



OPEN ACCESS

# When diagnostic testing leads to harm: a new outcomes-based approach for laboratory medicine

Paul L Epner,<sup>1</sup> Janet E Gans,<sup>2</sup> Mark L Graber<sup>3</sup>

<sup>1</sup>Paul Epner LLC, Evanston, Illinois, USA

<sup>2</sup>Evanston, Illinois, USA

<sup>3</sup>Healthcare Quality and Outcomes, RTI International SUNY Stony Brook School of Medicine, St James, New York, USA

## Correspondence to

Paul L Epner,  
1501 Hinman Ave, #7B,  
Evanston, IL 60201, USA;  
PEpner@ChicagoBooth.edu

Received 19 March 2013

Revised 26 July 2013

Accepted 29 July 2013

Published Online First

16 August 2013

Many diagnostic errors are associated with laboratory testing, and many of these are preventable. However, a reduction in testing-related diagnostic errors (TDE) is hindered by the absence of a well-defined relationship between diagnostic harm and the testing process (whether from laboratory or non-laboratory sources) as well as by a lack of relevant measures for evaluation. The goal of this paper is to review current models that describe the testing process, and then propose a different approach to facilitate the reduction of diagnostic errors and harm related to diagnostic testing. We then demonstrate how this approach can be used to develop measures that may improve patient outcomes and guide future research to reduce TDE. Finally, we highlight the need for collaboration between clinicians and laboratory physicians and scientists to achieve these goals.

## THE ROLE OF LABORATORY TESTING IN ESTABLISHING DIAGNOSES

Diagnoses typically result from the patient history and physical. However, diagnostic testing is often used to confirm initial impressions or rule out alternatives, and at least 10% of all diagnoses are not considered final until clinical laboratory testing is complete.<sup>1 2</sup> This number most likely underestimates the actual impact of testing on diagnosis. In the emergency room, clinical laboratory testing is ordered in more than 41% of all visits.<sup>3</sup> Family physicians order tests in 29% of all patient visits, and general internists, in 38% of visits.<sup>4</sup> These percentages would be even higher if the calculations were based only on the 33.9% of primary care visits that involve a new complaint.<sup>5</sup>

Advances in technology have also contributed to the increased importance of laboratory tests. In the past, laboratory

tests were used to identify organ and system dysfunctions or diseases. While this is still true, testing nowadays is used to diagnose disease subtypes, as occurs when pathology reports of cancer are accompanied by tumour-specific and patient-specific molecular analyses, data which help physicians determine optimum therapies and a patient's likely response to treatment.<sup>6 7</sup> Laboratory testing is also increasingly being used to diagnose treatment failures associated with newer measures of effective care, such as reduced hospital readmissions.<sup>8</sup> The clinical laboratory's growing significance may also reflect physicians' increasing reliance on objective data from diagnostic testing to partially compensate for reduced history and physical examination skills.<sup>9</sup>

## THE TOTAL TESTING PROCESS AND LIMITATIONS OF THE PROCESS-STEP APPROACH

The concept of the 'Total Testing Process' (TTP) was first defined by Gambino in 1970<sup>10</sup> and later became the familiar nine-step 'Brain-to-Brain turnaround time loop' described by George Lundberg in 1981<sup>11</sup> and modified in 2011 (see figure 1).<sup>12</sup>

In Lundberg's model, the value of laboratory results is influenced by events that occur before the sample reaches the laboratory and after the results are released from it. His model encompasses the physician's cognitive involvement at the start of the process and at the end. Some researchers have adapted Lundberg's TTP model to specific settings, as did Hickner for primary care physicians<sup>4</sup> and Raab for oncologic pathology diagnosis.<sup>13</sup> These process-step models are useful insofar as they simplify and clarify a complex process and identify nodes where errors can occur.

