A Scoping Review of Diagnostic Error: Executive Summary

Background

SCPOR undertook a scoping review on the subject of diagnostic error in health care, with the objective of exploring what research could be undertaken in Saskatchewan to address this issue.

Results

Diagnostic error is difficult to define, because diagnosis is both a process and the outcome of that process. Furthermore, it is difficult to judge when the process is complete – a physician may have a working diagnosis that seems to be the most likely reason for a given set of symptoms, but often only time will tell if that diagnosis was correct or not. Was his or her original diagnosis wrong, or merely in process? The Institute of Medicine has proposed the following definition: diagnostic error is the failure to (a) establish an accurate and timely explanation of the patient’s health problem(s) or (b) communicate that explanation to the patient. This frames diagnostic error from a patient’s perspective, reflecting the fact that patients bear the greatest burden and risk of harm from misdiagnosis.

While a number of data sources and methods have been used to understand the incidence and nature of diagnostic error, all have methodological issues. Taking all the data together, what is known is that both common and rare diseases are prone to missed or delayed diagnosis, and older patients and patients with complex and multiple disease states are more likely to be misdiagnosed. Many researchers feel almost everyone will experience a misdiagnosis in their lifetime, many resulting in serious injury. However, due to the difficulties of measuring diagnostic error, robust numbers simply aren’t available.

Researchers have tried to establish diagnostic error incidence through various means, including autopsy, medical chart reviews, physician and patient surveys and reporting, standardized patients, and using “triggers” in electronic medical chart systems. In the United States, malpractice claims databases have also been used. All of these approaches have methodological issues, and most are retrospective, which does not help the particular patient who risks potential injury from the error. Recently, detection methods have become more sophisticated. More advanced technology using medical charts is looking promising as a way to detect possible errors before harm is done.

About the Literature Review

Given the topic of diagnostic error is multifaceted, our strategy was to do a scoping review on the topic in order to map the literature, identify key concepts and theories, and look at sources of evidence. The aim was to provide an overview of the type, extent and quantity of research available on diagnostic error.

A search of the literature was done using MEDLINE, EMBASE, and Cinahl, and Google Scholar was used to identify relevant grey literature. Government reports and books relevant to the review objectives were also included. The literature was selected based on relevance to the topic of diagnostic error in terms of its prevalence, its origins, and research into strategies to reduce diagnostic error. Editorial and policy articles were scanned to identify controversies and trends in policy. Fifty-one sources were included in this scoping review.

Contact

To obtain more information about the Office of the Associate Vice-President Research – Health (UofS)/Vice President Research and Innovation (SHR), please click here.
Diagnostic error can have many origins, including lack of knowledge or cognitive error on the part of the physician, lack of communication between the health care team members or between the physician and the patient, and incorrect, missing or misinterpreted lab results. Time pressures and distractions amplify the chance of error. There is some indication that diagnostic error often has more than one cause—a mix of system and physician (cognitive) error.

Research into preventing diagnostic error is still in its nascent stage, and the few studies done in this area tend to be non experimental or quasi experimental. There is interest in improving medical education to improve diagnostic skill, including education on teamwork, cognitive forcing strategies (techniques to reduce bias and incorrect decision processes), and better feedback; more evaluation is needed for these interventions. Promoting patient participation in diagnosis, including strategies such as open medical records, have also been used and have yet to be robustly evaluated. Other promising areas include using diagnostic support tools (software that helps create a differential diagnosis) and including patients in research on diagnostic error, a viewpoint missing in most research on diagnostic errors.
Background

Saskatchewan was the first jurisdiction in Canada to formalize critical incident reporting through legislation that came into force on September 15, 2004. The province has an established network of professionals in place within health regions and the Saskatchewan Cancer Agency who identify events where a patient is harmed (or where there is a potential for harm) and report de-identified information to the Provincial Quality of Care Coordinators (PQCCs) in the Ministry of Health. Consistently since 2004, diagnostic error has been one of the most frequently reported adverse events, along with falls and medication errors (Ministry of Health Annual Report 2014-15, p. 60). This inspired the SCPOR office to investigate diagnostic error and how research into diagnostic error might align with the mandate of patient oriented research.

According to the Society to Improve Diagnosis in Medicine, it is estimated that one in every ten diagnoses is wrong, and one in every thousand ambulatory diagnostic encounters result in harm ("Society to Improve Diagnosis in Medicine," 2015). This statement is based on a wide variety of research studies that suggest that breakdowns in the diagnostic process result in a staggering toll of harm and patient deaths. Research on diagnostic error includes autopsy studies, case reviews, surveys of patient and physicians, voluntary reporting systems, using standardised patients, second reviews, diagnostic testing audits and closed claims reviews. Although these different approaches provide important information and unique insights regarding diagnostic errors, each has limitations and none is well suited to establishing the incidence of diagnostic error in actual practice, or the aggregate rate of error and harm (Mark L. Graber, 2013).

Objective

The objective of this scoping review is to summarize the recent medical literature on diagnostic error. Results of the review will be used to inform SCPOR stakeholders who are interested in pursuing research projects focusing on diagnostic error. This review will summarize the evolution of thinking regarding diagnostic error and current diagnostic error incidence research. In addition, the review will examine interventions to improve diagnostic error and the evidence supporting these interventions.

What is Diagnostic Error

A review of the literature reveals the definition of diagnostic error varies considerably. Schiff defined diagnostic error as “any mistake or failure in the diagnostic process leading to misdiagnosis, a missed diagnosis, or a delayed diagnosis.” (Schiff et al., 2009) There is disagreement about the precise meanings of delayed diagnosis, missed diagnosis, and misdiagnosis, however. Young (2015) argues that concept of “delayed diagnosis” needs to be curtailed, or even abandoned in primary care – he points out that to deliver lower cost medical care family physicians have to be comfortable with uncertainty and be comfortable applying probabilities to individual patient situations. In fact, he argues, with the exception of blatant negligence, the language of delayed or missed diagnosis has no place in the world of family medicine. Groszkruger (2014) agrees, writing: “Attempting to define diagnostic error by referring to mistake or failure seems to merely restate the central question. What acts or omissions constitute errors?” Like Young, he points out the preliminary, or working diagnosis is
often wrong, due to incomplete information. Nor does there seem to be a clear line between ‘error’ and ‘a difficult diagnosis’.

