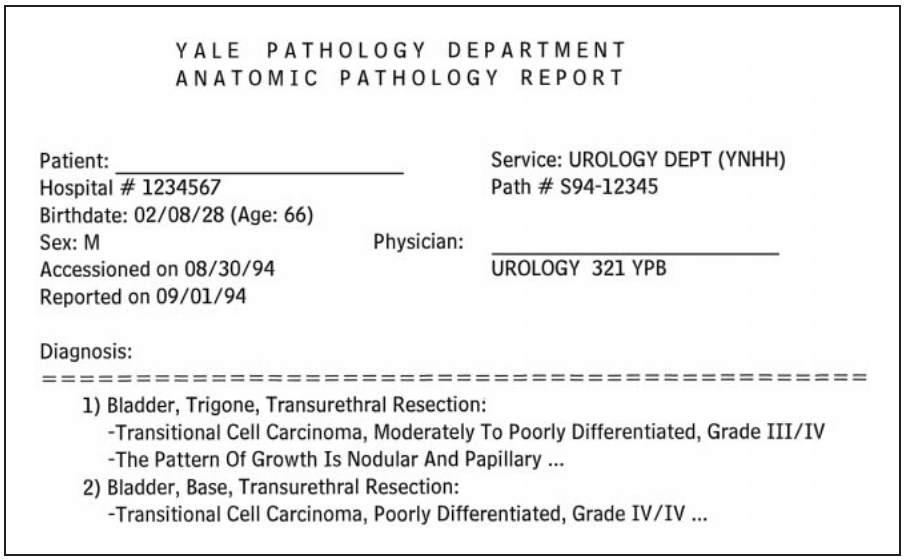
Figure 2. *New report format.*

Healthcare Systems, Waltham, Mass). We altered names and identifying information to protect confidentiality.

From each original report (Figure 1) we created a new streamlined version (Figure 2). Key pathologic findings appeared immediately after the patient's name. Nonclinical institutional information was placed at the bottom of the report. Compact Bell Centennial printer font[12,13] permitted us to present the name and essential findings in a larger type size, while keeping specimen preparation details legible. We changed the text of each report only as grammatically necessary for this new format. To control for changes in type style and page arrangement, we also created a partially modernized version of each original (Figure 3) with the original text set in Bell Centennial font and with final findings moved higher on the page.

We prepared a short clinical scenario for each case, then we prepared questions to be answered after reading a version of the report. Intended answers to these questions (yes, no, or not stated) were determined by 2 attending pathologists (J.C. and R.J.H.) after reviewing the reports. There was agreement between the pathologists in all cases. Figure 4 shows questions for the report appearing in Figures 1 through 3. Subjects were allowed 2 chances to answer these questions: first after reading an introduction to the case and the pathology report and setting them both aside, then immediately afterwards in an open-book-examination style

Figure 3. *Modernized report format.*


YALE PATHOLOGY DEPARTMENT
ANATOMIC PATHOLOGY REPORT

Patient: _____
Hospital # 1234567
Birthdate: 02/08/28 (Age: 66)
Sex: M
Accessioned on 08/30/94
Reported on 09/01/94

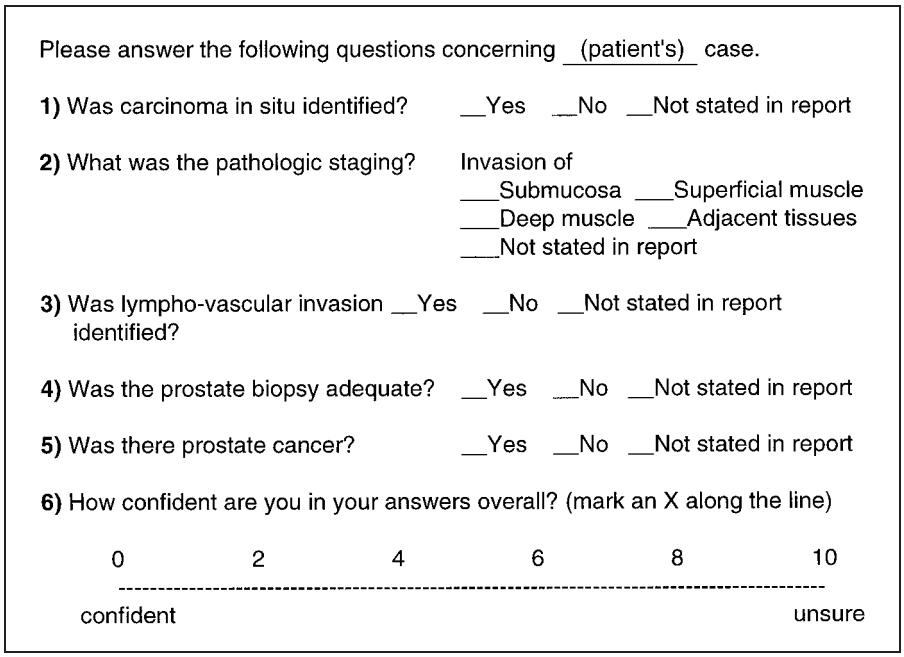
Service: UROLOGY DEPT (YNHH)
Path # S94-12345

