

Testing process errors and their harms and consequences reported from family medicine practices: a study of the American Academy of Family Physicians National Research Network

J Hickner,¹ D G Graham,² N C Elder,³ E Brandt,² C B Emsermann,⁴ S Dovey,⁵ R Phillips⁶

¹ Department of Family Medicine, The University of Chicago Pritzker School of Medicine, Chicago, Illinois, USA; ² American Academy of Family Physicians National Research Network, Leawood, Kansas, USA; ³ University of Cincinnati Department of Family Medicine, Cincinnati, Ohio, USA; ⁴ Department of Family Medicine, The University of Colorado Health Sciences Center, Aurora, Colorado, USA; ⁵ Royal New Zealand College of General Practitioners Research Unit, Dunedin, New Zealand; ⁶ The Robert Graham Center: Policy Studies in Family Medicine and Primary Care, Washington, DC, USA

Correspondence to: J Hickner, Department of Family Medicine, The University of Chicago Pritzker School of Medicine, 5841 S. Maryland Ave, MC 7110, Suite M-156, Chicago, IL 60637-1470, USA; jhickner@uchicago.edu

Accepted 13 July 2007

ABSTRACT

Context: Little is known about the types and outcomes of testing process errors that occur in primary care.

Objective: To describe types, predictors and outcomes of testing errors reported by family physicians and office staff.

Design: Events were reported anonymously. Each office completed a survey describing their testing processes prior to event reporting.

Setting and participants: 243 clinicians and office staff of eight family medicine offices.

Main outcome measures: Distribution of error types, associations with potential predictors; predictors of harm and consequences of the errors.

Results: Participants submitted 590 event reports with 966 testing process errors. Errors occurred in ordering tests (12.9%), implementing tests (17.9%), reporting results to clinicians (24.6%), clinicians responding to results (6.6%), notifying patient of results (6.8%), general administration (17.6%), communication (5.7%) and other categories (7.8%). Charting or filing errors accounted for 14.5% of errors. Significant associations ($p < 0.05$) existed between error types and type of reporter (clinician or staff), number of labs used by the practice, absence of a results follow-up system and patients' race/ethnicity. Adverse consequences included time lost and financial consequences (22%), delays in care (24%), pain/suffering (11%) and adverse clinical consequence (2%). Patients were unharmed in 54% of events; 18% resulted in some harm, and harm status was unknown for 28%. Using multilevel logistic regression analyses, adverse consequences or harm were more common in events that were clinician-reported, involved patients aged 45–64 years and involved test implementation errors. Minority patients were more likely than white, non-Hispanic patients to suffer adverse consequences or harm.

Conclusions: Errors occur throughout the testing process, most commonly involving test implementation and reporting results to clinicians. While significant physical harm was rare, adverse consequences for patients were common. The higher prevalence of harm and adverse consequences for minority patients is a troubling disparity needing further investigation.

Testing errors are common in primary care practice and may lead to patient harm and malpractice claims.^{1–5} The testing process is traditionally divided into preanalytic, analytic and postanalytic phases. In each phase, clinicians, patients, office staff and lab staff perform a series of tasks. These tasks can be grouped into the following expanded

categories: ordering the test, implementing the test, performing the test, reporting results to the clinician, clinician responding to the results, notifying the patient of the results and following up to ensure the patient took appropriate action based on the test results (fig 1). The complexity of the testing process, frequent separation of lab testing from clinic location and general lack of quality control systems in primary care offices make testing in primary care error-prone.

Systematic efforts to describe, understand and eliminate problems that occur in the management of the testing process in office practice lag behind those of hospitals. The first study of testing process errors reported from US primary care practices found a rate of 1.1 lab-related problems per 1000 office visits and suggested that one in four of these errors affected patient care.¹ In a more recent office-based study, 15 family physicians self-reported errors immediately after patient encounters.² The investigators noted an error or preventable adverse event in 24% of patient encounters. Fourteen per cent of those errors/adverse events involved the testing process, a rate of one testing process error for every 30 office visits. In three other primary care patient safety studies, 15% to 54% of errors reported by physicians were related to the testing process.^{3–6,7}

This study is the largest study to date of testing process errors reported from family physician offices in the United States. We sought to determine the distribution of error types, to describe the outcomes of these errors and to perform an exploratory analysis of predictors of error types and adverse outcomes. The aims of this study were to gain a deeper understanding of testing process errors in primary care practice and to lay a foundation for developing interventions to improve the safety of the testing process.

METHODS

This study took place in eight family physician offices of the American Academy of Family Physicians National Research Network, including four private practices and four family medicine residency clinics. The offices were purposefully selected from a list of 58 volunteers to maximise practice diversity. The offices were a mix of small and large practices. One residency clinic was a federally qualified community health centre. The practices were in seven states throughout the US.