Others have expanded Schiff’s definition to "a diagnosis that was unintentionally delayed (sufficient information was available earlier), wrong (another diagnosis was made before the correct one), or missed (no diagnosis was ever made), as judged from the eventual appreciation of more definitive information." (M. L. Graber, Franklin, & Gordon, 2005). Singh and colleagues have defined diagnostic error as "missed opportunities to make a timely or correct diagnosis based on available evidence." (Singh, Meyer, & Thomas, 2014).

Part of the reason for the varying definition lies in the dualistic nature of the term “diagnosis”, which refers to both the process and the result of that process (Balogh, Miller, & Ball, 2015). Another area of disagreement is that while some restrict diagnostic error to failure in the diagnostic process that could have been prevented, others also include unavoidable errors. (M. Graber, 2005; Newman-Toker & Pronovost, 2009; Schiff et al., 2009). The lack of definition prompted the Institute of Medicine (IOM) to formulate a new definition, which is: diagnostic error is the failure to (a) establish an accurate and timely explanation of the patient’s health problem(s) or (b) communicate that explanation to the patient (Balogh et al., 2015 p. 3-4).

The IOM’s definition frames diagnostic error from the patient’s perspective, because the patient bears the risk of harm from diagnostic error. If there is failure in either aspect – either a failure to correctly diagnose, or a failure to communicate that diagnosis – then a diagnostic error has occurred. While this adds a level of complexity to evaluation of the diagnostic process, it also supports the idea that each aspect may be evaluated separately, and then the two aspects aggregated.

Previous to the IOM’s report, there has been little in the literature about communication with the patient, and what little there is focused on patients not preparing appropriately for diagnostic tests due to inadequate communication, or failing to ensure the patient understood what was required of them. The IOM looked at the wider picture, arguing that an accurate and timely explanation of the health care problem is meaningless if the information does not reach the patient (or the patient’s family or proxy, in the case of those unable to communicate) so that the patient and health care professional can act on the information (Balogh et al., 2015 p.3-4).

Diagnostic Error: A Brief History

Discussion in the medical literature on diagnostic error is not new. A 1915 editorial in the Journal of the American Medical Association expressed concern that post mortem findings in America revealed a high percentage of incorrect clinical diagnoses. The editorial classified errors into two categories, social and clinical. Social errors included ‘bad deportment’ and ‘lack of tact’, while clinical errors stemmed from ignorance and faulty judgement, failure to think anatomically, and ‘failure to think at all’. While inherent difficulties in certain cases were acknowledged, for the most part the blame for incorrect diagnoses was placed squarely on the shoulders of the physician ("Why physicians err in diagnosis," 2015).
Fast forward to the year 2000, when the IOM published *To Err is Human*, a report that is generally credited with starting the modern patient safety movement (Kohn, Corrigan, & Donaldson, 2000). The report mentioned medication errors 70 times, while diagnostic error appeared only twice. This is surprising considering the IOM’s much quoted estimate of 44,000-98,000 yearly deaths from medical errors was taken from a Harvard Medical Practice study, which found that diagnostic errors constituted 17 percent of adverse events, many times more than medication errors (Mark L. Graber, Wachter, & Cassel, 2012; Robert M. Wachter, 2010). The reason for the lack of focus on diagnostic error in the IOM’s report lies in their approach, which focused on systems rather than individuals (R. M. Wachter, 2010). The first decade of the patient safety movement targeted the relatively low hanging fruit of improvements to systems of care, focusing on medication errors and errors of execution (for instance, preventing central line infections). By comparison, diagnostic error was perceived as difficult to address, because the reasons for their occurrence is often complex and multifaceted. (Berenson, Upadhyay, & Kaye, 2014; Croskerry, 2009; M. Graber, 2005; Schiff et al., 2005; L. Zwaan, Thijs, Wagner, van der Wal, & Timmermans, 2009)

Some have argued the systems approach to patient safety, while showing some real gains in reduced morbidity and mortality, resulted in a decade of inattention to diagnostic error. Resources and attention that could have been used to improve diagnosis instead were devoted to already over-represented management and system issues (Mark L. Graber, Wachter, et al., 2012). In *Improving Diagnosis in Health Care*, the IOM points out that few health care organizations have processes in place to identify diagnostic errors and near misses. (Balogh et al., 2015 p. S-8) The consequences of this may be more far-reaching than has been generally appreciated. When diagnostic accuracy is not taken into account in quality initiatives, a bad situation can be made worse. For example, the Centers for Medicare and Medicaid Services used to have a ‘door to antibiotic’ policy to expedite administration of antibiotics to patients with pneumonia. On evaluation, however, it was found many of the patients did not in fact have pneumonia, and the policy had to be changed (Robert M. Wachter, 2010).

Given these concerns, the IOM established a committee to look at the modern literature on diagnostic error and make recommendations. *Improving Diagnosis in Health Care* was published by the IOM in 2015. The report concluded that change is urgently needed to address the issue of diagnostic error and that errors continue to harm an unacceptable number of patients. The hope is that policy and priorities will change so that diagnosis will now become a major focus of health care research or practice (Balogh et al., 2015 p. 9-1).

Despite the title, the reader is hard-pressed to find any recommendations in *Improving Diagnosis in Health Care* that would have an immediate effect on the issue of diagnostic error (Goozner, 2015). Although the report has an ambitious policy agenda, the failure of health care to recognize and record misdiagnosis means it will be difficult to measure improvement. Some of the strategies that could be implemented immediately are:

1. Hospitals should find ways to bring together all the members of the medical team and use their collective wisdom in a team environment to create the best possible differential diagnoses.
2. Implementation of electronic diagnosis software such as ISABEL and DXplain, which have shown promise but have not gotten much traction outside of academic settings.
The other recommendations are matters of policy and putting resources toward research and education, and will take some time to bear fruit. It is clear much more research on measuring incidence, and testing of ameliorating measures, is necessary. This will require a shifting of research dollars and of the attention of health care leadership to the problem of diagnostic error.