Physician: _____
UROLOGY 321 YPB

Diagnosis:
==========

1) Bladder, Trigone, Transurethral Resection:
-Transitional Cell Carcinoma, Moderately To Poorly Differentiated, Grade III/IV
-The Pattern Of Growth Is Nodular And Papillary ...
2) Bladder, Base, Transurethral Resection:
-Transitional Cell Carcinoma, Poorly Differentiated, Grade IV/IV ...

Figure 4. *Sample report questions.*


Please answer the following questions concerning (patient's) case.

1) Was carcinoma in situ identified? __Yes __No __Not stated in report

2) What was the pathologic staging? Invasion of
___Submucosa ___Superficial muscle
___Deep muscle ___Adjacent tissues
___Not stated in report

3) Was lympho-vascular invasion identified? __Yes __No __Not stated in report

4) Was the prostate biopsy adequate? __Yes __No __Not stated in report

5) Was there prostate cancer? __Yes __No __Not stated in report

6) How confident are you in your answers overall? (mark an X along the line)

0 2 4 6 8 10
---
confident unsure

with report and questions side by side. All subjects reviewed all 6 cases, but subjects reviewed different combinations of formats (2 cases in each of the 3 formats). Only the more accurate, open-book responses were used in this analysis since charts can be kept open and reports reviewed any number of times during routine clinical practice. Initial review of closed-book answers did not reveal any significantly different trends, just more discrepancies.

Test subjects were 34 general surgical attending physicians and trainees. Testing was scheduled during time volunteered from regularly scheduled team conferences to allow group testing without extending anyone's workday. Average time for completion was 30 minutes. Question books were randomly passed around the room, unmarked and unsigned, except for a code to allow 1 trainee and 1 attending, chosen randomly from among the top 4 scores, to be awarded gift certificates for a local restaurant. Our Human Investigation Committee approved this protocol.

A seventh case was prepared and used as a sample for explaining the test procedure. It was not counted in any of the results.

RESULTS

Table 1 shows a disconcerting discordance between pathologists' intended meanings and interpretation by surgeons; the crude discrepancy rate was 30%. As expected, experience improved performance, although the discrepancy rate was still significant for both attending physicians and senior trainees reading reports in their original format (Table 1; Figure 5). All data tally responses with reports printed and open right next to the questionnaires. We did not conclude that our subjects were wrong, only that the pathologists' intended meanings were not effectively communicated. Discrepancy rates from Table 1 may unfairly magnify possibilities for clinical misadventure. In real hospital practice, surgeons can ask pathologists to clarify

Table 1. Discordance Rate (%) by Report Format and Clinical Experience					
	Format				
Experience*	Original	Modernized	New	All	n
Attending, pgy 6+	18	24	34	25	15
Housestaff, pgy 2–5	26	32	34	31	11
Student, pgy 1	33	28	48	37	8
All	24	28	37	30	34

* pgy indicates postgraduate year.