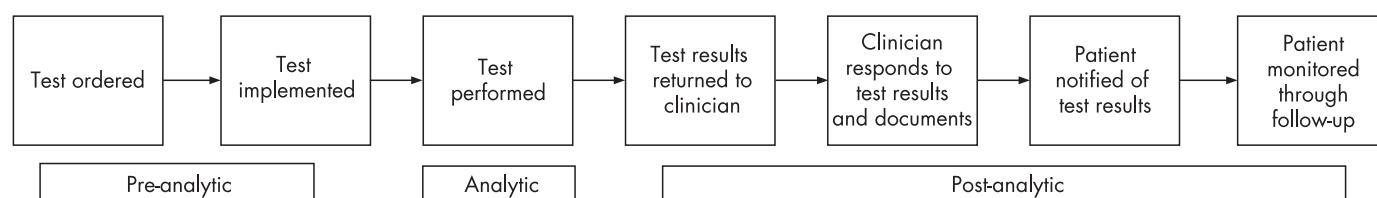


Figure 1 Conceptual framework of the testing process.

Four were rural, three were urban, and one was in a suburban location.

Physicians, residents, nurse practitioners, physician assistants and office staff submitted anonymous reports of errors they recognised or experienced during the course of their work day. We asked participants to report only errors related to the testing process, including lab tests, diagnostic imaging and other tests such as pulmonary function tests and electrocardiograms. We asked participants to report anything about the testing process that they observed or committed “that should not have happened and that you don’t want to happen again.”

Physicians and staff were given the option of reporting via the internet to a secure website or by mail. The reporting tool is consistent with the Institute of Medicine’s recommendations for error reporting, which are based on research on human performance measurement in industry.⁸ The fields of this error-reporting form are based on the conceptual framework of the Australian Incident Monitoring System and have been refined over the course of several primary care errors reporting studies (error-reporting form available on request).^{6,7,9} A physician and a designated study coordinator (usually a nurse or manager) from each of the participating sites attended a mandatory 1½-day group training session, and they trained other clinicians and staff and implemented the study at their respective offices.

Reports were submitted anonymously but included practice codes and reporter-type codes for four categories of reporters: physicians, nurse practitioners and physician assistants, office staff (including non-clinician nurses) and residents. Participants reported errors for 32 weeks in 2004. Each practice was assigned four intensive reporting weeks (1 week out of every 8 weeks) and 28 routine weeks. During the intensive reporting weeks, participants were asked to report every testing process error they identified, while during routine weeks, they were asked to report errors that they wanted us to know about, with particular attention to errors that had the potential to cause or actually did cause harm.

Before reporting began, a knowledgeable person in each office, generally the office manager, completed a survey describing the practice’s clinical testing processes.

ANALYSIS

Office lab survey

We summarised the responses to the office lab survey with frequency distributions. We dichotomised the offices into those with high-quality testing processes (offices reporting a specific system to ensure testing follow-up and who believe that more than 95% of patients are directly notified of significant abnormal lab results) and low-quality testing processes (offices that reported not having a specific follow-up system and who believe that less than 95% of patients are directly notified of significant abnormal lab results).

Coding the reports and errors

We studied the narrative sections of each report and assigned codes to each error and to the contributing factors, harms and consequences identified by the reporter. We coded these elements using the January 2005 version of the *International Taxonomy of Medical Errors in Primary Care (ITME-PC—Version 2)*.¹⁰ We expanded the testing process errors section of the taxonomy to allow more detailed classification. The main categories of testing process errors correspond to the boxes in fig 1 but do not include analytical errors. If a report identified multiple errors, they were numbered chronologically in order of occurrence.

DG, EB, SD and JH coded the event reports. Coding rules were logged, and when necessary, we recoded previously coded reports to maintain consistency. Any coding discrepancies were resolved by consensus among at least three coders. This process resulted in 100% agreement for coding to the first three levels of the taxonomy. The kappa coefficient for coding error types at the highest level the error was coded, generally at the fourth digit level, was 0.338 (95% CI: 0.24, 0.45; $p < 0.0001$) on a 10% sample of error reports, which indicates good general agreement among raters.

Outcomes

We coded negative outcomes as “adverse outcomes” and “harm.” The categories for adverse consequences are: no consequence, time, financial, delay in care, pain, suffering, clinical consequence and unknown. Because of the small sample sizes for two categories (time and financial), we grouped these consequences into one category for this analysis. The coders classified harm according to the United States Pharmacopoeia MedMARx Error Outcome Categories (8, p. 293), which is based on the risk-assessment index developed by the National Coordinating Council for Medication Error Reporting and Prevention.

Descriptive analysis

Descriptive analyses were performed using frequencies and percentages to summarise characteristics of the event reports, patients involved in the events, error types, contributing factors, and outcomes. Chi-square tests were performed to test associations between demographic and event report variables and the types of error reported.