The report emphasized a patient centered approach, and made the following recommendations regarding partnering with patients and families in the diagnostic process:

**IOM Committee on Diagnostic Error Recommendations:**
Health care professionals and organizations should partner with patients and their families as diagnostic team members and facilitate patient and family engagement in the diagnostic process, aligned with their needs, values, and preferences. To accomplish this, they should:

- Provide patients with opportunities to learn about the diagnostic process.
- Create environments in which patients and their families are comfortable engaging in the diagnostic process and sharing feedback and concerns about diagnostic errors and near misses.
- Ensure patient access to electronic health records, including clinical notes and diagnostic testing results, to facilitate patient engagement in the diagnostic process and patient review of health records for accuracy.
- Identify opportunities to include patients and their families in efforts to improve the diagnostic process by learning from diagnostic errors and near misses. (Balogh et al., 2015 p. S-7)

**What is the cause of diagnostic error?**
Traditional forms of medical education subscribed to the idea that diagnostic errors arise from knowledge deficits, and the common reaction to many types of diagnostic error is to insist on more education and training. Some feel the percentage of diagnostic errors directly attributable to knowledge deficits may be very small. This being the case, remedial education can improve only a fraction of diagnostic errors, based on current research (Groszkruger, 2014).

There is inherent uncertainty in diagnosis. The question a clinician must answer is whether sufficient information has been gathered to make a diagnosis. Treatment is often begun on the strength of a working diagnosis, weighing the harms and benefits of treatment versus waiting while more information is gathered (Balogh et al., 2015 p. 2-5). Experience, knowledge, time constraints, a modern emphasis on diagnostic testing over accurate and thorough history, language and literacy barriers, patients with mental health or cognitive issues, and stressful and distracting work places all contribute to diagnostic error (Balogh et al., 2015; Mark L. Graber, Kissam, et al., 2012). In fact, one internal medicine study found six factors on average combined to create a diagnostic error (M. L. Graber et al., 2005).

Errors in diagnosis are most often due to cognitive error or systems error (including poor communication) but more often a combination of system and cognitive error (M. L. Graber et al., 2005). Cognitive errors take place during the physician’s thinking process, and include latching on prematurely to a diagnosis and abandoning the search for evidence to the contrary. Systems errors occur between the inter-related pieces in the healthcare system. Examples include
Diagnostic Error: A Scoping Review

practitioners ‘dropping the ball’ in the referral-consultation process, or lost or unreported test results. The system’s error tends to be the catalyst for adverse events, because systems errors tend to replicate and amplify ("Facts Page, Society to Improve Diagnosis in Medicine,").

Diagnostic Error: What is the data?
The data on diagnostic error is sparse, and regardless of the method used, there are numerous challenges in reliably identifying or analyzing such errors. (Schiff et al., 2009) The answer to which diagnoses are the most likely to be misdiagnosed, how often they actually are missed or misdiagnosed, and the harm resulting from such errors continues to elude an accurate or even generalizable answer. The next section looks at diagnostic errors detected using various methods, and the advantages and limitations of each.

What are the most common diagnostic errors?
An analysis of 583 physician-reported errors at Brigham and Women’s Hospital showed the most common missed or delayed diagnoses were pulmonary embolism (4.5%), drug reactions or overdoses (4.5%), lung cancer (3.9%) colorectal cancer (3.3%), acute coronary syndrome (3.1%), breast cancer (3.1%), and stroke (2.6%). Errors occurred more frequently in the testing phase (failure to order, report, and follow up on results), followed by errors in the clinicians’ assessments, history taking, physical exam, and referral or consultation errors or delays. (Groszkruger, 2014)

In a large Veteran’s Affairs hospital and integrated private health care system in the USA, researchers reviewed medical charts that were identified through ‘triggers’ – that is, if a patient made an unexpected return visit after an initial primary care visit, the chart was analyzed for evidence of a diagnostic error (Singh et al., 2013). In 190 charts, 68 unique diagnoses were missed: 6.7% pneumonia, 5.7% decompensated heart failure, 5.3% acute renal failure, 5.3% cancer, and 4.8% urinary tract infection. In just over half the cases, the error was discovered because the symptoms, being untreated, did not resolve. In 35% of the cases there was an evolution of the original symptoms, and in 23% the patient returned because of new symptoms. Only 20% of the errors were caught during planned follow-up.

In Singh’s study, a wide variety of common conditions were misdiagnosed, as opposed to rare diseases or a few selected, difficult to diagnose illnesses. The study also found that 14% of the errors led to immediate or inevitable death, 19% to serious permanent injury, 16% to serious harm, 48% to considerable harm, and 10% to minor harm. The authors attributed most of the diagnostic errors made on practitioner time constraints.

Misdiagnosis of cancer used to be considered the most common misdiagnosis, because of its over-representation in American malpractice suits. According to the Society to Improve Diagnosis, today the top diagnostic errors resulting in US claims in adult primary care is myocardial infarction. Stroke is associated with diagnostic error 9% of the time ("Facts Page, Society to Improve Diagnosis in Medicine,").
How Prevalent is Diagnostic Error?

A number of data sources and methods have been used to understand the incidence and nature of diagnostic error, including postmortem examinations (autopsy), medical record reviews, malpractice claims, health insurance claims, diagnostic testing studies, and patient and clinician surveys, among others (Berner & Graber, 2008; Mark L. Graber, 2013; Singh & Sittig, 2015). Even the more reliable measures have methodological issues. Considerably less information is available on the harm caused by diagnostic error (Balogh et al., 2015 p. 9-1; Groszkruger, 2014). Existing tools to capture adverse events are insensitive to diagnostic error (Mark L. Graber et al., 2014). Part of the challenge is the variety of settings in which errors occur – hospitals, emergency departments, long term care, outpatient settings in their various forms – combined with the complexity of the diagnostic process itself.

