Diagnosis and diagnostic errors: time for a new paradigm

Gordon D Schiff

Correspondence to
Gordon Schiff, Center for Patient Safety Research and Practice, Division of General Internal Medicine, Brigham and Women's Hospital, Medicine Harvard Medical School, 1620 Tremont St. 3rd Fl, Boston, MA 02120, USA; gschiff@partners.org

Received 22 August 2013
Accepted 22 August 2013
Published Online First 19 September 2013

▸ http://dx.doi.org/10.1136/bmjqs-2013-001874

To cite: Schiff GD. *BMJ Qual Saf* 2014;**23**:1–3.

It looks like diagnosis triggers may be gaining traction. Building on their earlier efforts,[1 2] a team of investigators based in Houston reports on their latest effort to apply electronic screens—so called 'triggers'—to large clinical databases, to identify cases of potential diagnostic errors.[3] They searched nearly 300 000 patients' records over a 12-month period at two large health systems with comprehensive electronic health records. They sought patients who had one of four 'red flag' findings for prostate or colon cancer—elevated prostate specific antigen (PSA), positive fecal occult blood test (FOBT), rectal bleeding (haematochezia), and iron deficiency anaemia. They then used a refined electronic algorithm to cull out patients who (1) were already known to have prostate or colorectal cancer, or (2) had evidence of appropriate follow-up testing or referral. This process left roughly 1500 patients with one of the four red flags potentially unaddressed. Thus, searching an enormous haystack of 300 000 patients, they found roughly 1500 possible 'needles'–patients who may have had their diagnosis of colon or prostate cancer delayed or overlooked entirely.

Their next step was manual chart review. They had hoped that the yield of their electronic screen for diagnostic failures ('positive predictive value') might approach 35%, meaning that at least one out of every three 'screen positive' charts would have evidence for care improvement opportunities. Instead they were pleasantly surprised that fully 2/3 of the charts (positive predictive value of 60–70% depending on which screen for which cancer) had such opportunities, suggesting they could find an estimated 1000+ instances of delayed or missed follow-up representing an estimated 50 actual cancers each year.

The first thing that must be said is that, although the screen 'worked well' (to find care improvement opportunities), the outpatient systems of care obviously did not. Since there is no reason to believe their findings are not broadly representative of ambulatory care in general (and the fact that both the institutions had advanced electronic systems should, in theory, put them in a better position for reliable follow-up than those lacking such capability);, the findings mean that healthcare diagnosis, as measured by this one metric at least, is a long way from *six-sigma* quality (defined as one defect per 3.4 million). This study's rate translates into roughly 13 600 defects per 3.4 million patients. While one could quibble with some of the arbitrary cut-off intervals chosen for this study—a colonoscopy 61 days after a positive FOBT was failed care, whereas, one after 59 days was not; similarly with 91 vs 89 days for follow-up of an elevated PSA—the study unquestionably highlights undesirable delays that more efficient and more reliable care should be able to avoid.

The next important consideration to ponder is whether and how such retrospective 'triggers' can be used to minimise diagnostic errors prospectively. As we have noted previously, prospectively applying such triggers as safeguards to 'find and fix' actual or potential diagnostic errors and delays should be the ultimate application of such triggers.[4] Thus, as impressive as the results of the current application of these cancer electronic trigger screens are, we are still working in what quality improvement practitioners call the 'inspection' rather than the 're-engineering' or improvement mode.[5] In an earlier effort to pilot electronic screens, our diagnostic error research team screened records for potentially missed elevated thyroid stimulating hormone (TSH) levels and was able to intervene and treat multiple patients with overlooked hypothyroidism.[6] The prospect of prospectively intervening on the 1000 patients identified as being at risk for prostate or colorectal cancer in this retrospective study is a tantalising one, but one that awaits a different application and study