Exploratory univariate and multivariate analyses of predictors of outcomes

Exploratory unadjusted and adjusted analyses were performed to determine associations for adverse consequences and harm, and a defined set of 16 characteristics derived from the event reports and the practice lab surveys that the literature and our own research and clinical experience suggest might be associated with adverse consequences and harm. These include characteristics of the event reports (reporter type, intensive versus no

Table 1 Types of 966 testing process errors reported by family physicians and their office staffs in 590 event reports

	n (%)
Test ordering	125 (12.9)
Needed test not ordered	29 (3.0)
Wrong test ordered	11 (1.1)
Unnecessary test ordered	16 (1.7)
Ordered test at wrong time	2 (0.2)
Contraindicated test ordered	1 (0.1)
Wrong test/patient name recorded in log	7 (0.7)
Test not entered into log	3 (0.3)
Lab order misinterpreted	2 (0.2)
Incomplete or illegible lab order slip	40 (4.1)
Errors in ordering investigations (not otherwise specified)	14 (1.4)
Test implementation	173 (17.9)
Requested test not done (including specimen not drawn, image not booked)	65 (6.7)
Specimen improperly collected or stored/old or inadequate specimen	22 (2.3)
Specimen lost	23 (2.4)
Specimen/patient sent to wrong facility	5 (0.5)
Delay in obtaining specimen	3 (0.3)
Wrong specimen obtained	1 (0.1)
Stat or urgent test not processed or scheduled urgently	7 (0.7)
Wrong test performed or scheduled	21 (2.2)
Right test performed wrongly	3 (0.3)
Failure to instruct patient how to prepare for investigation	1 (0.1)
Test done, but results lost	5 (0.5)
Failure to alter medications for diagnostic procedure	4 (0.4)
Errors in implementing investigations (not otherwise specified)	13 (1.3)
Reporting results to the clinician	238 (24.6)
Failure to report test results in a timely manner	58 (6.0)
Failure to report correct results (wrong values on report)	11 (1.1)
Results never received by office	52 (5.4)
Incorrect interpretation of results by facility or laboratory	3 (0.3)
Previous results, images and specimens could not be found for comparison	1 (0.1)
Incorrect/incomplete information on report	49 (5.1)
Failure to report test results to provider requesting test	46 (4.8)
Errors in reporting investigations to office (not otherwise specified)	18 (1.9)
Clinician responding to the results	64 (6.6)
Responded incorrectly to test results	4 (0.4)
Failure to notice or respond to abnormal test results	16 (1.7)
Failure to notice or respond to abnormal test results in a timely manner	27 (2.8)
Inappropriately responded to incomplete test results	3 (0.3)
Failure to notice or respond to normal test results	1 (0.1)
Failure to notice or respond to normal test results in a timely manner	5 (0.5)
Responding to investigation results (not otherwise specified)	8 (0.8)
Notifying the patient of results	66 (6.8)
Failure to notify patient of test result	17 (1.8)
Failure to notify patient of test result in a timely fashion	37 (3.8)
Failure to notify patient of test result in a sensitive manner	1 (0.1)
Test results given to wrong patient	6 (0.6)
Informed patient about same result more than once	2 (0.2)
Incorrect test results given to patient	2 (0.2)
Notifying patients of investigation results (not otherwise specified)	1 (0.1)
Administrative	170 (17.6)
Filing system	95 (9.8)
Chart completeness and availability	51 (5.3)
Patient flow	6 (0.6)
Message handling	1 (0.1)
Appointments	17 (1.8)
Treatments	17 (1.8)
Medication errors	17 (1.8)

Continued

Table 1 Continued

	n (%)
Communication	55 (5.7)
Errors in communication with patients	16 (1.7)
Communication with other providers sharing patient care	26 (2.7)
Errors in communication between whole healthcare team	13 (1.3)
Other process errors	52 (5.4)
Building infrastructure or management errors	5 (0.5)
Equipment	32 (3.3)
Insurance-related errors	8 (0.8)
Wrongly charged	1 (0.1)
Workload management	6 (0.6)
Knowledge and skills	6 (0.6)
Failure to follow standard or recommended clinical procedure	5 (0.5)
Failure to follow standard/recommended administrative practice	1 (0.1)
Total	966 (100.0)

intensive reporting period); the errors (type and occurrence location), the patients (gender, age, race/ethnicity, presence of a chronic medical condition) and the practices (residency versus non-residency, number of labs used, presence or absence of an EMR, testing processes quality rating (high/low), and four general characteristics of good test tracking processes (present or absent).

