Medical error: Knowing (and herding) the elephants

John M. Gay, DVM, PhD, Dipl ACVPM
Department of Veterinary Clinical Sciences, College of Veterinary Medicine, Washington State University, PO Box 6610, Pullman, WA 99164-6610

Abstract

Medical error is an emerging issue, brought to the public's attention abruptly in 2000 by the National Institute of Medicine's *To Err is Human* report. Because human error is at its root, medical error is a complex and sensitive topic, is likely underreported, often requires process change to reduce, and is difficult to prevent. With the occurrence of at least a minimum level of human error now being regarded as inevitable, using lessons from high reliability organizations functioning in fault-intolerant industries, such as commercial aviation and nuclear power, the reduction of medical error is shifting from a focus on individual failings to a focus on the system in which it is occurring. This paper reviews the definitions of medical error, summarizes its frequency primarily in the diagnostic aspect of human and veterinary medicine, touches briefly upon approaches to reducing it, and provides references, many online, for pursuing the topic further.

Key words: medical error, diagnostic reasoning, observer variation, checklists, cognitive bias

Résumé

L'erreur médicale est un nouvel enjeu porté brusquement à la connaissance du public en 2000 par le rapport intitulé « *To err is human* » de la *National Institute of Medicine*. Parce que l'erreur humaine est à sa source, l'erreur médicale est un sujet complexe et délicat, probablement sous-estimé, qui nécessite souvent des modifications procédurales afin de l'éviter et qui est difficile à prévenir. On croit qu'il y aura toujours inévitablement un peu d'erreur humaine et on peut tirer des leçons d'organisations très fiables qui fonctionnent dans un environnement intolérant à la défaillance comme l'aviation commerciale et l'énergie nucléaire. Dans ce contexte, l'axe prioritaire de la réduction de l'erreur médicale laisse de côté les lacunes individuelles et se concentre sur le système dans lequel elle opère. Cet article se penche sur les définitions de l'erreur médicale, résume sa fréquence surtout au niveau du diagnostic en médecine humaine et vétérinaire, aborde brièvement les approches pour la prévenir et fournit des références dont plusieurs sont en ligne pour poursuivre le sujet plus à fond.

Introduction

Much research on describing, understanding, and reducing medical error and on the associated human cognitive processes is occurring. The National Library of Medicine (NLM) PubMed indexes some 4,000 papers under the MeSH term 'medical errors' annually, approximately 100[a] of which are under the veterinary science subset. In this literature the concepts, vocabulary, and definitions associated with 'medical error' are still fluid, the broader ones covering all aspects of clinical practice with no taxonomy definitively classifying the different types of medical error, because the field is still in its infancy. For example, a recent systematic search yielded 25 definitions for medical error exclusive of those focused on an error subtype, such as medication error or diagnostic error.[13] Some definitions focus on outcomes, requiring harm to have occurred, while others include 'near misses' or require only the potential for harm to have been present. To reduce the considerable stigma, fear, and jeopardy associated with the label 'error,' less evocative terms, such as 'discrepancy' or 'adverse event,' and related concepts, such as 'patient safety,' 'risk management', and 'quality assurance,' are being used in place of the term error. Being a sensitive topic, admission of a medical error event is often difficult in groups, particularly if the fear of punishment is present, and its occurrence is likely underreported, particularly in situations and cultures where its occurrence is viewed as an individual's failing rather than a system problem.[35]

What is Medical Error?

As commonly defined, medical error is essentially a manifestation in a medical environment of human error in 1 of its many forms. Although as humans most all of us are intimately familiar with the concept, human error is not easily or succinctly defined.[17] The NLM defines 'medical error' as errors or mistakes committed by health professionals that result in patient harm and includes errors in diagnosis, in medication administration, in surgical procedure performance, in the use of other therapy types, in equipment use, and in laboratory finding interpretation. This is distinguished from malpractice, regarding medical error as being 'honest mistakes or accidents,' and malpractice as being the consequence of 'negligence, reprehensible ignorance or criminal intent.'[31] Under medical error, NLM includes the MeSH terms 'diagnostic errors,' defined as: "incorrect diagnosis after clinical examination or technical diagnostic procedures," 'medication errors,' defined as "errors in prescribing, dispensing, or administering medication with the result that the patient fails to receive the correct drug or the indicated proper drug dosage", and 'observer variation', defined as:

> The failure by the observer to measure or identify a phenomenon accurately, which results in an error. Sources for this may be due to the observer's missing an abnormality, or to faulty technique resulting in incorrect test measurement, or to misinterpretation of the data. Two varieties are inter-observer variation (the amount observers vary from one another when reporting on the same material) and intraobserver variation (the amount 1 observer varies between observations when reporting more than once on the same material).