Nevertheless, the linearity and simplicity of these models understate significant



Open Access  
Scan to access more  
free content



► <http://dx.doi.org/10.1136/bmjqs-2013-002387>

**To cite:** Epner PL, Gans JE, Graber ML. *BMJ Qual Saf* 2013;**22**:ii6–ii10.

and intentional process variations that occur regularly as is the case with send-out testing (which is testing referred to another site), reflex testing (which refers to preauthorised follow-up testing linked to results observed in the lab), and add-on testing (which entails added testing not included on the original order but applied to the original sample). Each of these types of testing has important and potentially error-prone permutations. For example, when a laboratory lacks on-site capability to perform a test, a portion of the sample is sent to a reference lab for analysis (send-out test). This triggers additional process steps, such as dividing and repackaging samples for shipment. Aside from making the testing process more complex, test results typically become available to the physician later than other, co-ordered and locally analysed tests, thereby requiring multiple efforts to retrieve results. Furthermore, test result formats (eg, test names, significant digits or units of measure) can differ from those produced by the local information technology system, and complete results may not be visible except on scanned documents which can be difficult to retrieve in electronic medical record systems. A recent AHRQ-funded project identified 40 risk factors for diagnostic error associated with send-out testing. (Graber M, Morgan LC, Tant E, *et al.* Proactive risk assessment during the laboratory testing process to reduce diagnostic error: Literature Review. 2012. For the Agency for Healthcare Research and Quality. Contract HHS29032001T38 (unpublished)). Yet, we could find only one study that examined problems

related to send-out testing, and this focused solely on order entry problems.<sup>14</sup>

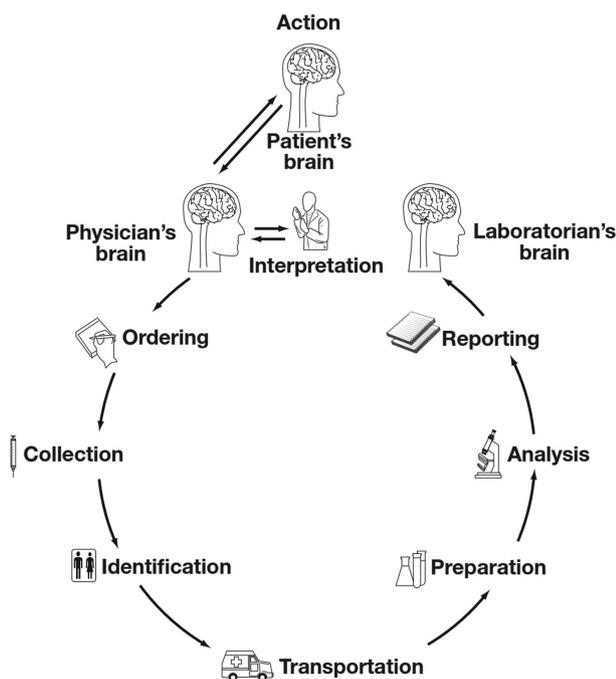
The process-step models currently in use also oversimplify the complexity of routine testing. One analysis identified over 80 distinct, planned process steps and a dozen handoffs of information or material, each one having the potential for failure and additional steps for remediation.<sup>15</sup> Additionally, these models rarely consider patient harm that occurs when the testing process is never initiated, as happens when a physician fails to order an appropriate test.

### CURRENT MEASURES OF PERFORMANCE IN THE TOTAL TESTING PROCESS

Clinical laboratories typically assess their performance based on measures of laboratory efficiency and internal quality rather than patient outcomes. For example, turnaround times typically measure in-laboratory sample receipt to the issuance of results. Measures of quality defects generally pertain to those that reduce productivity (eg, sample haemolysis, insufficient quantity of sample, missing sample identification, etc.).