For the univariate and multivariate analyses, we dichotomised the outcomes to harm/no harm and adverse consequence/no adverse consequence. Event reports for which harm or consequences could not be determined were excluded from these analyses. Univariate logistic regression analyses were first used to determine unadjusted associations between the event outcome and each practice, patient, error and lab test characteristic. To determine adjusted associations, multivariate logistic regression analyses were performed, adjusting for any practice, patient, error or lab characteristic that was associated with harm or adverse consequences at the alpha level of 0.20 from the univariate analyses. Because the interclass correlation coefficient was 0.9% for consequences and 2.0% for harm, multilevel analyses were performed. Fifty-one event reports did not identify a specific patient involved in the error. Therefore, any analysis that used patient characteristics excluded 51 event reports (n = 539).

Assignment of error type to each event report

For the report-level analysis of associations between type of error and the other 15 characteristics described above, it was necessary to assign a single error type to each event report. For the 288 event reports (49%) that contained more than one error, we tested three methods for assigning error type to these multi-error event reports. The first method selected the first error that occurred; the second method selected the last error that occurred; and the third method randomly selected one of the errors. The distributions of error types defined by the three methods were similar for the 590 event reports. Because the method of random selection eliminated bias toward the order of the reported error, we used this method for assignment of error type to the multi-error event reports. For the analyses, all errors were aggregated into the 10 mutually exclusive major categories listed in table 1.

All analyses were performed using SAS 9.1. For unadjusted and adjusted logistic regression analyses, the glimmix macro was used. This study was reviewed and approved by the University of Missouri-Kansas City Social Sciences Institutional

Table 2 Descriptors of 590 testing process event reports submitted by family physicians and their office staffs

Reporting period	n (%)
Non-intensive	370 (63)
Intensive*	204 (35)
Don't know	16 (3)
No. of errors per event report	
1	302 (51)
2	208 (36)
3	72 (12)
4	8 (1)
Error cascades†	
Single error	302 (51)
More than one error	
Independent errors	29 (5)
Cascade of errors	231 (39)
Cascade and independent errors	28 (5)
Place of occurrence	
Offsite	117 (20)
Onsite	360 (61)
Both	79 (13)
Unknown	34 (6)
Specific location of the error(s)‡	
Reporter's office	412 (60)
Nursing home	6 (1)
Hospital	39 (6)
Emergency department	4 (1)
Onsite lab	54 (8)
Offsite lab	126 (18)
Radiology	30 (4)
Other site (eg, home, other office)	11 (2)
Estimated frequency of occurrence in this practice	
This is the first time it has occurred	60 (10)
Seldom (once or twice per year)	111 (19)
Sometimes (three to 11 times per year)	209 (35)
Frequently (more than once per month)	204 (35)
Not reported	6 (1)
Seriousness	
Not very serious	97 (16)
Somewhat serious	216 (37)
Serious	143 (24)
Very serious	89 (15)
Extremely serious	39 (7)
Not reported	6 (1)
Actions taken as a result of the error‡	
No intervention/activity specific to event	71 (12)
Medical attention or intervention	117 (19)
Other intervention	298 (48)
Unable to determine	121 (20)
System, protocol, or practice change	8 (1)

*During intensive reporting periods, participants were instructed to report every error in the testing process that they observed. Intensive reporting occurred for four of the 32 reporting weeks for each practice at staggered intervals.

†An error "cascade" was noted if there was a clear causal connection between at least two of the errors reported in an event report; if a report included three or four errors, the report might include an error cascade and an independent error.

‡Totals to more than 590 because multiple locations/actions taken were reported in some event reports.

Review Board and by individual site institutional review boards as required.

RESULTS

Office testing process survey results

Four of the eight practices had an electronic medical record. Two practices used only one lab for testing, and six used two or

Table 3 Characteristics of 539 patients involved in testing process errors reported by family physicians and their office staffs*

	n (%)
Sex	
Female	339 (64)
Male	187 (36)
Total	526
Age	
0–17	36 (7)
18–44	169 (33)
45–64	195 (37)
≥65	119 (23)
Total	519
Race/ethnicity	
Hispanic	28 (7)
White	307 (74)
Black	68 (16)
Other	11 (3)
Total	414
Chronic medical problems	
Yes	243 (68)
No	112 (32)
Total	355
Complex medical problems	
Yes	190 (53)
No	169 (47)
Total	359

*Totals for each variable are less than 539, because reporters did not know or did not report the data for every variable for all reports. Percentages are the percentages of patients for which the variable was reported.

three labs. Six offices reported having a system for tracking abnormal results or results needing action until the action was completed. Five practices said they informed patients about normal labs at least 75% of the time, and four practices informed patients about clinically insignificant abnormal results at least 75% of the time. Five of the eight practices claimed to follow-up on abnormal labs greater than 95% of the time.

Event reports

During the 32-week reporting period, the 243 participants submitted 661 event reports. Of these, 590 events had 966 errors related to the testing process. Staff submitted 51% of the reports, physicians submitted 39%, nurse practitioners and physician assistants submitted 8%, and residents submitted 2%. Reporting volume varied from a high of 25.8 reports per reporter in one practice to 1.5 reports per reporter in another, with a mean of 4.1 reports per reporter across all practices. Other characteristics of the event reports are given in table 2.