One study combined the data from chart reviews in three outpatient facilities and found a combined error rate of 5.08%, an estimate that the authors consider conservative. Based on the author’s previous work, it was estimated about half of those errors could have resulted in harm. (Singh et al., 2014) Arthur Elstein, a cognitive psychologist whose interest was in how doctors process information, studied clinical decision making his entire career. He concluded diagnosis is wrong 10-15% of the time (Elstein, 1995). Graber (2013) agrees this is likely a good estimate. Autopsy studies identified major diagnostic errors in 10-20% of cases (Balogh et al., 2015 p. 3-17).

In the United States, diagnostic errors are the leading cause of paid medical malpractice claims, and those errors are almost twice as likely to have resulted in death compared to other claims. However, the data from these claims are biased toward more serious errors. In addition, a 2007 report from the USA suggested 6% of physicians are responsible for 58% of medical malpractice payments, thus further skewing data from this source (“The Great Medical Malpractice Hoax,” 2007). In Canada, both culture and legislation have resulted in awards against physicians being much less frequent, on a per capita basis, than in the USA. As a result it would not be useful to extrapolate diagnostic error data from malpractice claims in Canada (“Medical Malpractice Liability: Canada,”) However, one advantage of claims analysis is that it does incorporate the patient viewpoint, a viewpoint notably missing from most other data sources.

There are several reasons for the variation in prevalence seen in studies on diagnostic error. First, the prevalence will depend on where the error occurs – for instance, errors are less common in radiology than in clinical medicine. As well, errors are more common in older, more complex patients. The study methods also influence the estimates. Data sources that over-represent severe or lethal outcomes, such as autopsies or malpractice suits, could overstate the prevalence, while chart reviews may miss diagnostic errors. Certain diagnoses are more easily missed or confused. Lastly, there are differences in how diagnostic error is defined, as previously discussed, which can lead to differences in what gets counted as an error.
Diagnostic Error: A Scoping Review

Diagnostic Error: the Data Sources

Autopsies (Post mortem examinations): Post mortem examinations can provide extensive information on diagnostic error because the examination can be very thorough. However, a very limited number of these exams are performed annually, little information is available for characterizing the relationship between those who receive post-mortem exams and the potential number of eligible cases, and those who undergo autopsy are more likely to have died due to a diagnostic error, which creates selection bias (Balogh et al., 2015 p. 3-16).

Medical Records: Researchers have attempted to identify diagnostic errors by reviewing medical records – either by random selection, or by using triggers to identify records with a higher likelihood of error. For example, a trigger could be an unplanned hospitalization following a primary care visit. This method is very labour intensive, as it requires clinicians to review the files. Information may be missing from the file. Nor does this method ensure consensus, as even expert reviewers sometimes don’t agree on whether a diagnostic error has been made. Although there is no universally agreed definition of missed diagnostic opportunities, retrospective chart reviews are still considered the gold standard for detecting such errors. (Cheraghi-Sohi; Sudeh; Singh, 2015)

Clinician Surveys: Physicians have been surveyed on their experience of diagnostic errors, either their own or their colleagues’. This method is subject to non response bias, reporting bias, and the lack of ability to confirm the validity of rates based on self report. If the surveys are anonymous, there may be a greater likelihood of doctors reporting their own errors. (IOM, 3-27) In one study, 47% of clinicians encountered diagnostic errors at least monthly, while 64% said that 10% of these errors resulted in patient harm. It was interesting that in this particular survey, 1% of physicians reported never having encountered a diagnostic error. (MacDonald, 2011)

Clinician surveys that have been reported in the literature are seldom prospective, but one study is planned in Spain that looks at dyspnoea in family physician practice (Minué et al., 2014). Doctors will fill in questionnaires about their diagnostic process in each case of dyspnoea, the first asking about initial diagnostic impression, the three most likely diagnoses, and the diagnosis reached after the medical history and physical exam. The second questionnaire will record the confirmed diagnosis. The results will have a volunteer selection bias and a Hawthorne effect, and with no objective criteria for optimal dyspnoea care, they will have no gold standard for comparison. In addition, some errors may not be detected because the participating doctor deemed the diagnostic process finished when it was not.

Patient surveys: The particular value of patient surveys is that they allow researchers to understand failures at the front end of the diagnostic process and in the process of communicating with patients. The health care consumer is uniquely positioned to report on this component of diagnostic performance. A survey of this type in Massachusetts indicated 23% of people surveyed had experienced a medical error, and half of these errors were diagnostic (Balogh et al., 2015 p. 3-28). One issue with many of these surveys, as with many of the physician surveys, is that there is no way to corroborate the error. Weissman and his colleagues compared patient survey-reported adverse events with the patient’s medical record (Weissman et al., 2008). Twenty-three percent of the patients who responded reported at least one adverse
event, compared to 11 percent identified in the medical record. Although this study looked at adverse events, it did not look at diagnostic error, which is somewhat surprising – it focused on adverse events such as hospital acquired infection, adverse drug event, and deep vein thrombosis. However, Weissman et al suggested hospitals could use surveys to detect adverse events to identify patients for a targeted record review, in place of medical record reviews or as a complementary search strategy. It is possible this could be a useful strategy with diagnostic errors as well.

**Second reviews:** In the visual sub-specialties such as radiology and pathology, film or specimens can be reviewed by a second specialist to check for error. In studies where only abnormal specimens are used, it is interesting to note that two reviewers increase the risk of diagnostic error by 10-50%, and in a small number of cases specialist even disagreed with their own prior interpretation. In ‘real world’ situations, the majority of examinations are normal. Under these conditions, a critical abnormality is detected by a second reviewer in 2–5% of cases (Mark L. Graber, 2013).

**Diagnostic testing audits:** Diagnostic error is rarely the result of the clinical test itself. Lab results are misleading in 2-4% of cases, roughly the same as the percentages found in radiology. This does not take into account errors in interpretation, misdirection of results or failure to communicate the results to the patient.