For more than a decade, the College of American Pathologists (CAP) has offered a performance monitoring service for clinical laboratories known as Q-Probes and Q-Tracks. Q-Probes are periodic surveys used to develop key benchmarks, and Q-Tracks provide longitudinal performance monitoring.<sup>16</sup> Since 2009, the four Q-Probes for clinical pathology have focused on laboratory management (two probes), labelling errors, and clinician satisfaction. The current catalogue for Q-Tracks lists 11 measures for the clinical laboratory, eight of which CAP deems to be related to patient safety, but most of the measures assess only the laboratory's portion of TTP. For example, the measure of 'Stat Test Turnaround Time Outliers' ignores all time before the sample reaches the laboratory and after the result is released from the laboratory, that is, from result release to clinician action. The 'Critical Values Reporting' survey measures the documentation of successful notification according to the given laboratory's policy, but does not assess the timeliness of actions directed to patient care (<http://www.cap.org>).

Such narrow approaches to process monitoring overlook important sources of patient harm. The disconnect between currently monitored error types and patient harm makes it difficult to set priorities that would improve quality of care and reduce patient harm and suffering. In a recent review article, Plebani argued that quality improvement efforts to reduce laboratory errors are strongly influenced and limited by data collection goals and methods.<sup>17</sup> The issue of priorities is important. More than six billion tests are ordered in the USA each year,<sup>18</sup> and while the majority of process defects may have little impact on patient outcomes, even small percentages of testing-related error can translate into significant harm.



**Figure 1** The 'Brain-to-Brain' loop, depicting the steps in the process of considering, performing and using laboratory tests for diagnosis.

## AN OUTCOMES-BASED APPROACH TO REDUCE PATIENT HARM

For more than a decade, some laboratorians and patient safety researchers have suggested that quality improvement efforts should seek to reduce patient harm rather than process defects where the relationship to patient harm is unclear.<sup>17 19–21</sup> We believe that a unified approach has yet to be developed that can identify, classify and measure outcomes and errors, applicable in both research and clinical settings. A patient-harm-based approach would more likely lead to quality improvement interventions that would reduce testing-related diagnostic error. This approach would encourage the development of measurement tools to systematically monitor the testing process for TDE and also evaluate the effectiveness of potential interventions.

## AN OUTCOMES-BASED APPROACH TO CLASSIFYING TESTING-RELATED DIAGNOSTIC ERROR

Astion *et al* were among the first to recognise and develop a classification system that could be used to prioritise quality improvement initiatives based on actual or potential patient impact associated with the testing process.<sup>22</sup> Schiff *et al* created the Diagnostic Error Evaluation and Research (DEER) classification system that was designed to identify errors at any and every point along the diagnostic process including the testing process.<sup>23</sup> However, the classification schemes developed by Astion and Schiff have yet to be widely used and evaluated, especially for routine clinical use, making opportunities for further innovation appropriate.

Classification systems should be relevant to the task, exhaustive, consist of mutually exclusive categories, and have a high degree of inter-rater reliability when events are categorised. After reviewing the diagnostic error and testing literature, we identified five causes that taken together may explain all important sources of diagnostic error and harm related to the testing process (see box 1). While occurrences of the five causes will not *always* result in diagnostic error, patient harm related to diagnostic testing is highly likely to stem from one of these five causes.

The mechanisms by which these causes lead to diagnostic error can be easily explained. When an *inappropriate test is ordered*, a false positive result can lead to diagnostic error. It may also lead the clinician to interpret results as actionable which, in turn, can lead to unnecessary tests, procedures or treatments which may result in patient harm. As important, *when an appropriate test is not ordered*, the clinician misses key information important to a correct diagnosis.

Even when test ordering is appropriate, the *misapplication of test results* can result from cognitive failures by the clinician, whether from misunderstanding the clinical implications of a result, or from failing to understand the limitations of the test methodology (ie, statistical variations, performance limitations, or

### Box 1: Five causes taxonomy of testing-related diagnostic error

- ▶ An inappropriate test is ordered
- ▶ An appropriate test is not ordered
- ▶ An appropriate test result is misapplied
- ▶ An appropriate test is ordered, but a delay occurs somewhere in the total testing process
- ▶ The result of an appropriately ordered test is inaccurate

interfering substances). Misapplication can also occur when a patient provides erroneous or incomplete information needed to correctly interpret the result. Regardless of origin, any misapplication of results may lead to an erroneous diagnosis.