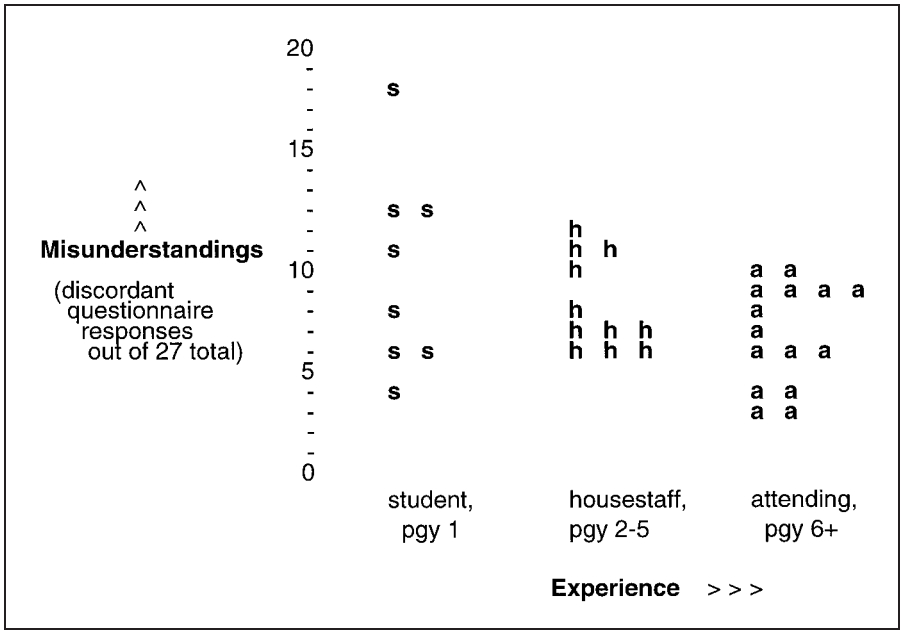


Figure 5. *Individual scores versus experience.*

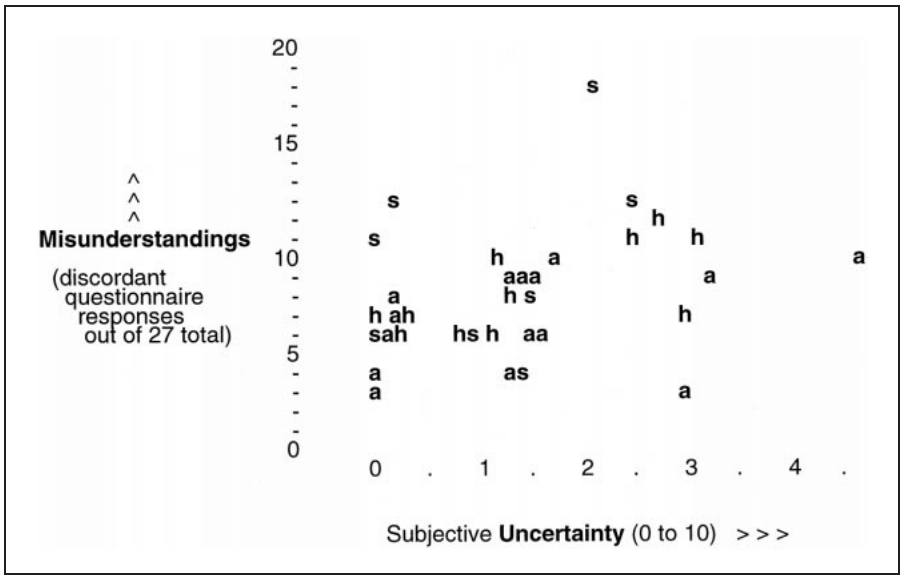


Figure 6. *Individual scores versus uncertainty. s indicates student or postgraduate year (pgy) 1; h, housestaff (ie, pgy 2–5); and a, attending physician or pgy 6+.*

ambiguities or resolve unanswered questions if the surgeons recognize their clinical uncertainty. Unfortunately, we found no reliable correlation between discordance and uncertainty. Figure 6 shows discordant responses plotted against uncertainty ratings. (These are overall case uncertainty ratings. Subjects were asked only to indicate their uncertainty in answering all questions about a case, not in answering individual questions.) A scattering of uncertainty and discordance measurements is seen, and uncertainty is not related in any simple way to inexperience.