**Standardized patients:** In these studies, real or simulated patients with classical presentations of common diseases such as rheumatoid arthritis or asthma are sent anonymously into real practice settings. (Mark L. Graber, 2013) One study involved 23 rheumatologists who were visited by a patient with known psoriatic arthritis. The diagnosis was missed or wrong in 9 visits (39%) (Laura Zwaan, Schiff, & Singh, 2013). Overall, standardized patient research show a 13-15% error rate and have substantial face validity, because the studies are both prospective and in a real clinical setting. Besides the resource intensiveness of this type of study, the approach has the limitation of presenting a smaller subset of conditions than would be seen in usual practice. Case complexity increases the rate of diagnostic error. (Kostopoulou, Delaney, & Munro, 2008) However, this approach has the advantage of allowing some insight into what contributes or detracts from diagnostic accuracy.

**Case Review:** This method is used to study specific symptoms or diseases, for example, ER patients who experienced a stroke, or pediatric asthma. Delayed or wrong diagnoses of 10-50% have been identified in studies of coronary artery disease, HIV associated complications, tuberculosis and many malignancies. The limitation with this approach, as with any chart review, is that it relies only on information in the medical record. Not only is information often missing, but the chart does not record what the doctor was thinking at the time of diagnosis (Mark L. Graber, 2013)

**Studies about context:** Various studies have been done looking at working conditions and how they affect diagnostic accuracy. One study looked at telephone call interruptions in the radiology reading room. Findings indicated that call interruptions may negatively impact diagnostic accuracy in residents (Balint et al., 2014).
The IOM committee (p 3-30), after reviewing the evidence, felt that of the various methods to assess diagnostic error frequency, post-mortem exams, medical record reviews and malpractice claims data showed the most promise. However, none of these alone will give a valid estimate, and multiple approaches will be necessary to give a more thorough understanding of the prevalence of the problem. The committee’s addition of successfully communicating the diagnosis to the patient adds a level of complexity – for the diagnosis to be considered correct, it must meet all the elements of a correct diagnosis (timeliness, accuracy and communication) (Balogh et al., 2015 pp. 3-13 - 3-14).

Methods for Estimating the Incidence of Diagnostic Error (adapted from Balogh, Improving Diagnostic Error, Table 3-1)

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Key Features of the Data Source</th>
<th>Method(s) for Selecting Cases for Review (Denominator)</th>
<th>Method for Determining Error Occurred (Numerator)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Postmorten examination (Autopsy)</td>
<td>Deaths only Limited number of reviews Selection bias (typically focused on unexpected deaths) Limited workforce</td>
<td>Consecutive series with criteria Convenience samples Pre-specified criteria Requests (from clinicians)</td>
<td>Comparison to another data source (medical record, interview, location/circumstance of death) Cause of death determination Effects or indication of disease</td>
</tr>
<tr>
<td>Medical records</td>
<td>Rely on documentation (what was recorded, such as clinical history and interview, physical exam, and diagnostic testing)</td>
<td>Pre-specified criteria (e.g., trigger tool) Random sample</td>
<td>Implicit review/expert assessment Explicit criteria</td>
</tr>
<tr>
<td>Diagnostic testing</td>
<td>Source data available for review Applies only to diagnoses for which diagnostic testing data are a key factor Focus on interpretation</td>
<td>Random sample Pre-specified criteria</td>
<td>Expert assessment compared to original</td>
</tr>
<tr>
<td>Medical imaging</td>
<td>Source data available for review Applies only to diagnoses for which medical imaging data are a key factor Focus on interpretation</td>
<td>Random sample Pre-specified criteria</td>
<td>Expert assessment compared to original</td>
</tr>
<tr>
<td>Surveys of clinicians</td>
<td>Subject to non-response bias</td>
<td>Sample receiving survey</td>
<td>Descriptive statistics on self-report</td>
</tr>
</tbody>
</table>
Interventions to reduce diagnostic error

A review of the literature shows there are numerous opinions about, and suggestions for, interventions that can reduce diagnostic error. However, there is a paucity of studies providing evidence for these interventions (Balogh et al., 2015 p. 3-42; Robert M. Wachter, 2010). Interventions can be grouped into two categories: system related and cognitive. Two recent reviews looked at system related interventions and cognitive interventions on the reduction of diagnostic error.

System related interventions include encouraging team work, changing procedures, introducing better technology, or using clinical decision support software. Singh and colleagues (Singh et al., 2012) concluded that few empirical studies have tested system related interventions of any sort to reduce diagnostic error in the last decade. A systematic review done by Singh in 2012 found only six tested interventions, and those studies were non experimental or quasi experimental and samples were small. Advancing the research on this issue will require not only more robust study design but also rigorous definitions of diagnostic processes and outcomes.

Cognitive interventions include educational interventions and metacognition approaches (training clinicians to think about how they think in order to avoid errors and improve decision making). Graber and colleagues (Mark L. Graber, Kissam, et al., 2012) found empirical testing on these approaches were sparse, and often involved observing trainees in an artificial setting rather than actual practice. They concluded that methodological refinements are needed in cognitive intervention research before any approaches can be endorsed. A review of the literature also shows there has been a failure to engage with scientists who are expert in human performance, perception, cognition and decision making (Henriksen & Brady, 2013).

In the emerging literature cognitive issues and biases are treated separately from system failure. Both camps have their advocates, even though both these approaches yield less than impressive track records. It can be argued that supporting one camp or the other does disservice to the complexity of the diagnostic process (Henriksen & Brady, 2013). The next section is a more detailed look at various types of interventions and their success or failure in research and in practice.