For their 2000 report *To Err is Human: Building a safer health system*, which greatly raised public awareness of medical error, the Institute of Medicine (IOM) Committee on Quality of Healthcare in America modified James Reason's definition of medical error to "the failure of a planned action to be completed as intended (i.e., error of execution) or the use of a wrong plan to achieve an aim (i.e., error of planning)."[29] Grober and Bohen defined medical error as "an act of omission or commission in planning or execution that contributes or could contribute to an unintended result."[16] For their 2015 report *Improving Diagnosis in Health Care*, the IOM Committee on Diagnostic Error in Health Care defined diagnostic error as "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."[30] Both IOM reports are freely available in full online.

How Frequent is Medical Error?

The following is a summary from papers primarily addressing diagnostic medical error, many being reviews. These were selected to provide an initial vocabulary, some background frequency estimates, and to serve entry points into the scientific literature for those interested in pursuing the topic further. For easier recall, percentages are rounded to whole numbers.

Recently Makary and Daniel estimated that medical error was the third leading cause of human death in the US.[23] If this ranking is true, then based on CDC National Center for Health Statistics, medical error ranks third to heart disease and cancer but ahead of chronic lower respiratory disease and accidents as the leading causes of mortality.[9] This estimate generated considerable controversy, 1 online responder pointing out that were this the case, half of all hospital deaths would be attributable to medical error,[23] and Shojania and Dixon-Wood alleging that the authors used data sources inappropriately and used improper procedures to arrive at their estimate.[48] Of note is that the death certification process upon which the CDC mortality statistics are based is itself prone to considerable but seldom publicized error.[8] Counting and summarizing anything at a national scale with reasonable accuracy is challenging; breaking the data down into subcategories is even more so.

The medical error that comes immediately to most clinician's minds is diagnostic error. The 2015 IOM Committee on Diagnostic Error in Health Care concluded that 5% of outpatients experience a diagnostic error annually, and that diagnostic error contributes to 10% of patient deaths and to between 6 and 17% of hospital adverse events.[30] One approach to estimating the frequency of diagnostic error is to compare antemortem clinical diagnosis to postmortem autopsy diagnosis, which numerous studies have done. A major caveat of many of these is that because the autopsy subjects were not randomly selected but were instead a deliberately chosen small and declining proportion of mortalities, significant selection bias is likely present. In a recent survey in a human teaching hospital of 1,800 human autopsies with both clinical and pathological diagnoses, of the 300 autopsies randomly selected from those occurring in 2008, 11% had class I diagnostic discrepancies (antemortem detection of error would likely would have prolonged survival or would have cured), 9% class II (major discrepancy present but antemortem error detection would not likely have prolonged survival or have cured), 27% class III (minor discrepancy that would have changed treatment if detected antemortem), 4% class IV (minor discrepancy that would not have changed treatment), and 49% class V (no discrepancy).[53] In a study of 407 clinician-requested hospital autopsies, the overall diagnostic discrepancy was 30% for the 1,467 diagnoses represented, an average of 3.6 diagnoses per case.[19] Treating the clinical exam as a diagnostic test and the autopsy as the gold standard, the overall sensitivity of clinical diagnosis was 0.74 and the positive predictive value was 0.93 in those circumstances.

The evidence from veterinary medicine necropsy-based diagnostic discrepancy studies parallels that from human medicine and, subject to the same caveats, several examples are the following. In a retrospective study of 3 annual cohorts a decade apart (1989, 1999, 2009) of all dogs hospitalized at least 1 night in a veterinary teaching hospital, Dank et al found that of the 7% (1,854/24,997) that died or were euthanized, the proportion necropsied declined from 59% (339/576) to 22% (148/690) in the 20 years, and that the proportion of these in which 'substantial disagreement' between the clinical and the necropsy diagnosis occurred decreased from 40% (135/339) to 15% (22/148).[11] The authors provided no judgement on the consequences of the discrepancies, concluding that "there was a marked improvement in antemortem diagnosis" but that "... despite the advances in diagnostic modalities (e.g. availability of computerized tomography and magnetic resonance imaging), remarkable diagnostic errors still occur." In a retrospective study of 1,370 bovine cases occurring over 6 years that were presented to the veterinary teaching hospital alive but subsequently died, whether naturally or by euthanasia, and had both a clinical and a necropsy-based pathological diagnosis, complete agreement occurred in 21% (286/1370), a minor finding was missed in the clinical exam of 35% (482/1370), a major finding that may have had bearing on the death was missed in 36% (489/1370), and complete discrepancy occurred in 8% (113/1370).[52]

