Anatomic Pathology Workload and Error

Stephen S. Raab, MD, and Dana M. Grzybicki, MD, PhD

The Institute of Medicine defined medical error as the failure of a planned action to be completed as intended or the use of a wrong plan to achieve an aim. During the last 2 decades, the pathology patient safety literature mainly has focused on reporting the frequency of error, and this frequency depends on the detection method. Few anatomic pathology studies have focused on the performance of error root cause analysis or on redesigning pathology systems so that fewer errors occur.

Determining the root cause of error is an important step in error prevention through system redesign. A problem in error reduction is that most errors are secondary to multiple causes and difficult to target with a single initiative. In other areas of medicine, excessive workload has been correlated with higher error frequencies, and these investigations have led to changes in how workforce is used. In this issue of the Journal, Renshaw and Gould report that pathologist workload was not a significant driver of pathology error. This editorial examines the findings of Renshaw and Gould in the perspective of the psychology of error and other studies examining workload and medical error.

Psychology of Error

A component of patient safety research involves the study of the psychology of error and the relationship of error to work. Types of error may be differentiated by their connection to action. Errors may occur when using automatic, reflex-like behavior (ie, schematic control mode) or when using conscious thought or processing (ie, attention control mode). Automatic processing is when continued practice leads to a quick choice of standard procedure, and this results in a level of behavioral autonomy. In pathology, cytology screening and grossing small tissues, at times, may be performed in the schematic control mode. When using conscious thought or processing, attention selects information for further use, and working memory stores information that has been attended recently. Making a histologic diagnosis or grossly examining a complicated specimen often is performed in the attention control mode.

Reason subclassified errors as mistakes, slips, and lapses. A mistake occurs because an individual chooses the wrong course of action. A mistake may be the lack of awareness of a disease entity or the misapplication of rules of disease classification. A slip occurs when the correct action is chosen but is executed incorrectly. Pani and Chariker wrote that a lapse occurs when incorrect execution involves failure of memory. Sirota indicated that a lapse is an omission of an automatic action. These error types are not exclusive of each other; a lapse can lead to a slip, that then may result in a mistake.

Sirota provided examples of slips, lapses, and mistakes. A slip is when a slide is marked as read or reviewed when it was not observed at all. A lapse is when key words are inadvertently left off a report. A mistake is when the diagnosis of a rare disease entity is not made because the pathologist did not recall that entity when forming his or her differential diagnosis. All types of workers, ranging from transport personnel to pathologists, make slips, lapses, and mistakes in all subdivisions of an anatomic pathology laboratory. The majority of anatomic pathology errors are detected before sign-out, and the errors that are detected after sign-out usually are not associated with major untoward clinical consequences.

Excessive workload is a cause of error in several situations, such as working for too long a time or being asked to do...
too much in a fixed period. Exhaustion and stress may result in the inability to effectively function in the schematic control mode. If workload is too high, elements of a problem are unattended or pushed out of working memory, and cognition moves forward with missing data. We intuitively recognize that more errors may occur when we are tired. People are encouraged to not drive when overly tired (as an error reduction plan to limit accidents). In addition, much of the American workforce performs shift work (ie, set number of working hours per day), which also helps limit work-related errors and personnel injury.

**Workload and Patient Safety Literature**

Recently published articles, outside the field of pathology, serve as a guide for measuring the relationship between workload and error.

**Working Hours of Hospital Registered Nurses and Patient Safety**

Rogers et al\(^1\) hypothesized that registered nurses (RNs) who worked longer shifts made more errors. In this study, 393 full-time hospital-based RNs completed logbooks on hours worked, overtime, and errors. Logbooks have been used to collect data during field studies of individuals such as airplane pilots, air traffic controllers, space shuttle flight controllers, and emergency room physicians. Rogers et al\(^1\) measured the number of shifts worked (n = 5,317) and classified shifts into less than 8.5 hours, 8.5 to 12 hours, and greater than 12.5 hours. During the study period, the RNs documented 199 errors; 58% of errors involved medication administration, 18% of errors were procedural, and 12% were charting. Rogers et al\(^1\) found that the likelihood of making an error increased as RNs worked longer hours and was 3 times higher when RNs worked shifts lasting 12.5 hours or more (odds ratio, 3.29; \(P = .005\)). Working overtime increased the odds of making at least 1 error, regardless of how long the shift originally was scheduled (odds ratio, 2.06; \(P = .005\)). Working more than 40 hours per week and working more than 50 hours per week significantly increased the risk of making an error. Rogers et al\(^1\) concluded that longer work durations, possibly owing to fatigue and stress, are associated with increased error frequency.

**Working Hours of Interns and Patient Safety**

Landrigan et al\(^4\) hypothesized that interns in intensive care units who worked with an extended schedule (≥24 hours) made more errors than interns who did not work extended shifts. For some interns, an intervention was introduced that eliminated extended work shifts and reduced the number of hours worked per week. Errors were identified by direct, continuous observation. Landrigan et al\(^4\) reported that during a total of 2,203 patient days and 634 admissions, interns made 35.9% more serious errors during the extended schedule (\(P < .001\)). Interns made 5.6 times as many diagnostic errors during the extended schedule (18.6 vs 3.3 per 1,000 patient days; \(P < .001\)). Landrigan et al\(^4\) concluded that reducing the number of extended work shifts and reducing the number of hours worked per week can reduce the number of serious medical errors in intensive care units.

