Frequency of Failure to Inform Patients of Clinically Significant Outpatient Test Results

Lawrence P. Casalino, MD, PhD; Daniel Dunham, MD, MPH; Marshall H. Chin, MD, MPH; Rebecca Bienlang, MD; Emily O. Kistner, PhD; Theodore G. Karrison, PhD; Michael K. Ong, MD, PhD; Urmimala Sarkar, MD, MPH; Margaret A. McLaughlin, MD; David O. Meltzer, MD, PhD

**Background:** Failing to inform a patient of an abnormal outpatient test result can be a serious error, but little is known about the frequency of such errors or the processes for managing results that may reduce errors.

**Methods:** We conducted a retrospective medical record review of 5434 randomly selected patients aged 50 to 69 years in 19 community-based and 4 academic medical center primary care practices. Primary care practice physicians were surveyed about their processes for managing test results, and individual physicians were notified of apparent failures to inform and asked whether they had informed the patient. Blinded reviewers calculated a “process score” ranging from 0 to 5 for each practice using survey responses.

**Results:** The rate of apparent failures to inform or to document informing the patient was 7.1% (135 failures divided by 1889 abnormal results), with a range of 0% to 26.2%. The mean process score was 3.8 (range, 0.9-5.0). In mixed-effects logistic regression, higher process scores were associated with lower failure rates (odds ratio, 0.68; P < .001). Use of a “partial electronic medical record” (paper-based progress notes and electronic test results or vice versa) was associated with higher failure rates compared with not having an electronic medical record (odds ratio, 1.92; P = .03) or with having an electronic medical record that included both progress notes and test results (odds ratio, 2.37; P = .007).

**Conclusions:** Failures to inform patients or to document informing patients of abnormal outpatient test results are common; use of simple processes for managing results is associated with lower failure rates.

Arch Intern Med. 2009;169(12):1123-1129
chance that a test result will eventually show up on a physician's desk, but in a poorly organized EMR-based practice, the physician may never realize that the result has been received.

**METHODS**

**DEVELOPMENT OF THE MEDICAL RECORD REVIEW PROTOCOL AND STUDY INSTRUMENTS**

We selected 11 blood tests and 3 screening tests (mammography, Papanicolaou smear, and fecal occult blood) commonly performed in the outpatient setting (eTable 1; http://www.archinternmed.com). After discussion within the team and consultation with physicians in appropriate specialties, we defined a range of “clinically significantly abnormal” values for each test (eTable 1). These values were mostly well out of the reference range for each test; our intent was to define results as clinically significantly abnormal only when our team and consultants believed that nearly any physician would agree that the patient should be informed of the result, either because it indicated an immediate danger or because it had potential implications for the patient’s health over time (eg, a high total cholesterol level).

For each abnormal result, we searched the patient’s medical record for 13 types of events (eBox 1) suggesting that the patient had been informed and scored the physician as having been informed if any one of the events had occurred within a predefined time interval. For example, we considered the patient informed if there was a note stating that the patient had been informed, if the abnormal test was repeated, or if a relevant consultation or procedure (eg, referral to a urologist and/or results of a prostate biopsy) was performed. In most cases, we defined 90 days as the interval within which the patient should be informed; in a few cases (eg, exceptionally high or low serum sodium or potassium level) we defined the interval as 21 days. We used these relatively long time intervals to ensure that there was time, for example, for a procedure to have been scheduled and performed and the results to have appeared in the record.

To our knowledge, no guidelines exist to delineate the processes that practices should use to manage test results. Based on our review of the literature23-25 and a pilot study that we conducted, we defined “good processes” for managing test results as the following: (1) all results are routed to the responsible physician; (2) the physician signs off on all results; (3) the practice informs patients of all results, normal and abnormal, at least in general terms; (4) the practice documents that the patient has been informed; and (5) patients are told to call after a certain time interval if they have not been notified of their results.

We developed a written 6-question survey for distribution to physicians at each study site (available at http://www.archinternmed.com). The first 5 questions asked about the processes used by the physician to manage test results. The sixth question asked how satisfied the physician was with these methods. We also developed a short semistructured protocol for interviewing with a physician leader of the practice (eBox). The protocol included questions similar to the survey questions, as well as questions about the extent, if any, to which the practice used an EMR.

We developed a “physician notification form” (PNF) to mail when we found an apparent failure to inform. The PNF included the patient’s name, date of birth, medical record number, test result, and date of the test so that the physician could take corrective action if he or she believed it appropriate to do so. We asked the physician to return the form to us after checking a response indicating whether he or she believed that (1) the patient had been informed but that this had not been documented, (2) the patient had not been informed and that they planned to notify the patient, (3) the patient had not been informed because the physician did not consider the result clinically significant, or (4) the record contained information showing that the patient had been informed.