As for the effects of changing report format, the larger the change, the larger the discordance rates (Table 1). Our modernized format (Figure 3) yielded higher discordance rates than the original format (Figure 1) overall. Our rad-

Table 2. Question Summary by Discordance Rate

Rate, %	Difficulty	Case No.•Question No./Question Summary	Clinical Procedure†
0	1	1•2/How much prostate resected?	TURP for BPH
0	1	3•2/Presence of cholestasis?	Liver biopsy
0	1	5•5/Prostate cancer?	TURB and prostate biopsy
3	1	1•1/Malignancy?	···
6	1	3•1/Cirrhosis?	···
6	1	4•2/*Helicobacter pylori*?	Gastric biopsy
6	2	4•1/Malignancy?	···
12	1	5•3/Lymphovascular invasion?	···
12	2	6•4/Any thymic tissue present?	Parathyroidectomy
15	1	5•1/Carcinoma in situ?	···
15	2	6•3/Any nerve tissue present?	···
21	2	4•4/Ulcer present?	···
21	2	6•2/Any thyroid tissue present?	···
26	1	5•2/Pathologic staging?	···
26	2	3•3/Single etiology for liver disease?	···
26	2	6•1/Frozen sections confirmed by final diagnosis	···
38	1	1•4/Was all tissue reviewed?	···
38	3	2•2/Acute rejection?	Transplant renal biopsy
41	3	4•5/Barrett's epithelim?	···
44	1	6•5/How many lymph nodes removed?	···
56	2	2•4/Diabetic glomerulopathy?	···
56	2	2•5/Infection?	···
59	2	2•1/Adequate speciman?	···
59	2	5•4/Prostate biopsy adequate?	···
65	1	1•3/Procedural complications?	···
74	3	2•3/Chronic interstitial rejection?	···
79	3	4•3/Reflux?	···

* Difficulty scale: 1, can be read directly from text of report or counted; 2, requires some interpretation of report; 3, requires higher level of understanding or terminology.

† TURP indicates transurethral resection of prostate; BPH, benign prostatic hypertrophy; and TURB, transurethral resection of bladder.

ically new format (Figure 2) yielded even higher discordance rates. Despite our intended improvements, Table 1 shows that test subjects reading either of our new formats were overall less likely to understand our pathologists' intentions than were those reading reports in the original format.

Since all observed discordance rates seemed unacceptable, we abandoned further analysis related to specific format. We sought, instead, to identify the characteristics of questions—or of reports in general—that lead to discrepancies. We divided questions into the following 3 categories: those which can be answered (1) directly from the text of a report, (2) after some interpretation of a report, or (3) only with a higher level of (clinical) understanding or terminology, including the operation of a pathology laboratory. Table 2 lists abbreviated questions in order of increasing discordance. In general, the more interpretation required, the more discrepancies occurred.

Questions that were answered by an explicit statement in the diagnosis section of the report were usually answered correctly (eg, question 4•2, presence or absence of *Helicobacter pylori* in a gastric biopsy). Surprisingly, question 1•2, which required examination of the gross description section of the report, was always answered correctly. (Perhaps ''40 grams'' stood out as this report's only numeric finding.) Questions concerning the presence or absence of carcinoma were virtually always answered correctly. Errors in this category were limited to 1 surgical intern, who may have just finished night call,[14–16] and 2 attending physicians. The 2 attending physicians actually answered ''not stated,'' and under normal circumstances could have asked for clarification.

Discrepancies increased in cases with multipart specimens, even if our questions were answered explicitly in a report's diagnosis section (questions 5•1 and 6•5). Pathologists generally note only the presence of certain tissues in the specimen; presumably, readers understand that tissues not mentioned were not found. Our clinician readers appeared not to share this convention (questions 6•4 and 6•2). There were similar misunderstandings about specimen adequacy. In 2 such questions (questions 2•1 and 5•4), 60% of clinicians did not recognize a specimen to be adequate even when a specific diagnosis was made by the pathologist. The Banff system for kidney biopsy reports explicitly states that diagnosis should not be made in the absence of adequate tissue.

Two often misunderstood questions required a knowledge of how pathology laboratories function (questions 1•3 and 1•4). In the first case, full-thickness bowel was found on the slide but was stated to be a contaminant. Many clinicians interpreted this to mean that there was a procedural complication. In the second case, the report stated that only representative prostate chips were examined histologically, but clinicians failed to register this. Both cases have medicolegal implications.

Other questions with medicolegal implications included mistaken impression of a frozen section–final diagnosis difference and mistaken impression of having removed thymus tissue (questions 6•1 and 6•4). Our hospital convention is to label each part with the specimen label given by the surgeon. The surgeon labeled the specimen ''thymus''; this convention may have left the impression that thymus had in fact been removed.