**Study Characteristics**

These 2 studies shared a number of methods in common, which are representative of the majority of workforce studies examining factors associated with error. First, these studies measured the independent variable of workload with time metrics, such as amount of time worked, or shift time, and not as the amount of work produced. Nursing studies have correlated the number of beds covered (ie, business of shift) with error frequency, although this correlation has been performed as a subanalysis of data demonstrating a primary correlation between work-shift time and error. One of the reasons that studies focus on work-shift time is that the work performed by the majority of decision-making medical personnel is highly varied and difficult to quantitate; work-shift time is the simplest metric that is comparable across many individuals. Nurses, physician trainees, and physicians perform many tasks, and patient or caseload is only 1 factor that influences the activities of a work shift. Once the potential correlation between work-shift time and error frequency is determined, researchers may examine additional workload elements, such as case complexity or volume. Measuring these factors is challenging and requires incorporation of time-motion studies that necessitate self-reporting or observation.

Another characteristic of these clinical workload studies is that they measure workload as shift length per day. Intuitively, errors increase in frequency as individuals become tired, as daily work times increase. Measuring error frequency over longer periods (other than a shift or a 24-hour period) eliminates the investigators’ ability to reveal the previously described relationship between errors and work time, if it exists. These studies focused on all types of medical error that occurred in real time and did not use retrospective data review as their error detection method, which generally is acknowledged to involve significant bias. These studies also evaluated fairly large numbers of participants and large data sets.

Nonmedical fields also have focused on the length of work shift as the main metric of workload although work productivity has been evaluated as a variable (ie, air traffic controllers only handle a fixed number of planes at one time).

**Cytotechnology Workload and Error**

Workload limits already exist in anatomic pathology, and the Clinical Laboratory Improvement Amendments of 1988
prescribed a workload limit of 100 slides per day for cytotechnologists. In contrast with the studies just discussed, this limit focused on productivity and not work time. The justification for this practice may have been based on the fact that most cytotechnologists already worked shifts so that this limit would affect cytotechnologist screening speed per day. The limit would prevent laboratories from enacting exceedingly high productivity requirements, limit overtime, and restrict cytotechnologists from screening too many slides by working 2 jobs. As Allen pointed out, the 100-slide limit was not meant as a performance target and represented a maximum number that many believed (at the time) was very high. Studies showed that the majority of cytotechnologists screened between 40 and 80 conventional Papanicolaou tests a day, and the limit was established mainly to prevent cytotechnologists from screening unreasonably high numbers, which was part of the media and public uproar leading to governmental regulation in the first place.

Arguably, the 100-slide limit does not seem strongly data-driven and was based more on intuition. A limitation in the regulation is that new technologies were not factored in, and the advent of liquid-based and other technologies seriously challenges the efficacy of the 100-slide limit. A problem is that the limit prevents a rigorous study evaluating cytotechnologist workload and error frequency. Another point that should be made is that the 100-slide limit applied to cytotechnologists who mainly screened slides during the day. Cytotechnologists who performed other duties would be assumed to screen fewer slides.

Pathologists and Workload

The first question to ask is “Which metric(s) is/are the most valid measure(s) of pathologist workload?” The answer is that it depends, although neither metric would seem to fully characterize pathologist work. Unlike emergency department physicians (in whom error frequency and workload has been studied), most pathologists do not work shifts, and the work of many pathologists is highly varied and unable to be characterized by case volume. For example, measuring surgical pathology caseload would not characterize pathologists who perform clinical pathology duties, sign out Papanicolaou tests, perform autopsies, attend hospital conferences and meetings, examine intradepartmental consultations, teach trainees, or conduct a host of other duties. If one chose to examine surgical pathology workload for these pathologists, the most realistic method would involve measure of these other factors simultaneously with performance of a multivariable regression analysis. Researchers could perform time-motion studies to determine pathologist activities and then measure actual pathologist sign-out times.

However, if pathologists believe that the absolute number of cases examined is an accurate estimation of their work, the measure of case examination rate may be an accurate metric to correlate with error frequency. As mentioned earlier, Renshaw and Gould did not show a statistically significant correlation between workload (in terms of average number of cases evaluated over time) and error frequency. Ironically, the conclusions by Renshaw and Gould may be confounded by type II error, and the authors discuss this possibility for some situations. Type II error is an error that occurs when a sample’s observations fail to demonstrate statistical significance when a true association or difference exists in the larger population. Possible causes of a type II error in the study by Renshaw and Gould include measure of error as workload as case examination rate over an extended time (as opposed to daily work shift), measure of a single error type (ie, secondary review or amended report error, rather than many error types), small size, measure of error in a single practice, failure to measure other sign-out responsibilities (eg, gynecologic cytology), and lack of measure of case complexity.

Renshaw and Gould performed a bivariable analysis of correlation examining the mean number of cases signed out per month vs error frequency in a convenience sample of cases during the entire study period (21 months). It is not surprising that this study showed a lack of significant association, based on what we know from previous workforce studies examining factors associated with errors, the diversity of daily pathology practice, and the specific variables and data analyses used by these authors.

The problem in establishing workload limits for pathologists is that currently we have no evidence that accurately identifies and describes the workload metrics that correlate with errors. For nurses, medical interns, and emergency department physicians, researchers hypothesized that error and workload were intimately linked, and studies were designed to establish the correlation between shift time and error frequency. Intuitively, we believe that pathology workload and error frequency are linked. But demonstrating the association will require a rigorous examination of multiple work-related variables, which is much more challenging than simply looking at the pathologist mean case volume per month and error frequencies in convenience samples of signed-out cases.

From the Department of Pathology, University of Pittsburgh School of Medicine, Pittsburgh, PA.

References