Table 1 New paradigms for better diagnosis

Traditional ways of thinking about diagnosis, and diagnostic error	New paradigms/better ways to think about diagnosis, and diagnosis improvement
General concepts	
Good diagnosticians get it right 1st time, most all of the time	Diagnosis is an inexact science, fraught with uncertainty. Goal is to lower error rates and delays via more reliable systems and follow-up
Lore, academic model of the master/skilful diagnostician who knows/recalls everything	Quality diagnosis is based on well coordinated distributed network/team of people and reliable processes. Relying less on human memory
Diagnosis is the doctor's job	Co-production of diagnosis between clinicians (including nurses, social workers, specialists), lab/radiology, and especially, the patient and family
Patients often seen as anxious, exaggerating, overly questioning, with at times unreasonable demands and expectations	Patients as key allies in making diagnosis. Need to address understandable/legitimate fears, desires for explanations. Use their questions to stimulate rethinking of diagnoses.
Diagnosis and treatment as separate stages in patient care	Prioritising diagnostic efforts to target treatable conditions; more integrated strategies and timing for testing and treatment
Clinical practices	
Order lots of tests to avoid missing diagnoses	Judicious ordering: targeted, well organised data and testing. Appreciation of test limitations (false±, incidental findings, risks)
More specialty referrals on one hand, but utilisation barriers (co-pays, prior authorisation) on the other.	Pull systems to lower barriers for raising questions, real-time virtual consults; collaborative approaches to enable watch and wait strategies where appropriate
Frequent empirical drug trials when uncertain of diagnosis	Conservative use of drugs to avoid confusing clinical picture
MD attention/efforts to ensure disease screening	Automating, delegating clerical functions; teamwork, to free MD cognitive time.
Thinking about diagnosis errors and challenges	
Errors classified as either system or cognitive	Most errors/delays rooted in processes and system design/failures. Errors multifactorial with interwoven, interacting and inseparable cognitive—system factors
Errors infrequent; hit-and-miss hearing about errors	Systematically and proactively surveilling of high-risk situations and diagnoses' performance and outcomes
Clinicians reactions: denial, defensive, others to blame, others also making similar errors	Culture of actively and non-defensively seeking to uncover, dig deep to learn from and share errors
Dreading complex diagnostic dilemmas	Welcoming intellectual/professional challenges; getting support (time, help) for more complex patients
Diagnoses as distinct labels, events	Diagnosis as often fuzzy, multifactorial, evolving over time
Documentation/communication	
Documentation: time-consuming/wasting, mindless, mainly to CYA (covering your back)	Documentation as useful tool, friend, CYA=canvass for assessment to reflect and share assessments, unanswered questions
Say and write as little as possible as it could be used against you in malpractice allegation	Share uncertainties to maximise communication and engagement with other care givers, patients
Eschew/hide uncertainty	Leverage, disclose, learn from uncertainty
Don't let patient know about errors so they don't become angry, mistrustful, or sue	Patients have right to honest disclosure; often find out about errors anyway (cancer evolves); anticipate, engage concerns
Patients advised to call if not better; no news is good news (test results: 'we'll call if anything abnormal')	Systematic proactive feedback and follow-up. Calling/emailing to check how patient is doing; survey patient outcomes
Global remedies	
Knowing more medical knowledge	Knowing the patient (including psychosocial, environmental contexts)
Attention to the 'facts' to objectively make diagnosis	Acknowledgement of ubiquitous cognitive biases; efforts to anticipate, recognise, counteract, via various de-biasing strategies
Exhortations to have 'high index of suspicion' of various diagnoses	Less reliance on memory, recall of lectures/reading. Affordances, alerts engineered into work flow; delineation of 'don't miss' diagnoses with design of context-relevant reminders
Ensuring MD is cc'd everything, thorough/voluminous notes, widespread reminders/alerts	Appreciation of detrimental consequences of information and alert overload; strategies to minimise
Redundancies, double checks	Recognition that highly reliable systems are safer than multiple halfway systems
Fear of malpractice to motivate MDs to be more careful and practice defensive medicine	Drive out fear, making it safe, joyful to learn from, share errors. Situational awareness of where pitfalls lurk
More accountability, 'P4P' payment incentives and punishments tied to performance metrics	Clinician engagement in improvement based on trust, collaboration Metric modesty as many best practices yet to be defined/proven
More rules, requirements; target outliers for better compliance	Standardisation with flexibility; learning from deviations
More time with patients	Better time spent with patients: offloading distractions, more efficient history collection/organisation, longitudinal continuity, and where needed, additional time to talk/think during, before, after visits
Reflex changes in response to errors	Avoiding tampering; understanding/diagnosing difference between special versus common cause variation

design (the authors did feed back to the providers any outstanding failed follow-up patients, but the 2-year lag in the study period precluded more 'real time' feedback). In addition to the logistical challenges of such massive chart reviews are challenges that application of the electronic screen would face related to the question of timing—when should the screens/triggers be run? If run too early (eg, 2 weeks after the time of documentation of a +FOBT), firing reminders or instituting interventions risks needlessly harassing physicians and patients just embarking on a work-up; if too late (eg, after 6 or 12 months) the protocol misses an opportunity for more timely diagnosis of a growing colon cancer.