Approximately one-third of the event reports were submitted during the four intensive reporting weeks. During the routine and intensive reporting period, there were 1.7 and 6.4 event reports per practice per week, respectively. Half of the reports (49%) contained more than one error, and 90% of these reports had "error cascades" in which one error led directly to the next.¹¹

The demographic characteristics reported for the 539 patients involved in the events are summarised in table 3.

Among patients with available data, 64% were female, 70% were 18–64 years old, and 57% were white; 68% were reported to have chronic health problems, and 53% had complex health problems.

Table 4 Summary of testing process errors by practice site (column %)

	1	2	3	4	5*	6*	7*	8*	Total
Test ordering	25 (15)	10 (10)	12 (13)	18 (28)	43 (15)	5 (4)	7 (10)	5 (8)	125 (13)
Test implementation	35 (21)	19 (19)	23 (24)	15 (23)	28 (10)	45 (35)	3 (4)	5 (8)	173 (18)
Reporting results to clinician	35 (21)	31 (31)	13 (14)	7 (11)	89 (32)	36 (28)	8 (12)	19 (29)	238 (25)
Clinician responding to results	7 (4)	8 (8)	9 (10)	1 (2)	15 (5)	9 (7)	9 (13)	6 (9)	64 (7)
Notifying the patient of results	6 (4)	7 (7)	7 (7)	3 (5)	15 (5)	6 (5)	14 (21)	8 (12)	66 (7)
Administrative	20 (12)	9 (9)	21 (22)	11 (17)	49 (18)	22 (17)	22 (32)	16 (24)	170 (18)
Treatments	5 (3)	3 (3)	0 (0)	1 (2)	5 (2)	1 (1)	2 (3)	0 (0)	17 (2)
Communications	13 (8)	10 (10)	5 (5)	4 (6)	17 (6)	0 (0)	2 (3)	4 (6)	55 (6)
Other process errors	18 (11)	2 (2)	3 (3)	4 (6)	17 (6)	5 (4)	0 (0)	3 (5)	52 (5)
Knowledge and skills	2 (1)	0 (0)	2 (2)	0 (0)	1 (0)	0 (0)	1 (2)	0 (0)	6 (1)
Total no. of errors for each practice	166 (100)	99 (100)	95 (100)	64 (100)	279 (100)	129 (100)	68 (100)	66 (100)	966 (100)
Percentage of errors contributed by each practice	(17)	(10)	(10)	(7)	(29)	(13)	(7)	(7)	(100)

*Family medicine residency clinics.

Error types and associations

Table 1 shows the distribution of the types of 966 reported testing process errors. The most common types of errors were related to reporting results to the clinician (24.6%), implementing tests (17.9%), general administrative errors such as filing and chart availability (17.6%) and test ordering (12.9%). There was considerable variability in the distribution of error types reported from each practice site (table 4).

Clinicians reported a greater proportion of administrative errors (18% vs 13%) and errors in responding to test results (8% vs 3%), and office staff reported more test ordering errors (18% vs 11%) ($\chi^2 = 25.4$, $p = 0.001$). Participants from the residency practices reported a greater proportion of administrative errors (19% vs 11%) and errors related to reporting results to the clinician (32% vs 24%), while participants from the non-residency practices reported more errors in test ordering (19% vs 12%) and test implementation (24% vs 18%) ($\chi^2 = 21.5$, $p = 0.003$). Compared with routine reporting, intensive reporting periods had a higher proportion of test ordering (19% vs 13%) and test implementation (27% vs 17%) errors and a slightly lower proportion of errors reported in all other categories ($\chi^2 = 18.0$, $p = 0.012$).

Table 5 Adverse consequences and harms of 590 testing process events reported by family physicians and their staffs

	n (%)
Adverse consequence*	
No consequence	151 (26)
Time/financial	128 (22)
Delay in care	144 (24)
Pain/suffering	63 (11)
Adverse clinical consequence	14 (2)
Unknown	90 (15)
Harm	
No harm	318 (54)
Harm	109 (18)
Emotional harm	33 (6)
Temporary physical harm	69 (11)
Temporary physical harm that required hospitalisation	1
Permanent harm	3 (0.5)
Unspecified harm	3 (0.5)
Unknown	163 (28)

*If more than one adverse consequence was noted, the most serious consequence was assigned to the event report.

†Classified according to the US Pharmacopoeia severity of harm classification.

Among the 416 event reports in which race/ethnicity was reported, minority patients were more likely to have errors of test implementation (32% vs 18%) and less likely to have administrative errors (8% vs 17%) ($\chi^2 = 16.8$, $p = 0.018$). There was no difference in error-type distribution by patients' age.