The medical record review protocol, survey instrument, PNF, and interview protocol were tested during a review of 1066 patient records in a general internal medicine practice at an academic medical center; revisions were made based on this pilot testing. This study was approved by the institutional review board at each participating academic medical center and by Chesapeake Research Review, which served as the institutional review board for private practice sites.

**STUDY PARTICIPANTS**

No census of physician practices exists in the United States. We selected community-based practices randomly using preferred provider organization directories of physicians available online from large health insurance plans in the Midwest and on the West Coast. Only primary care practices or multispecialty groups with at least 20% primary care physicians were eligible. The principal investigator called practices to invite participation; of the 98 practices that were approached, 19 agreed. General internal medicine clinics at 4 academic medical centers (2 in the Midwest and 2 on the West Coast) were selected on a convenience basis and agreed to participate.

At each practice, research assistants randomly selected patients aged between 50 and 69 years, who had been seen by a primary care physician during the 90 to 360 days prior to the date of the review. Patients of residents and fellows were excluded. We selected this age range to include patients likely to have had more tests and more abnormal test results than younger patients. To exclude patients likely to have a short life expectancy (ie, patients for whom it could be argued that a high total cholesterol or hemoglobin A1c, for example, might not be a significant finding), we excluded patients older than 70 years and patients with diseases likely to be fatal in the short term (eg, metastatic cancer). We also excluded patients with medical conditions (eg, chronic renal failure) likely to make it difficult to define clinically significant results. We excluded individual tests in patients with a condition likely to complicate interpretation of that test (eg, we excluded prostate specific antigen tests in patients with benign prostatic hypertrophy). Reviews were performed between June 2005 and February 2006.

A total of 176 surveys were distributed to all primary care physicians in smaller practices and to as many as 15 randomly selected primary care physicians in larger practices.

**MEASURES**

We defined results as abnormal if they fell outside our predefined “normal” range. We defined “apparent failures to inform” as abnormal results for which the reviewer could not find evidence within the medical record that the patient had been informed within the defined time interval. We defined “failures to document” as apparent failures to inform for which the physician stated in the PNF that the patient had been informed but that this had not been documented. We defined “failures to inform” as apparent failures to inform for which the physician did not return a PNF or stated on the form that the patient had not been informed.

Two physician members of the team who were blinded to the identities of the practices and failure rates used survey re-
responses and notes from research assistants’ interviews with practice leaders to independently give each practice a score of zero to 1 for each of the 5 processes. Zero indicated that the practice did not use the process; 1, that it appeared to use the process routinely; and intermediate scores indicated that the practice appeared to use the process to some extent but not routinely. The reviewers’ scores were generally very close; when they differed, they discussed their reasoning and then rescored (weighted $\kappa = 0.56-0.72$ for the final total scores and individual items). In the few cases when the reviewers’ scores differed after rescoring, the practice was assigned the average of the 2 scores. The “process score” for each practice was then calculated as the sum of the scores for the 5 processes.

We categorized practices as having a “full EMR” if both test results and progress notes were available to physicians in electronic form; as having a “partial EMR” if results or notes, but not both, were available electronically; and as not having an EMR if neither was available electronically.

**MEDICAL RECORD REVIEW**

A review of 5434 medical records was conducted using laptop computers with the protocol automatically displayed, by 3 fourth-year and 4 second-year medical students. Each reviewer was trained by the lead author. At 18 of the 23 sites, reviews were conducted by pairs of reviewers, with each reviewing approximately half the records. Randomly selected records were reviewed by both reviewers; for 97% of these records the reviewers’ findings were in agreement.

The lead author (L.P.C.) reviewed all 182 apparent failures to inform using diagnoses and notes recorded by the reviewers and information from the physician notification forms. Reviewers were judged to have erred in 14 of the 182 cases (7.7%). In 8 of these cases, the patient should have been excluded from the study; in 4, the physician’s response to the PNF stated that the medical record indicated the way in which the patient had been informed (eg, by a telephone call); in 2 cases, the test result was not clinically significantly abnormal.