<table>
<thead>
<tr>
<th>Survey Type</th>
<th>Potential Issues</th>
<th>Sample Characteristics</th>
<th>Data Collection Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surveys of Patients</td>
<td>Subject to non-response bias</td>
<td>Sample receiving survey</td>
<td>Descriptive statistics on self-report</td>
</tr>
<tr>
<td>Standardized Patients</td>
<td>Traditionally have only been used for specific conditions; very resource intensive</td>
<td>Random sample</td>
<td>Pre-determined diagnosis; clinician either arrives at the diagnosis or does not.</td>
</tr>
</tbody>
</table>
Teamwork

Numerous studies have shown teamwork and emphasizing a culture of collaboration can produce better, more cost effective results (Balogh et al., 2015 p. 4-3). These findings are consistent with other sectors, such as aviation and nuclear power industries. However, compared to teamwork in other areas of health care, teamwork in the process of medical diagnosis has received little attention. One study found that medical students working in teams made fewer diagnostic errors (Hautz, Kämmer, Schauber, Spies, & Gaissmaier, 2015), while collaboration between clinicians and clinical pathology teams resulted in better diagnostic test selection (Seegmiller et al., 2013).

In a systematic review, Graber (2012) reported that groups can make better decisions if its members are allowed to function independently. But another well designed controlled study found performance did not improve by using a team. Given the evidence, the IOM recommends reframing diagnosis as a team activity, although they note that this will take time and likely encounter some resistance.

Mechanisms to Measure and Promote Diagnostic Skill in Practicing Physicians

Reporting System with Root Cause Analysis

Two programs in the United States have been developed for the specific purpose of measuring and promoting diagnostic skill (R. M. Wachter, 2010). The first is a three pronged program at a Maine Medical Centre. It consists of an educational campaign, a physician based reporting system, and a re-designed root cause analysis. Although successful in its implementation, the system is resource intensive and its long term sustainability is in question. The actual effect it had on diagnostic reliability remains unclear, as well. However, the initiative reflects a trend that has gained some traction in the US to encourage and facilitate reporting from physicians (Laura Zwaan et al., 2013). This may require physician leadership to act as a champion to encourage participation from colleagues.

Triggers

The second program was developed by the Kaiser Permanente Health Organization. This initiative employed a multidisciplinary team to look at the organization’s database of adverse events (for example, failure to follow up on abnormal test results). The system was designed to catch errors before harm occurred, but didn’t prevent it from happening in the first place. For some limited number of errors – failure to follow up on rectal bleeding and failure to rule out colon cancer in patients with suspected iron deficiency anemia – a semi-automated algorithm was developed. How errors that were discovered were reported back the physician, and whether this reporting improved their diagnostic skill, is unclear.

Triggers may be an efficient way to identify problematic charts, but this method is still retrospective. Switching algorithms to be more proactive brings up issues of logistics and timing. If the algorithm is done too early, it may result in the needless harassment of doctors and patients. If it is done too late, the opportunity may be missed for a more timely diagnosis (Schiff, 2014). The use of triggers is also tied to electronic health records that are shared between
primary care and hospitals and are structured in such a way that access to information is easy to filter and organize (Newman-Toker & Pronovost, 2009; Schiff, 2014). Some fear too much data may have the unintended effect of causing data overload and hence more diagnostic error (Mark L. Graber, Kissam, et al., 2012).

Checklists

Checklists have been successful in hazardous industries such as aviation, and have found their way into healthcare. They have been used to advantage in discrete, observable tasks, such as central line insertions and surgical procedures. Some are suggesting checklists could also reduce diagnostic error. This approach isn’t without its critics. Much of diagnostic process requires perceiving, thinking and interpreting. A checklist may actually discourage those activities (Henriksen & Brady, 2013).

Ely and colleagues (2011) developed three types of checklists:

1. A general checklist to prompt physicians to optimize their cognitive approach
2. A differential dx checklist to help physicians avoid the most common cause of diagnostic error – failure to consider the correct diagnosis
3. A checklist of common pitfalls and reminders to improve evaluation of selected diseases.

The authors contended that theoretically, checklists could reduce biases and failed heuristics that lead to diagnostic errors (Ely et al., 2011) The lists might seem familiar and even insultingly obvious, but pilots are habituated to their own obvious lists and apparently do not feel insulted by reminders to, for instance, release the elevator locks.

However, none of these checklists have been subject to study, and Ely et al cautioned that checklists for diagnosis do not correspond tightly to actionable procedures in aviation. They also cautioned that the checklists they developed were not developed using rigorous or reproducible methods. In fact, they ventured a theory that if there is any advantage at all in using them, it may be less tied to checklists than to a diagnostic time out – a brief pause to reflect on reasoning and affective state. Their article serves to promote the need to study and tests checklists as a means to reduce diagnostic error.

One research team borrowed from aviation and applied an evidence based checklist with the effect of reducing central line associated bloodstream infections (Winters, Aswani, & Pronovost, 2011). In both Michigan and Rhode Island ICUs a 70% reduction in bloodstream infections was achieved. Winters cautioned, however, that while the program was commonly described as implementing a checklist, in fact it was much more complicated than merely ticking off tasks. The program required extensive behavioural and cultural change, which was a difficult and complicated process. In addition, robust measurement of infections was necessary to obtain good data. The checklist was only part of the process.

Decision Support Tools
Clinical decision support systems (CDSS) are computer systems designed to assist clinicians with clinical decision making tasks. Their purpose is to link what is known about the patient with a large database of health knowledge to influence the choices clinicians make about the patient. Of the newer web based decision support tools, ISABEL has been the most extensively evaluated (Mark L. Graber & Mathew, 2008). ISABEL is a web-based clinical decision support system designed to suggest the correct diagnosis in complex medical cases involving adults. Compared to previous generations of this type of computerized decision support tool, ISABEL is quick to use, and suggested the correct diagnosis in 48 of 50 cases with key findings entry. In contrast, Google searches suggest the correct diagnosis in only 58% of cases. However, the physicians entering the key findings already knew the correct diagnosis, which were taken from case studies published in the New England Journal of Medicine. ISABEL has yet to be tested in non-academic clinical settings.

More recently, Bond compared four CDSS and found them somewhat assistive for diagnosis and education, with a number of caveats, including simple availability, and the ability of the program to integrate with electronic health records (Bond et al., 2012). The authors comment that further studies are needed to test whether these tools support safety in the diagnostic process, and whether their tendency to generate long lists of diagnoses makes them unusable for practitioners and confusing for students.