Error types also varied depending on lab survey responses. Practices that used more than one laboratory were much more likely to experience errors in testing ordering (17% vs 5%), but much less likely to report errors related to test implementation (18% vs 30%) ($\chi^2 = 21.68$, $p = 0.003$). Practices with a specific system to monitor the testing process reported a much lower proportion of errors related to test implementation (15% vs 33%)

Table 6 Multivariable analysis: adverse consequence versus no adverse consequence (365 reports were included in this analysis)

Characteristic	Odds ratio 95% CI (lower to upper)	p Value
Clinician type		
Staff	1.00 (1.00 to 1.00)	0.108
Clinician	1.64 (0.97 to 2.77)	
Age category		
0–17 years	1.00 (1.00 to 1.00)	0.200
18–44 years	1.91 (0.66 to 5.53)	
45–64 years	3.07 (1.06 to 8.93)	
≥65	2.27 (0.77 to 6.65)	
Race		
White, not Hispanic	1.00 (1.00 to 1.00)	0.035
All others	3.07 (1.32 to 7.14)	
Error type		
Test ordering	1.00 (1.00 to 1.00)	0.041
Administration	0.80 (0.35 to 1.85)	
Test implementation	4.18 (1.64 to 10.67)	
Reporting results to the clinician	1.11 (0.54 to 2.29)	
Clinician responding to the results	1.41 (0.44 to 4.48)	
Notifying the patient of results	2.54 (0.70 to 9.30)	
Other errors*	1.53 (0.56 to 4.16)	
Quality of office system for follow-up of abnormal test results		
Low quality†	1.00 (1.00 to 1.00)	0.995
High quality‡	1.00 (0.31 to 3.17)	

*Includes "communication errors" and "other process errors," combined because of the small sample size.

†Less than 95% of patients are directly notified of significant abnormal lab results, and the practice does not have a specific system to ensure testing follow-up.

‡More than 95% of patients are directly notified of significant abnormal lab results, and the practice does have a specific system to ensure testing follow-up.

($\chi^2 = 29.2$, $p < 0.001$). We found no statistical associations between having an electronic medical record and error types.

Outcomes

Event outcomes are displayed in table 5. A quarter of the errors resulted in delays in care, and 13% caused pain, suffering, or a definite adverse clinical consequence. Harm resulted from 18% of the events, and the coders could not determine if there was patient harm in 28%.

Predictors of outcomes

Adverse consequences

Race/ethnicity, age, and error type were associated with adverse consequence at the univariate level. Minority patients had higher odds of experiencing an adverse consequence compared with white, non-Hispanic patients (OR: 2.74, 95% CI (1.45, 5.18); $p = 0.017$), as did patients between the ages of 45 and 64 years of age compared with 0- to 17-year-olds (OR: 2.97, 95% CI (1.30, 6.80); $p = 0.047$). Patients with errors of test implementation had higher odds of adverse consequences compared with patients with errors of test ordering (OR: 3.27, 95% CI (1.55, 6.87); $p = 0.001$). In the adjusted model (table 6), minority race/ethnicity status (OR 3.07) and implementation errors (OR 4.18) remained significant predictors of adverse consequences.

Harm

Race/ethnicity and error type were associated with harm at the univariate level. The odds of harm from errors in implementing tests were higher compared with errors in test ordering (OR: 5.06, 95% CI (2.39, 10.75); $p < 0.001$), and minority patients had higher odds of experiencing harm compared with white, non-Hispanic patients (OR: 2.42, 95% CI (1.40, 4.19); $p = 0.016$). In the adjusted analysis (table 7), errors in implementing tests remained a significant predictor of harm (OR: 5.32, 95% CI (2.22, 12.76); $p < 0.001$), and there was a trend toward increased harm among minority patients (OR: 2.27, 95% CI (1.09, 4.73); $p = 0.066$).

DISCUSSION

We have examined errors in the testing process reported from eight diverse family medicine offices from across the United States. Family physicians and their office staffs were willing and able to identify and report in reasonable detail a wide variety of errors in the testing process. Although one cannot determine true error rates from error reporting studies, testing process errors appear to be common. Since many errors are undetected or unreported, we can assume this is an extreme lower bound and believe the volume of errors reported during the intensive reporting weeks supports this supposition.

Each practice reported errors across the spectrum of the testing process, regardless of the level of sophistication of their testing process management systems. About one-third of the errors related to ordering tests and getting them done. One-quarter of the errors were related to getting the results back to the ordering clinician in a timely fashion. One-fifth of the errors were general administrative errors such as misfiling. These administrative errors cut across the spectrum of the testing process and could not be attributed to a specific step in the testing process.