*Delays in the TTP* may occur at the preanalytical, analytical or postanalytical stage, from initial ordering through the timely retrieval and application of results. Delays are problematic if a patient's health deteriorates during the delay, or if the effectiveness of treatment is compromised.

Finally, the *result of an appropriately ordered test can be inaccurate* due to analytical issues, such as an improperly calibrated instrument, or non-analytical issues, as occurs when a result is assigned to the wrong patient. Both can lead to inappropriate diagnoses.

## AN OUTCOMES-BASED APPROACH TO MEASURING TESTING-RELATED DIAGNOSTIC ERROR

The outcomes-based approach to TDE proposed here represents an important shift for many laboratory personnel, clinicians and others interested in improving the quality of TTP, and requires the development of measures linked to each of the five causes. Examples of such measures are sprinkled throughout the medical literature, but more systematic development is required. For instance, the *ordering of inappropriate tests* is a recognised problem, but is usually considered in relation to cost savings and the goal of reducing test volume.<sup>24–26</sup> With few exceptions,<sup>27</sup> the impact of inappropriate testing on patient outcomes is rarely reported, and the impact on diagnostic error, undocumented. Some laboratories have addressed the risk of false positives by monitoring the ordering of tests that are prone to misordering, such as certain thyroid tests, and some have developed ways to detect inappropriate repetition of test orders, usually associated with standing orders.

The development and implementation of measures that reflect the *failure to order appropriate tests* during a diagnostic work-up will be more straightforward in some instances than in others. For presenting complaints that have standard protocols (eg, troponin tests for chest pain), measuring compliance with the

protocol should be relatively clear-cut, but limitations in information technology may still impede implementation. However, when standard protocols do not exist, the development of measures of the failure to order appropriate tests is more complicated. That is because it is harder to systematically determine a clinician's diagnostic reasoning and then evaluate the appropriateness of the corresponding test selection.

The failure to follow-up actionable test results is one example of an *appropriate test result that is misapplied* and known to be an important problem.<sup>28–30</sup> Singh *et al* measured instances of failure to take appropriate action on abnormal fecal occult blood tests as well as other laboratory tests.<sup>28–31</sup> Kanter did the same for patients without a 90-day follow-up of abnormal creatinine results, and found that 51% of patients who were contacted and retested had undiagnosed chronic kidney disease. (Kanter M. *The Kaiser Permanente Safety Net Program*. Oral presentation at *Diagnostic Error in Medicine 2012*, Baltimore, MD (unpublished)). These studies illustrate how specific measures can be developed using electronic data. Such measures could help monitor TDE performance and generate data needed to prioritise interventions.

A Corrected Results report tracks erroneous results released by the clinical laboratory. It is one measure of *inaccurate test results* that is routinely used, but it is not sufficient. Such reports are typically generated from failures detected by the laboratories themselves. Systematic means of obtaining feedback from other sources, for example, from clinicians receiving absurd test results, is often missing.<sup>32</sup>

Measures of diagnostic errors associated with *testing-related delays* should begin with the clinician's order and end with the action taken by the clinician. (Ideally, measures would also include the time taken by the patient to implement a recommended action.) Currently, the time between the initial clinician-patient encounter and the test order, between the test order and receipt of the sample by the laboratory, and between the test result availability and clinician action based on results, is rarely available or reviewed. The development of these measures hinges on the use of information technology to record and retrieve time-stamps of the testing process.

Although measures to systematically monitor TDE are few in number, some researchers have found innovative ways to identify and examine likely instances of diagnostic error. This is important because reliance on chart reviews for a random sample of patient diagnostic encounters would be inefficient and probably ineffective for routine use. Singh *et al* used a 'trigger' algorithm to identify situations where diagnostic errors were more likely to be found, thus improving efficiency of error detection.<sup>33</sup> Efforts to develop additional decision rules that are sensitive for TDE will be important to establish practical routine monitoring strategies.