A 2014 systematic review of computerized decision support tools that focused only on randomized clinical trials in actual practice did not find any difference in patient mortality between care that used these tools and care that did not (Moja et al., 2014). However, there was some weak evidence for improvement in morbidity. None of the studies reviewed looked at diagnostic accuracy, only patient outcome, which reflects the comparative difficulty of ascertaining all-cause morbidity as opposed to diagnostic accuracy in particular.

Open Medical Records

One way to encourage patient engagement is to make health information accessible, including the patient’s medical record. Sunnybrook Health Sciences Centre in Toronto has been providing an online personal health record management service called MyChart to patients since 2006. MyChart provides patients with online access to test results, some medical history appointment management tools and diary for personal health information. Patients can use MyChart to track personal health information such as family history and medication history. Although extensive evaluations of this system have not been published, one small qualitative study indicated physicians felt open medical records may improve physician-patient communication (Yau, Williams, & Brown, 2011).

In the United States the OpenNotes initiative made patient charts available to 5 million Americans. In an analysis of patients invited to read their notes over the course of a year, 70-80% did so and reported feeling better prepared and having a better understanding of their care plan (Balogh et al., 2015 p. 4-25). Doctors reported few if any disruptions to their practice. Initiatives like OpenNotes have the potential to reduce diagnostic error by enabling patients to catch errors in the notes, encouraging them to speak up, and preventing diagnostic delay by
helping patients remember recommendations for tests and procedures (Balogh et al., 2015 p. 4-25). There are some concerns that allowing patients access to their test results without a doctor there to interpret the information may be inappropriate in some cases, although this could be overcome by introducing time delays.

Developing Partnerships with Patients and Families.

Patients and families are an important and largely neglected resource for improving outcomes (Singh et al., 2012). The literature reviewed reflected both the recognition that patients and their families should be more actively included in the diagnostic process, while also acknowledging the difficulties of doing so. The health system can place high health literacy demands on patients (Balogh et al., 2015 p. 4-23). Patient engagement is well documented in chronic disease management and decision making, but their participation in diagnostic error prevention has not been explored.

The definition of patient participation is not self-evident. Patient participation, patient collaboration, patient involvement, partnership, empowerment and patient centred care are all terms used interchangeably in the literature (Willeboordse, Hugtenburg, Schellevis, & Elders, 2014). In Improving Diagnosis in Health Care, the IOM committee recommended that health care professionals and organizations partner with patients and their families as diagnostic team members and facilitate patient and family engagement in the diagnostic process, aligned with their needs, values, and preferences. Health care professionals need to embrace patients and their families as partners in the diagnostic process (Balogh et al., 2015 p. 4-17). Recent efforts to increase patient involvement include a patient education initiative from the IOM. The four page patient education brochure available on the IOM website encourages patients to question their diagnosis, tell their story clearly, and make sure they get their test results. How successful these efforts have been is unclear, and an unintended consequence of these initiatives is that patients may feel as if they are at fault if there is an adverse event.

One approach for identifying diagnostic errors is to ask patients recently seen in Emergency, or discharged from hospital, if they have been misdiagnosed. Studies have shown patients are able, willing and motivated to participate in error reporting systems (Mark L. Graber et al., 2014) Patients and families may have important information not otherwise available to professionals doing a root cause analysis (Singh et al., 2012). This is a logical and practical approach to explore. The chief limitation is the verification needed from health care providers or organizations.

The IOM has several recommendations for involving patients in the diagnostic process (Balogh et al., 2015 p. 8-5, Box 8-1):

Patient and Family Engagement and the Diagnostic Process

• Effective strategies for partnering with patients in the diagnostic process; approaches for reaching diverse population groups, including those who are diverse in language, culture, and individual values, preferences, and needs.
• Development of patient-focused educational resources and shared decision-making tools/strategies in the diagnostic process.
- Patient-centered priorities in reducing diagnostic errors.
- Identification of multiple perspectives to better understand and mitigate diagnostic error (including the patient, family, primary care clinicians, specialists, other health care professionals, organizational leaders, risk management perspectives, and others).
- The impact of patient variables on the diagnostic process and outcomes.
- Disparities in timely and accurate diagnosis and on the populations at highest risk, including those with health literacy limitations, socioeconomic disadvantages, limited English proficiency, and racial/ethnic minority populations.

**Health Care Professional Education**

The literature reviewed talked extensively about improving educational approaches for medical students in order to improve diagnostic abilities. Clinical reasoning, learning to work in teams, and communication were all included as areas that could be improved in medical pedagogy. (IOM s-6).

**Cognitive Forcing Strategies and Metacognition**

Metacognition means thinking about thinking. It describes an individual's ability to stand apart from his or her own thinking, to observe it, and to recognize opportunities for using strategies to improve the logic of the thinking process (Croskerry, 2003).

The more common cognitive errors are referred to as biases, and some appear to be very powerful and universal, affecting all walks of life (Croskerry, 2003) Cognitive scientists have described cognitive biases in a wide range of experimental demonstrations, which has led to predictions of the conditions under which they are likely to occur. Their incidence increases under conditions of uncertainty, or when thinking is hurried, thus pressuring the thinker to use heuristics (shortcuts or abbreviated thinking strategies). Many cognitive errors are made only when the right latent conditions are present, such as in an environment of extreme stress.

A prerequisite to minimizing cognitive error is to develop a general working knowledge of cognitive error theory. Often, clinicians have little insight or awareness of their own cognitive processes and biases that might affect their thinking. Cognitive forcing strategies are specific debiasing techniques that introduce self-monitoring of decision making. Their intent is to prevent clinicians from pursuing a pattern-recognition path that can lead to error (Croskerry, 2003).