Another convention at our institution is to not include a histologic description except in unusual or difficult cases. This practice might lead to higher rates of misunderstanding; however, among the 6 cases studied, more severe problems of comprehension arose when clinicians

had to interpret histologic descriptions or understand histologic terminology (questions 2•2, 2•3, 2•4, and 4•5). Histologic descriptions may be of little use to surgeons. For example, question 2•4 assumed a basic knowledge of histologic and ultrastructural characteristics of diabetic glomerulopathy (diabetic patient with a transplanted kidney). Interestingly, this particular case, which was reported in a formal terminology (Banff system), was the least understood overall. A shorthand system such as the Banff system is completely opaque to nonspecialists.

Subjects were solicited for comments after each session. No subject suggested that the questions or test format was clinically unfair. A number of our surgical oncology team members did ask that a structured report format be added to considerations for new computer reports. A number of surgeons indicated a preference for ''the darker printing'' of the modernized reports.

COMMENT

An overall discordance rate of 30% for surgeons and surgical trainees answering open-book questions about anatomic pathology reports was surprising. It is reassuring that more experienced attending physicians performed better than less experienced trainees, and that presence or absence of invasive cancer was virtually always recognized. However, the discrepancies that remained were not just about minor findings. Reading a kidney transplant biopsy report, 13 (38%) of 34 clinicians had the mistaken impression of no acute rejection. With respect to a bladder biopsy report, there was widespread failure to recognize deeply invasive bladder cancer and carcinoma in situ, both of which are key features in determining therapy.

Our study took place in an optimal setting with printed paper reports and a block of time set aside from hospital ward demands. In clinical practice, reports may be seen only on a computer screen in a busy, noisy ward. Our study shows that on average trainees understand reports less well than attending surgeons. Unfortunately, in academic settings, reports may be retrieved by residents, to be reported back to attending physicians. Nurse-practitioners or physician assistants may screen pathology reports prior to physician review in private practice settings (Priscilla Jencks, MPH, MSN, personal communication, June 3, 1998). While we have not studied these midlevel practitioners, similar difficulties in comprehension are likely, as their formal training in pathology is minimal or nonexistent.

Most of our subjects had been reading reports in our original format for years. Changing the report format may have caused a negative short-term effect on these readers (ie, report disorientation). A long-term study might show improvement above the starting baseline. Students and subjects in their first postgraduate year, who were less familiar with any report format, demonstrated less degradation with new formats (Table 1). We suspect it takes a year to establish short-term report preference based on format. Based on these data, the new formats may ultimately be shown to be superior. Our improvements in report format comprised changes to type font, wording, and page layout (position). Perhaps these factors must be varied and tested separately.

Other factors confound our findings. It is possible our test setting eliminated critical clinical motivation for report interpretation. Moreover, we have no way of knowing which discrepancies in interpretation would lead to real medical errors. A prospective study of report interpretation in clinical settings can be imagined, although it would be difficult to execute. Such an investigation would require separation of clinical error from poor outcome, and its results would likely be confounded by the observers themselves (Hawthorne effect).

Essentially no empirical literature exists on pathology report comprehension to provide a guide for improvement. Professional societies and authors have recommended various styles for clinical reports.[2–4] It is unknown if any of the proposed standardized reports actually improve understanding by clinicians. However, standardized reports encourage explicit reporting, and our findings suggest that the less interpretation by readers the better. Adopting standardized systems comes at a price: new nomenclatures can lead to balkanization of knowledge. Standardized systems are not a panacea. They codify presentation of common findings in routine clinical situations, but cannot reflect the wide range of findings commonly encountered in clinical practice. Our problematic cases could not have been reported in any standardized way.