## QUESTIONS TO GUIDE FUTURE RESEARCH

Clearly, laboratorians and clinicians should forge stronger links between diagnostic testing and patient outcomes. Without those links, the clinical laboratory will continue to be driven primarily by cost, volume and process measures, similar to the way a factory manages inputs and outputs. By developing measures of patient impact, the relative effectiveness of interventions to reduce diagnostic error can be assessed. To guide the development of new measures and new interventions, additional research is needed that should take into account the following questions:

- ▶ What specific measures can be developed and validated to assess and monitor the harm of testing-related diagnostic error?
- ▶ How often and under what circumstances do the five types of errors proposed in our approach lead to harm associated with an erroneous diagnosis, a missed diagnosis or a delay in diagnosis?
- ▶ What practices would optimise the appropriate ordering of laboratory tests and application of laboratory test results to improve patient outcomes?

## SUMMARY AND CONCLUSIONS

Failures in the ordering of laboratory tests and the application of laboratory test results are major contributors to diagnostic errors, along with residual problems in test performance per se. The Five Cause Taxonomy and the strategy for defining appropriate measures presented here address gaps that have limited significant reductions in TDE and patient harm. Only through a concerted and coordinated effort by laboratory and clinical staff will the benefits associated with this approach be realised. Neither group, if focused only on their separate domains, will be successful. The TTP is too complex, the causes of errors too diverse, and the continuing development of new testing modalities and uses, too rapid.

Our approach offers an opportunity for clinical laboratory physicians and scientists to greatly expand their mission from a factory model focused almost exclusively on providing accurate, timely test results at the lowest possible cost, to a mission that rapidly and efficiently enables the accurate diagnosis of conditions, the selection of appropriate treatments and the effective monitoring of health status. The expertise they bring to the TTP can benefit clinicians and patients enormously, and their leadership could be crucial to success. However, only by working together with clinicians, can the goal of improving the safety of laboratory-supported diagnostic evaluation be achieved.

**Acknowledgements** The authors would like to thank Hardeep Singh, MD, MPH for his thoughtful comments and constructive critique throughout the development of this manuscript.