The small proportion of errors we classified as communication errors (6%) needs comment. Because we chose to use a classification system with mutually exclusive categories, it does

not accurately represent the contribution of communication to patient safety events. Communication difficulties are inherent in many, perhaps most errors in physician offices.¹¹ In another error-reporting study from family physician offices that used a multi-axial coding system, a communication problem was present in 71% of event reports.³ No analytical errors were reported. Analytic errors represent less than 10% of testing process errors and are unlikely to be observed in primary care offices.¹²⁻¹³ Of note, there were no reported errors of failure to monitor patient response to an action based on a test result, the last step of the testing process (fig 1). For example, a patient's coumadin dose might be adjusted based on INR result, and the physician may have failed to repeat the INR at the appropriate interval after the adjustment. Absence of reported errors from this last testing process step could reflect that clinicians and staff do not consider this a testing process step but rather a treatment or patient adherence error.

The error-type distribution was not associated with most of the measured variables related to event report, practice or patient. There was minor variation in the types of errors reported by clinicians and staff and from residency and non-residency practices, but these seem to have little practical importance. "Intensive" reporting appears to pick up more "front end" problems. Forty-six per cent of the errors during the intensive reporting periods were errors of ordering and implementation compared with 30% during the non-intensive periods.

There was considerable variation, however, in the types of errors reported from each practice. For example, test ordering errors comprised 28% of the errors reported from one practice but only 4% from another. This suggests that each practice must examine its own testing process to discover the weak links. For example, practices lacking a specific system to monitor the testing process were twice as likely to report errors related to test implementation. Curiously, having an EMR appeared to have no effect on the type of errors reported. The lack of effect of an EMR supports recent studies showing that

Table 7 Multivariable analysis: harm versus no harm (312 event reports are included in this analysis)

Characteristic	Odds ratio 95% CI (lower to upper)	p Value
Clinician type		
Staff	1.00 (1.00 to 1.00)	0.246
Clinician	1.45 (0.82 to 2.56)	
Race		
White, not Hispanic	1.00 (1.00 to 1.00)	0.066
All others	2.27 (1.09 to 4.73)	
Error type*		
Test ordering	1.00 (1.00 to 1.00)	<0.001
Administration	0.65 (0.21 to 2.00)	
Test implementation	5.32 (2.22 to 12.76)	
Reporting results to the clinician	0.89 (0.34 to 2.27)	
Clinician responding to the results	0.66 (0.15 to 2.85)	
Notifying the patient of results	3.55 (1.19 to 10.63)	
Other errors*	3.37 (1.23 to 9.20)	
Quality of office system for follow-up of abnormal test results		
Low quality†	1.00 (1.00 to 1.00)	0.872
High quality‡	0.92 (0.34 to 2.50)	

*Other errors category combined communication errors with process errors.

†Less than 95% of patients are directly notified of significant abnormal lab results, and the practice does not have a specific system to ensure testing follow-up.

‡More than 95% of patients are directly notified of significant abnormal lab results, and the practice does have a specific system to ensure testing follow-up.

having an EMR can actually lower quality if they are not configured to appropriately support care.¹⁴

The most important association with error type and adverse outcomes is the race/ethnicity of the patient involved. Errors of test implementation were nearly double for minority groups compared with non-Hispanic whites (32% vs 18%). Minority patients were also more likely to experience adverse consequences. The odds of a minority patient suffering an adverse consequence from a testing process error was three times that of a white non-Hispanic patient, even after adjusting for the association between implementation errors and race/ethnicity. Minority patients were twice as likely to experience harm compared with white, non-Hispanic patients. Coders were blinded to racial/ethnic status during the coding procedures, so these findings are not due to coder bias. Further investigation in a more representative sample of practices is required.

A significant portion of testing process errors result in adverse outcomes. We classified outcomes in two ways, and depending on which of these outcome categories one uses, adverse consequences or harm occurred in 18% to 59% of the events. What could be labelled as inefficiency and inconvenience (lost time, greater cost, delays in care) occurred in nearly half of the events. Thirteen per cent of events resulted in adverse clinical outcomes to patients, and 18% resulted in physical or emotional harm. Our analysis of predictors of adverse outcomes suggest that clinicians and staff should focus improvement efforts on the implementation and notification steps of the testing process, and researchers should focus on race/ethnicity.