**STATISTICAL ANALYSES**

We conducted 4 linear regression analyses using (1) the failure to inform or document rate of each practice as the outcome variable and the practice’s process score as the predictor variable, with each practice weighted by its number of abnormal results; (2) weighted linear regression with practices’ failure rate as the outcome variable and average physician satisfaction with its processes for managing test results as the predictor variable; (3) each practice’s average physician satisfaction as the outcome variable and the practice’s process score as the predictor; and (4) multivariate mixed-effects logistic regression analysis with failure to inform or document as the dichotomous outcome and process score and EMR type as the predictor variables. Statistical analyses were performed with SAS version 9.1 software (SAS Institute Inc, Cary, North Carolina).

**RESULTS**

Reviewers recorded 1889 abnormal results and 182 apparent failures to inform (Figure). A review of the apparent failures by the lead author indicated that 10 should be excluded as ineligible and that an additional 10 were ambiguous, given the protocol; ambiguous cases were counted as notified. Physician notification forms were sent to 105 physicians for the remaining 162 apparent failures to inform. Fifty-one physicians (49%) returned 74 forms (45%). For 17 apparent failures, the physician stated that the patient had been informed; in 18 cases the physician stated that the patient had been informed but that this was not documented; in 18 cases the physician stated that the patient was not responsible for the test; and in 6, the physician stated that the patient had not been informed and that the physician planned to do so. We counted the cases in which the physician stated that he or she was not re-
Table 1. Results by Practice Site

<table>
<thead>
<tr>
<th>Primary Care Practice No.</th>
<th>Type/Size of Practice</th>
<th>No. of Primary Care MDs</th>
<th>Medical Records Reviewed/Excluded</th>
<th>Abnormal Results</th>
<th>Failures to Inform</th>
<th>Failures to Document</th>
<th>Failure Rate, %</th>
<th>Process Score (Range, 0-5)</th>
<th>MD Satisfaction Score (Range, 1-4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Small site of large practice</td>
<td>6</td>
<td>196/6</td>
<td>59</td>
<td>0</td>
<td>0</td>
<td>0.4</td>
<td>4.0</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>Small</td>
<td>2</td>
<td>151/3</td>
<td>38</td>
<td>0</td>
<td>0</td>
<td>0.4</td>
<td>3.5</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>Small</td>
<td>8</td>
<td>205/4</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>0.4</td>
<td>4.1</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>Small</td>
<td>5</td>
<td>150/3</td>
<td>73</td>
<td>1</td>
<td>0</td>
<td>1.4</td>
<td>3.9</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>Small</td>
<td>6</td>
<td>186/7</td>
<td>56</td>
<td>1</td>
<td>0</td>
<td>1.8</td>
<td>3.0</td>
<td>No</td>
</tr>
<tr>
<td>6</td>
<td>Large</td>
<td>26</td>
<td>169/8</td>
<td>247</td>
<td>7</td>
<td>0</td>
<td>2.8</td>
<td>3.8</td>
<td>Yes</td>
</tr>
<tr>
<td>7</td>
<td>Small</td>
<td>2</td>
<td>149/2</td>
<td>61</td>
<td>0</td>
<td>2</td>
<td>3.3</td>
<td>4.2</td>
<td>No</td>
</tr>
<tr>
<td>8</td>
<td>Small</td>
<td>3</td>
<td>151/0</td>
<td>29</td>
<td>0</td>
<td>1</td>
<td>3.4</td>
<td>5.0</td>
<td>No</td>
</tr>
<tr>
<td>9</td>
<td>Small</td>
<td>1</td>
<td>131/3</td>
<td>42</td>
<td>2</td>
<td>0</td>
<td>4.8</td>
<td>3.6</td>
<td>No</td>
</tr>
<tr>
<td>10</td>
<td>AMC</td>
<td>47</td>
<td>402/19</td>
<td>98</td>
<td>5</td>
<td>0</td>
<td>5.1</td>
<td>4.4</td>
<td>Yes</td>
</tr>
<tr>
<td>11</td>
<td>Small</td>
<td>5</td>
<td>149/3</td>
<td>57</td>
<td>3</td>
<td>0</td>
<td>5.3</td>
<td>4.1</td>
<td>No</td>
</tr>
<tr>
<td>12</td>
<td>Large</td>
<td>44</td>
<td>346/1</td>
<td>74</td>
<td>4</td>
<td>0</td>
<td>5.4</td>
<td>4.6</td>
<td>Partial</td>
</tr>
<tr>
<td>13</td>
<td>Small</td>
<td>7</td>
<td>198/3</td>
<td>54</td>
<td>3</td>
<td>0</td>
<td>5.6</td>
<td>3.3</td>
<td>No</td>
</tr>
<tr>
<td>14</td>
<td>Large</td>
<td>17</td>
<td>489/18</td>
<td>234</td>
<td>11</td>
<td>2</td>
<td>5.6</td>
<td>3.5</td>
<td>Yes</td>
</tr>
<tr>
<td>15</td>
<td>Small site of large practice</td>
<td>13</td>
<td>320/5</td>
<td>126</td>
<td>6</td>
<td>2</td>
<td>6.3</td>
<td>4.5</td>
<td>Yes</td>
</tr>
<tr>
<td>16</td>
<td>Small</td>
<td>4</td>
<td>150/6</td>
<td>35</td>
<td>3</td>
<td>0</td>
<td>8.6</td>
<td>4.4</td>
<td>No</td>
</tr>
<tr>
<td>17</td>
<td>Large</td>
<td>13</td>
<td>346/2</td>
<td>103</td>
<td>8</td>
<td>1</td>
<td>8.7</td>
<td>4.5</td>
<td>Partial</td>
</tr>
<tr>
<td>18</td>
<td>Small</td>
<td>3</td>
<td>151/0</td>
<td>34</td>
<td>3</td>
<td>0</td>
<td>8.8</td>
<td>3.9</td>
<td>No</td>
</tr>
<tr>
<td>19</td>
<td>AMC</td>
<td>23</td>
<td>369/9</td>
<td>158</td>
<td>13</td>
<td>2</td>
<td>9.5</td>
<td>3.1</td>
<td>No</td>
</tr>
<tr>
<td>20</td>
<td>Small</td>
<td>6</td>
<td>150/1</td>
<td>64</td>
<td>5</td>
<td>2</td>
<td>10.9</td>
<td>4.3</td>
<td>No</td>
</tr>
<tr>
<td>21</td>
<td>Small</td>
<td>4</td>
<td>196/1</td>
<td>39</td>
<td>5</td>
<td>1</td>
<td>15.4</td>
<td>3.1</td>
<td>No</td>
</tr>
<tr>
<td>22</td>
<td>AMC</td>
<td>41</td>
<td>349/14</td>
<td>121</td>
<td>24</td>
<td>2</td>
<td>21.5</td>
<td>0.9</td>
<td>Partial</td>
</tr>
<tr>
<td>23</td>
<td>AMC</td>
<td>33</td>
<td>329/11</td>
<td>61</td>
<td>13</td>
<td>3</td>
<td>26.2</td>
<td>2.0</td>
<td>Partial</td>
</tr>
</tbody>
</table>