**Contributors** All authors equally contributed to the submission of this paper.

**Competing interests** None.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>

## REFERENCES

- Peterson MC, Holbrook JH, Von Hales D, *et al.* Contributions of the history, physical examination, and laboratory investigation in making medical diagnoses. The Western Journal of Medicine. *BMJ* 1992;156:163–5.
- Wahner-Roedler DL, Chaliki SS, Bauer BA, *et al.* Who makes the diagnosis? The role of clinical skills and diagnostic test results. *J Eval Clin Pract* 2007;13:321–5.
- National hospital ambulatory medical care survey. Centers for Disease Control and Prevention 2010. [http://www.cdc.gov/nchs/data/ahcd/nhamcs\\_emergency/2010\\_ed\\_web\\_tables.pdf](http://www.cdc.gov/nchs/data/ahcd/nhamcs_emergency/2010_ed_web_tables.pdf)
- Hickner JM, Fernald DH, Harris DM, *et al.* Issues and initiatives in the testing process in primary care physician offices. *Jt Comm J Qual Patient Saf* 2005;31:81–9.
- Woodwell DA, Cherry DK. National Ambulatory Medical Care Survey: 2002 summary. Advance data. 2004;(346):1–44. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15460863>.
- Igbokwe A, Lopez-Terrada DH. Molecular testing of solid tumors. *Arch Pathol Lab Med* 2011;135:67–82.
- Mehta S, Shelling A, Muthukaruppan A, *et al.* Predictive and prognostic molecular markers for cancer medicine. *Ther Adv Med Oncol.* 2010;2:125–48.
- Bettencourt P, Azevedo A, Pimenta J, *et al.* N-terminal-pro-brain natriuretic peptide predicts outcome after hospital discharge in heart failure patients. *Circulation* 2004;110:2168–74.
- Feddock CA. The lost art of clinical skills. *Am J Med* 2007;120:374–8.
- Gambino SR. Met and unmet needs of the automated clinical laboratory. *Trans N Y Acad Sci* 1970;32(7 Series II):816–20.
- Lundberg GD. Acting on significant laboratory results. *JAMA* 1981;245:1762–3.
- Plebani M, Laposata M, Lundberg GD. The brain-to-brain loop concept for laboratory testing 40 years after its introduction. *Am J Clin Pathol* 2011;136:829–33.
- Raab SS, Grzybicki DM. Quality in cancer diagnosis. *CA Cancer J Clin* 2010;60:139–65.
- Valenstein PN, Walsh MK, Stankovic AK. Accuracy of send-out test ordering: a College of American Pathologists Q-Probes study of ordering accuracy in 97 clinical laboratories. *Arch Pathol Lab Med* 2008;132:206–10.
- Epner P. Owning the total testing process. *Adv Administrators Lab* 2008;17:65–6.
- Zarbo RJ, Jones BA, Friedberg RC, *et al.* Q-tracks: a College of American Pathologists program of continuous laboratory monitoring and longitudinal tracking. *Arch Pathol Lab Med* 2002;126:1036–44.
- Plebani M. Errors in clinical laboratories or errors in laboratory medicine? *Clin Chem Lab Med* 2006;44:750–9.
- Wolcott J, Schwartz A, Goodman C. *Laboratory medicine: a national status report.* 2008,385.
- Bonini P, Plebani M, Ceriotti F, *et al.* Errors in laboratory medicine. *Clin Chem* 2002;48:691–8.
- Lippi G, Fostini R, Guidi GC. Quality improvement in laboratory medicine: extra-analytical issues. *Clin Lab Med* 2008;28:285–94, vii.
- Plebani M. The detection and prevention of errors in laboratory medicine. *Ann Clin Biochem* 2010;47(Pt 2):101–10.
- Astion M, Shojania KG, Hamill TR, *et al.* Classifying laboratory incident reports to identify problems that jeopardize patient safety. *Am J Clin Pathol* 2003;120:18–26.
- Schiff GD, Kim S, Abrams R, *et al.* *Diagnosing diagnosis errors: lessons from a multi-institutional collaborative project.* Advances in Patient Safety: From Research to Implementation. 2005;2:255–278.
- Liu Z, Abdullah A, Baskin L, *et al.* An Intervention to reduce laboratory utilization of referred-out tests. *Lab Med* 2012;43:164–7.
- Feldman LS, Shihab HM, Thiemann D, *et al.* Impact of providing fee data on laboratory test ordering: a controlled clinical trial. *JAMA Intern Med* 2013;173:903–8.
- Stuebing EA, Miner TJ. Surgical vampires and rising health care expenditure: reducing the cost of daily phlebotomy. *Arch Surg* 2011;146:524–7.
- Neilson EG, Johnson KB, Rosenbloom ST, *et al.* The impact of peer management on test-ordering behavior. *Ann Intern Med* 2004;141:196–204.
- Singh H, Thomas EJ, Sittig DF, *et al.* Notification of abnormal lab test results in an electronic medical record: do any safety concerns remain? *Am J Med*; 2009;123:238–44.
- Callen JL, Westbrook JJ, Georgiou A, *et al.* Failure to follow-up test results for ambulatory patients: a systematic review. *J Gen Intern Med* 2012;27:1334–48.
- Callen J, Georgiou A, Li J, *et al.* The safety implications of missed test results for hospitalised patients: a systematic review. *BMJ Qual Saf* 2011;20:194–9.
- Singh H, Wilson L, Petersen LA, *et al.* Improving follow-up of abnormal cancer screens using electronic health records: trust but verify test result communication. *BMC Med Inform Decis Mak* 2009;9:49.
- Yuan S, Astion ML, Schapiro J, *et al.* Clinical impact associated with corrected results in clinical microbiology testing. *J Clinical Microbiol* 2005;43:2188–93.
- Singh H, Giardina TD, Meyer AND, *et al.* Types and origins of diagnostic errors in primary care settings. *JAMA Intern Med* 2013;173:418–25.