Study limitations

There are several limitations to this study. We did not use a standard definition of medical error, and we did this intentionally. The definition we used, anything you observe “that should not have happened and that you don’t want to happen again,” has been highly successful in eliciting error reports in three of our prior studies of errors reported by family physicians. Next, the study has all of the usual limitations of error-reporting studies, which cannot give a true distribution or frequency of all errors that occur.¹⁵ For example, the lab and radiology facilities used by the offices would have observed other errors that were not reported by the office staff and physicians. Chart review would have provided yet another set of errors. While the true distribution of testing process errors in family physician offices may be different, our data lead to some testable hypotheses and needed exploration. Second, the sample of practices is small and may not be representative, but we were purposeful about our selections to improve generalizability. Third, the demographic and outcomes data from reporters were incomplete (racial/ethnic status was not reported for 23% of patients), so the results of the analysis of patient variables and outcomes may have been different with a complete data set. We used an anonymous system, rather than a confidential system, so we could not follow-up with the reporter to ask questions and get additional information. Fourth, the reports are from physicians and office staff. Lab staff and patients may have identified other errors, and the frequency distributions of errors they observe and report could be much different. Fifth, despite this being the largest study to date of testing process errors reported from primary-care offices, the sample size was relatively small for testing associations and predictors of adverse consequences and

harm. Finally, the classification system we used provides for mutually exclusive classification of errors which provides a broad brush picture of error encounters and is, therefore, best for research projects but lacks the texture of multi-axial coding systems, which are best used for case-by-case analysis in active quality improvement systems.

We believe that when taken as a whole, this study in conjunction with others (refer to American Academy of Family Physicians/Elder prior studies and Applied Strategies for Improving Patient Safety, etc) strongly supports the need for office-by-office improvements in the overall testing process within primary care. Even the lower-bound estimates of frequency and harm provided by this report are unacceptable. Given the volume of lab and imaging studies performed or ordered through the primary-care system, the extent of harm, inconvenience and waste caused by errors is significant. Clinicians, office administrators, and office staff need to assess their internal systems in these areas in an ongoing fashion and address discovered weaknesses in a timely manner. Further work to develop solutions at both the monitoring and improvement stages is needed.

Acknowledgements: JH 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. We wish to acknowledge the staff and physicians of the eight family medicine offices who submitted event reports, W Pace for his critique, and E Staton and K Hitchcock for manuscript preparation and review.

Funding: The study was funded in part by federal grants R21 HS13554-01 and P20 HS11584-02 from the Agency for Health Care Research and Quality.

Competing interests: None.

Ethics approval: Ethics approval was obtained.

REFERENCES

1. **Nutting PA**, Main DS, Fischer PM, *et al*. Toward optimal laboratory use. Problems in laboratory testing in primary care. *JAMA* 1996;**275**:635–9.
2. **Elder NC**, Vonder Meulen M, Cassidy A. The identification of medical errors by family physicians during outpatient visits. *Ann Fam Med* 2004;**2**:125–9.
3. **Fernald DH**, Pace WD, Harris DM, *et al*. Event reporting to a primary care patient safety reporting system: a report from the ASIPS Collaborative. *Ann Fam Med* 2004;**2**:327–32.
4. **Ross J**, Boone D. Assessing the effect of mistakes in the total testing process on the quality of patient care. In: Martin LM, ed, for the Institute on Critical Issues in Health Laboratory Practice. Proceedings of the 1989 Institute on Critical Issues in Health Laboratory Practice: Improving the Quality of Health Management Through Clinician and Laboratorian Teamwork. Wilmington, DE: DuPont, 1991.
5. **Hickner JM**, Fernald DH, Harris DM, *et al*. Issues and initiatives in the testing process in primary care physician offices. *Jt Comm J Qual Saf* 2005;**31**:81–9.
6. **Dovey SM**, Meyers DS, Phillips RL Jr, *et al*. A preliminary taxonomy of medical errors in family practice. *Qual Saf Health Care* 2002;**11**:233–8.
7. **Makeham M**, Dovey SM, County M, *et al*. An international taxonomy for errors in general practice: a pilot study. *Med J Aust* 2002;**177**:68–72.
8. **Aspden P**, for the Institute of Medicine. *Patient Safety: Achieving a New Standard for Care*. Washington, DC: National Academy Press, 2004.
9. **Rosser W**, Dovey S, Bordman R, *et al*. Medical errors in primary care: results of an international study of family practice. *Can Fam Physician* 2005;**51**:386–7.
10. **The Linnaeus-PC Collaboration**. *International Taxonomy of Medical Errors in Primary Care—Version 2*. Washington, DC: The Robert Graham Center, 2002.
11. **Woolf SH**, Kuzel AJ, Dovey SM, *et al*. A string of mistakes: the importance of cascade analysis in describing, counting, and preventing medical errors. *Ann Fam Med* 2004;**2**:317–26.
12. **Lapworth R**, Teal TK. Laboratory blunders revisited. *Ann Clin Biochem* 1994;**31**:78–84.
13. **Chambers AM**, Elder J, O’Reilly DS. The blunder-rate in a clinical biochemistry service. *Ann Clin Biochem* 1986;**23**:470–3.
14. **Crosson JC**, Stroebel C, Scott JG, *et al*. Implementing an electronic medical record in a family medicine practice: communication, decision making, and conflict. *Ann Fam Med* 2005;**3**:307–11.
15. **Thomas EJ**, Petersen LA. Measuring errors and adverse events in health care. *J Gen Intern Med* 2003;**18**:61–7.