Traditionally, the specific situations in which errors are likely to be made have been referred to as “pitfalls.” Awareness of pitfalls, a form of expert, insider knowledge, naturally develops over time in any sphere of activity where specific problems re-occur. All disciplines of medicine have their specific pitfalls, as well as methods for avoiding them. A classic example is a patient complaining of joint pain. There is a tendency for inexperienced clinicians to focus their attention on the joint in question. More experienced clinicians know that pain in a joint might be referred from the joint above or below the joint that is painful. Therefore, the rule is to always
examine the joint above and the joint below. Similarly, in the reading of radiographs in the ED, a pitfall is to call off the search once a salient injury has been found. This has led to the maxim “the most commonly missed fracture in the ED is the second.” The cognitive forcing strategy in this case is that when a fracture or significant soft tissue abnormality is found, the search should be continued for other findings. These pitfalls refer to cognitive errors, and the cardinal rules are the cognitive forcing strategies (Croskerry, 2003).

Given all this, are cognitive errors in medicine necessarily inevitable or might there be opportunities and strategies for avoiding them? The fact that clinical acumen and expertise come with experience would suggest there may be. In many fields, such expertise typically takes about 10 years to develop. Clinicians concerned with medical education have therefore been vocal in support of using cognitive forcing strategies to improve diagnostic accuracy.

However logical the approach may be, the evidence has not supported it. Sherbino et al suggests metacognitive approaches (attempting to improve decision making processes) are ineffective (Sherbino, Dore, Siu, & Norman, 2011). Another study suggested taking time for reflection might improve diagnostic accuracy in complex cases, but the evidence was preliminary at best. As the authors wrote, “Research on cognitive diagnostic mistakes remains a quite novel line of investigation.” (van den Berge & Mamede, 2013)

In summary, while there is a lot of enthusiasm for cognitive interventions, there is a paucity of actual interventions and the studies done were on trainees in a lab setting, reducing generalizability to real world settings (Mark L. Graber, Kissam, et al., 2012).

Other Educational Approaches

Simulation has been shown to improve manual and procedural skills, but it has yet to show improvement in clinical reasoning (Mark L. Graber, Kissam, et al., 2012) and it remains to be shown that it can replace actual clinical experience. However, one study showed diagnostic accuracy was improved when simulation was combined with a diagnostic support tool, while another used simulation to successfully introduce the use of cognitive forcing strategies in emergency department residents. Graber also found that intensive, focused feedback has been shown to improve performance in all areas, including diagnostic accuracy. Detailed feedback to trainees on why their diagnosis was correct or incorrect improved later diagnostic accuracy.

However, educators need to ensure diagnostic process is addressed across the career trajectory. The continuing education system needs to focus on educational outcomes and competencies, not just credit hours. One author recommends competency based evaluation (CBE) as this is a better predictor of future performance (Holmboe et al., 2011).

Educational interventions in particular are difficult to evaluate. One major issue is the difficulty of demonstrating that diagnosis can be improved by any approach in real-world settings. Definitions of diagnostic error are not standardised and error designations are typically subjective judgements, often confounded by hindsight bias. Measurement instruments and methods to evaluate cognitive intervention effects are not well developed. Additionally, because diagnostic error reflects the interplay of system-related and patient-dependent factors, the true
effect of a purely cognitive intervention might be difficult to demonstrate. All of these factors pose challenges in the design of future interventions in this area (Mark L. Graber, Kissam, et al., 2012). Singh (2013) has suggested the medical community work closely with cognitive psychologists to translate scientific knowledge about cognitive processes into practice.

**Literature Search Strategy**

Given the topic of diagnostic error, which is multifaceted, our strategy was to do a scoping review on the topic in order to map the literature, identify key concepts and theories, and look at sources of evidence. The aim was to provide an overview of the type, extent and quantity of research available on diagnostic error.

A search of the literature was done using MEDLINE, EMBASE, and Cinahl, and Google Scholar was used to identify relevant grey literature. Government reports and books relevant to the review objectives were also obtained. The literature was selected based on relevance to the topic of diagnostic error in terms of its prevalence, its origins, and research into strategies to reduce diagnostic error. Editorial and policy articles were scanned to identify controversies and trends in policy. Seventy-two articles were found, and 51 were found sufficiently relevant and were retrieved and included in the review, along with two reports from the Institute of Medicine and one report from the Ministry of Health, Saskatchewan. Information from the Society to Improve Diagnosis in Medicine in the United States and from the Library of Congress was also scanned for relevant information.

**Keywords and Terms**

The following key words were used:

- Diagnostic error
- Diagnostic accuracy
- Medical education and diagnosis
- Cognitive error in medicine
- Clinical reasoning

**Summary**

Available research estimates are inadequate to describe the causes, prevalence and resulting injury of diagnostic error, however the medical community is in agreement that there is an urgent need for research both on diagnostic process and errors. The Institute of Medicine’s most recent report, *Improving Diagnosis in Health Care*, also emphasized the need to frame diagnostic accuracy from a patient perspective and include patients not only in discussions of their own diagnosis, but in research. Given that patients are the ones most harmed by diagnostic error, this approach seems not only logical but just.

The challenges of measuring diagnostic errors have been, and will continue to be, an unavoidable obstacle to progress. However, a comprehensive quality management system requires measuring the incidence of diagnostic error in all areas of health care. The focus of research has shifted from the blunt instruments of autopsy reporting and (in the USA) malpractice claims to more
representative and sophisticated data sources. Patient participation, along with more advanced technology, give us reason to be optimistic progress can be made.

**
Author: Barbara Colvin, Health Research Facilitator and Innovation Analyst
Office of the Associate Vice-President Research - Health, University of Saskatchewan
December 14, 2015
References


Cheraghi-Sohi; Sudeh; Singh, H. R., David; Stocks, Jill; Morris, Rebecca; Esmail, Aneez; Campbell, Stephen; de Wet, Carl. (2015). Missed diagnostic opportunities and English general practice: a study to determine their incidence, confounding and contributing factors and potential impact on patients through retrospective review of electronic medical records. *Implementation Science, 10*(1), 1-9p. doi: 10.1186/s13012-015-0296-z


Diagnostic Error: A Scoping Review