Total or mean: 5434/129 / 1889 / 117 / 18
7.1% (135 failures divided by 1889 abnormal results).

Failure rates ranged from 0% in 3 practices to 26.2% (Table 1). Patients were not informed of results of a total cholesterol level as high as 318 mg/dL (to convert to millimoles per liter, multiply by 0.0259) or a hemoglobin A<sub>1c</sub> level as high as 18.9% (to convert to pro-portion, multiply by 0.01) or a potassium level as low as 2.6 mEq/L (to convert to millimoles per liter, multiply by 1) or a hematocrit level as low as 28.6% (Table 2). The mean process score was 3.8 on a 0 to 5 scale, with 5 indicating that the practice routinely used all 5 processes that we hypothesized would be associated with a lower failure rate. Process scores ranged from 0.9 to 5.0.

We performed linear regression analysis with practices’ failure to inform or document rates as the outcome variable and their process scores as the predictor variable, with each of the 23 practices weighted by its number of abnormal results. Process scores were significantly associated with failure to inform rates (β [SE], −0.05 [0.01]; P < .001). Table 2 displays descriptive data on this association, showing how failure rates varied by low, medium, and high process scores.

Of 176 physician surveys, 99 were returned (56.2% response rate). On average, physicians were moderately satisfied with their system for managing test results: on a 4-point scale, ranging from “very satisfied” to “moderately satisfied” to “very dissatisfied,” the mean response was 3.2 (“very satisfied” was scored as a 4). Practices that had higher process scores had higher physician satisfaction (β [SE], 0.45 [0.10]; P < .001) and practices that had higher satisfaction had lower failure rates (β [SE], −0.06 [0.02]; P = .009).
had the highest failure rates. We did not find a significant difference between practices that had a “complete” EMR and those that used paper records; this may be because there is no difference or because the number of practices included was not large enough to detect a difference.