Results obtained here may or may not apply to other consultation reports (eg, radiology or psychiatry). Do primary care physicians understand reports from subspecialists to whom they refer patients for diagnosis and treatment recommendations? Based on their understanding or misunderstanding, primary clinicians often implement and follow up treatment on their own, because of managed care constraints. Leape[17] did not focus on professional communication as a source of medical errors in his review in 1994. More recently, Bhasale and colleagues[18] described poor communication between health professionals as a contributing factor in 19 of 100 incidents registered by Australian general practitioners. Their findings reinforce the idea that communication failure is an important source of medical error.[19]

Technical writers have a similar challenge when conveying commercial or industrial information.[20,21] Unfortunately, it is not clear how guidelines for instructing a homeowner to assemble a lawnmower, for example, should be applied to diagnostic reports for trained professionals. New information systems allow inclusion of line drawings, gross photographs, and photomicrographs in reports. Other purely stylistic improvements also are possible, for example, placing important statements up front or using a more legible type font. However, our study demonstrates that stylistic improvements alone may reduce rather than increase comprehension. Clinicians and pathologists need to improve medical communication for patient care, while avoiding the disruption of established familiar report formats currently protecting clinical content.

We thank the members of the Department of Surgery, Yale University School of Medicine, New Haven, Conn, for their time and cooperation in this study. We also appreciate an initial review and help coordinating test sessions provided by Dana Andersen, MD. Elizabeth C. Yen, JD, was of great assistance assembling test books. We also thank the late Robert Byck, MD, for his thoughtful critique of our experimental design and early drafts of our manuscript.

References

1. Broder S. Rapid communication: the Bethesda System for reporting cervical/vaginal cytologic diagnoses: report of the 1991 Bethesda workshop. *JAMA*. 1992; 267:1892.

2. Markel SF, Hirsch SD. Synoptic surgical pathology reporting. *Hum Pathol*. 1991;22:807–810.

3. Leslie KO, Rosai J. Standardization of the surgical pathology report: formats, templates, and synoptic reports. *Semin Diagn Pathol.* 1994;11:253–257.
4. College of American Pathologists. Reporting on cancer specimens: protocols and case summaries. Available at: http://www.cap.org/html/publications/cancerfactsheet.html. Accessed October 28, 1999.
5. Wyatt JC, Wright P. Design should help use of patients' data. *Lancet.* 1998; 352:1375–1378.
6. Nygren E, Wyatt JC, Wright P. Helping clinicians to find data and avoid delays. *Lancet.* 1998;352:1462–1466.
7. Wright P, Jansen C, Wyatt JC. How to limit clinical errors in interpretation of data. *Lancet.* 1998;352:1539–1543.
8. Powsner SM, Wright P, Wyatt JC. Medical records: opportunities for and challenges of computerisation. Lancet. 1998;352:1617–1622.
9. Patel VL, Kushniruk AW. Understanding, navigating and communicating knowledge: issues and challenges. *Methods Inf Med.* 1998;37:460–470.
10. Fletcher RH. Auditing problem-oriented records and traditional records: a controlled comparison of speed, accuracy and identification of errors in medical care. *N Engl J Med.* 1974;290:829–833.
11. Gaber LW, Moore LW, Gaber AO, et al. Utility of standardized histological classification in the management of acute rejection: 1995 Efficacy Endpoints Conference. *Transplantation.* 1998;65:376–380.
12. Carter M. Bell Centennial. In: *Type & Technology Monograph I.* New York, NY: Cooper Union Center for Design & Typography; 1982.
13. Tufte ER. The visual display of quantitative information. Cheshire, Conn: Graphics Press; 1983.
14. Deary IJ, Tait R. Effects of sleep disruption on cognitive performance and mood in medical house officers. *BMJ.* 1987;295:1513–1516.
15. Baldwin RJ, Dodd M, Wrate RM. Junior doctors making mistakes. *Lancet.* 1998;351:804.
16. Taffinder NJ, McManus IC, Gul Y, Russell RCG, Darzi A. Effect of sleep deprivation on surgeon's dexterity on laparoscopy simulator. *Lancet.* 1998;352: 1191.
17. Leap LL. Error in medicine. *JAMA.* 1994;272:1851–1857.
18. Bhasale AL, Miller GC, Reid SE, Britt HC. Analysing potential harm in Australian general practice: an incident-monitoring study. *Med J Aust.* 1998;169: 73–76.
19. Arnstein F. Catalogue of human error. *Br J Anaesth.* 1997;79:645–656.
20. Wright P. Writing to be understood: why use sentences? *Appl Ergonomics.* 1971;207–209.
21. Wright P. "The instructions clearly state. . . ." Can't people read? *Consumer Ergonomics.* 1981;12(3):131–141.