READY, AIM, IMPROVE: NEW PARADIGMS TO TRIGGER BETTER DIAGNOSIS

Thus, we see from Murphy *et al* that we have widespread diagnostic errors and delays, at least for these two diagnoses, confirming a growing body of literature demonstrating suboptimal diagnosis.[7 8] We also see a glimpse of ways new tools might aid in overcoming limitations of care systems and human memory and performance reliability.[9–11] Over the past decade a small but growing cadre of researchers, educators, and practitioners, have begun to grapple with the millennium-old problem of medical diagnosis in new ways, informed by a larger error-prevention movement outside and within medicine.[12–16] Much of this work has coalesced in a series of international conferences on Diagnostic Error in Medicine (now in their 6th year). These conferences (selected proceedings from which appeared in a recent supplement to *BMJ Quality & Safety*) have planted the seeds for new approaches to diagnostic error.

What will it take to jump-start new thinking, approaches and practices to help fulfil the promise of better diagnosis? Historically, efforts to improve diagnosis have been directed toward improving diagnostic technology—more and better lab and imaging tests. A parallel, potentially offsetting and challenging recent trend is changes in traditional physician–patient relationships. Patients and physicians were previously more likely to intimately know each other over time, and (according to physicians and patients at least) physicians had more time to talk to, examine, and think about their patients.[17] Without delving into a host of important related controversies (such as, whether and how technologies are being overused, and ways to ensure they are used more cost-effectively, whether medical homes will make things better or worse), there are ways we need to begin rethinking how we approach diagnosis and diagnosis errors.

From our work with the earlier AHRQ funded Diagnostic Error Evaluation and Research (DEER)[12] Project, and more recent opportunities to study malpractice and diagnostic errors with Harvard's malpractice insurer,[4 18] I offer a series of possibly provocative and certainly oversimplified bullets to contrast where we have come from and where we need to go (table 1).

While these artificially dichotomised contrasting paradigms each warrant much more evidence and discussion, they can stimulate discussion about what and how we are thinking, teaching and practicing related to medical diagnosis. We welcome the 'needles' Murphy *et al* have uncovered, and hope some of the provocative jabs offered here can serve to puncture our complacency and force us to rethink our collective approach to better diagnosis.

Competing interests None.

Provenance and peer review Commissioned; internally peer reviewed.

REFERENCES

1 Singh H, Thomas EJ, Khan MM, *et al*. Identifying diagnostic errors in primary care using an electronic screening algorithm. *Arch Intern Med* 2007;167:302.
2 Singh H, Giardina TD, Forjuoh SN, *et al*. Electronic health record-based surveillance of diagnostic errors in primary care. *BMJ Qual Saf* 2012;21:93–100.
3 Murphy DR, Laxmisan A, Reis BA, *et al*. Electronic health record-based triggers to detect potential delays in cancer diagnosis. *BMJ Qual Saf* 2014;23:8–16
4 Schiff GD. Finding and fixing diagnosis errors: can triggers help? *BMJ Qual Saf* 2012;21:89–92.
5 Berwick DM. Continuous improvement as an ideal in health care. *N Engl J Med* 1989;320:53.
6 Schiff GD, Kim S, Krosnjar N, *et al*. Missed hypothyroidism diagnosis uncovered by linking laboratory and pharmacy data. *Arch Intern Med* 2005;165:574.
7 Graber ML. The incidence of diagnostic error in medicine. *BMJ Qual Saf* 2013;22:ii21–ii27.
8 Schiff GD, Hasan O, Kim S, *et al*. Diagnostic error in medicine: analysis of 583 physician-reported errors. *Arch Intern Med* 2009;169:1881.
9 Schiff GD, Bates DW. Can electronic clinical documentation help prevent diagnostic errors? *N Engl J Med* 2010;362:1066–9.
10 El-Kareh R, Hasan O, Schiff GD. Use of health information technology to reduce diagnostic errors. *BMJ Qual Saf* 2013;22:ii40–ii51.
11 Singh H, Graber ML, Kissam SM, *et al*. System-related interventions to reduce diagnostic errors: a narrative review. *BMJ Qual Saf* 2012;21:160–70.
12 Schiff GD, Kim S, Abrams R, *et al*. *Diagnosing diagnosis errors: lessons from a multi-institutional collaborative project*: DTIC Document. 2005.
13 Ely JW, Kaldjian LC, D'Alessandro DM. Diagnostic errors in primary care: lessons learned. *J Am Board Fam Med* 2012;25:87–97.
14 Berner ES. Diagnostic error in medicine: introduction. *Adv Health Sci Educ* 2009;14:1–5.
15 Elstein AS. Thinking about diagnostic thinking: a 30-year perspective. *Adv Health Sci Educ* 2009;14:7–18.
16 Wachter RM. Why diagnostic errors don't get any respect—and what can be done about them. *Health Aff* 2010;29:1605–10.
17 Dugdale DC, Epstein R, Pantilat SZ. Time and the patient–physician relationship. *J Gen Intern Med* 1999;14(S1):34–40.
18 Schiff GD, Leape LL. Commentary: how can we make diagnosis safer? *Acad Med* 2012;87:135–8.