Most practices did not use all 5 of the relatively simple processes suggested in the literature as basic to managing test results. Most did not have explicit rules for notifying patients of results, and many used the dangerous practice of telling patients that “no news is good news”—an assumption that the Agency for Healthcare Research and Quality counsels patients not to make.26

To our knowledge, this is the first study to estimate the failure to inform rate across a variety of tests and types of medical practice. Other studies have suggested that failures to inform are common. In a national survey, 11% of patients stated that they had experienced delays during the previous year in receiving abnormal test results.27 In a study of 126 patients with abnormal mammograms in 10 academically affiliated practices in 1996-1997, the physician documented having discussed the result with the patient in 71% of cases.28 Of 48 patients at a single academic center with an abnormal dual-energy x-ray absorptiometry scan, it appeared that 33% had not been informed.15 Single-site studies suggest that physicians frequently fail to follow up on abnormal thyrotropin,29,30 potassium,31 and blood glucose levels,32 though these studies were not designed specifically to estimate failure to inform rates.

Our study has several limitations. First, it is possible that in some of the cases counted as failures to inform, the patient actually had been informed. It is likely that such cases were few, if they existed at all, since we counted the patient as informed if any one of 13 types of evidence that the patient had been informed appeared in the

In this study, failures to inform patients of clinically significant abnormal test results or to document that they have been informed appear to be relatively common, occurring in 1 of every 14 tests. Failure rates varied widely among practices, from 0% to 26%; practices that used better processes to manage results had lower failure rates and had physicians who were more satisfied with the processes used. Practices that used a combination of paper and electronic records, that is, that had a “partial EMR,”
medical record or if the physician responded to the PNF by stating that the patient had been informed. Second, because practices were included only if they agreed to participate, the failure rates we found may differ from what would be found in a random sample of practices; unfortunately, such random sampling is likely to be impossible. Third, medical chart reviews were performed by medical students, which may have led to some errors; use of a detailed computer protocol, review by the lead author, and use of the PNFs were intended to minimize this possibility. Fourth, we studied only primary care physicians, and in a limited number of practices (n=23) on the West Coast and in the Midwest.

Limited research suggests that most patients and physicians believe that patients should be informed of both abnormal and normal test results.7,21,33,34 Failures to inform patients of abnormal test results or to document that they have been informed can harm patients and expose physicians to indefensible malpractice liability.

One approach to reducing failure rates would be to rely on the efforts of individual physicians and to exhort them to try harder to notify patients. Alternatively, failures to inform could be approached as a systems problem—a problem of organization and incentives—rather than as a failing of individual physicians. We observed practices that use EMRs in which the only way to see test results is by searching the record of each patient for whom a physician has ordered a test; in these practices we found individual physicians devising their own methods, such as Excel spreadsheets, to help them remember to check for results. At the opposite extreme, some practices used EMRs in which all results are routed to the electronic mailbox of the responsible physician. Abnormal results are highlighted and the system records the fact that the physician has clicked on the results.

Some elements of medical care (eg, diagnosis) are an art as well as a science, depend heavily on the cognitive skills and effort of individual physicians, involve much uncertainty, and will probably always have relatively high error rates. However, notifying patients of test results does not appear to be such a process; with appropriate within-practice systems, low rates of failure to inform should be possible.33 For practices that want to improve, suggestions are available.16,36-39 Individual practices40 and at least 1 regional collaborative41 are experimenting with ways to improve the management of test results.

Accepted for Publication: March 6, 2009.

Correspondence: Lawrence P. Casalino, MD, PhD, Department of Public Health, Weill Cornell Medical College, 402 E 67th St, New York, NY 10065-6304 (lac2021@med.cornell.edu).

Author Contributions: Dr Casalino had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Casalino, Dunham, Chin, Bielang, Karrison, and Meltzer. Acquisition of data: Casalino, Bielang, Ong, Sarkar, and McLaughlin. Analysis and interpretation of data: Casalino, Dunham, Chin, Kistner, Karrison, Ong, and Meltzer. Drafting of the manuscript: Casalino, Kistner, and Karrison. Critical revision of the manuscript for important intellectual content: Casalino, Dunham, Chin, Bielang, Kistner, Karrison, Ong, Sarkar, McLaughlin, and Meltzer.

REFERENCES


Correction
Errors in Contributions. In the article titled “Frequency of Failure to Inform Patients of Clinically Significant Outpatient Test Results” by Casalino et al, published in the June 22 issue of the Archives (2009;169[12]:1123-1129), David O. Meltzer, MD, PhD, should have been included in the “Author Contributions” section on page 1128 for study concept and design; analysis and interpretation of data; and critical revision of the manuscript for important intellectual content. Also, Kari Fitzgerald Jerge, MD, should have been included in “Additional Contributions” for conducting medical record reviews and interviews.