Of particular note are 2 studies comparing the pre-necropsy cause of death attributed by the livestock with necropsy findings. In a year-long study of a commercial dairy in which 94 of the 2,067 enrolled cows died, the producer's perception of the cause of death was wrong in 45% (42).[25] In a necropsy study of 79 cows from 69 high-mortality Danish dairy herds, the producer's perception was wrong in 50%.[50] One might conclude that without necropsying a sufficient proportion of mortalities on an operation, everyone is likely seriously fooling themselves when attributing the causes of that mortality.

Any diagnostic processes requiring subjective observation and interpretation by clinicians functioning as trained observers may be sources of medical error. Examples include applying a scoring system for body condition, using a microscope to score a stained semen smear, using ultrasonography to replace manual reproductive palpation with visualization, as well as applying traditional techniques of auscultation, ballottement, palpation, and percussion during a physical exam. Most all tests and observations are subject to variation; observer variation is the failure of an observer to measure the same material or to identify the same phenomenon identically to another observation by themselves (intraobserver variation - repeatability) or by another observer (interobserver variation - reproducibility).[36] Observer variability of categorical scoring is frequently quantified by a form of Cohen's Kappa statistic. A Kappa value of 1 is perfect agreement, 0 is agreement only due to random chance (coin flip), values less than 0.4 are typically regarded as poor, 0.4 to 0.75 as fair to good and over 0.75 as excellent. Generally Kappas are found to be lower and the standard error of measurement larger than experience-based intuition suggests. Assessing observer variability when measuring continuous variables requires different statistics, such as standard error of measurement.[37]

An example of a scoring system evaluation is that of Morin et al, in which they found that 3 observers with a median of 9 years of body scoring experience had an average Kappa of 0.82 (excellent) for 2 scoring sessions, 1 in early lactation, and 1 at peak milk, but only 0.49 (fair) for the change in body condition between the 2 sessions.[28]

Imaging is an advancing area of diagnostic medicine requiring considerable subjective interpretation, and thus is particularly susceptible to multiple forms of medical error. Alexander, a veterinary radiologist, classifies differences between observations of identical material when neither are shown to be wrong only as observer variation rather than as error. In turn, she classifies "when a mistake has been made" into technical (e.g., motion, positioning, exposure), perception or analysis error.[1] Brady defines interpretive radiological error as "any discrepancy in interpretation that deviates substantially for a consensus of one's peers" and a discrepancy as "reasonable differences of opinions between conscientious practitioners."[4] Bruno et al estimated that the true prevalence of radiologic error ranges from 4% when a high proportion of normal findings are present to 30% when most findings present are abnormal.[5] They estimate that 60% to 80% of these errors are perceptual, which they define as missing an abnormality that was present on the diagnostic image, stating that "all too often, a finding that is readily apparent in retrospect is inexplicably missed." In a study of 656 difficult radiological examinations in which the diagnosis was missed and delayed until a subsequent exam, Kim and Mansfield classified 42% of the errors as being due to underreading, which is the abnormal finding simply being missed, 22% as being due to satisfaction of search, which is failing to continue searching after the first abnormal finding, 9% as being due to faulty reasoning, which is attributing the abnormal finding to the wrong diagnosis, and 7% as being due to location, which is missing the abnormal finding because it was outside the area of interest and attention.[22] Of note is that error due to lack of knowledge was only 3% of all errors, confirming the old aphorism that more is missed for not looking than for not knowing. In a meta-analysis of 58 studies in which the discrepancy between 2 radiologists reading human CT scans was evaluated, Wu et al estimated an overall discrepancy rate of 8% and a major discrepancy rate of 2%, a major discrepancy being defined as 1 that could potentially or did contribute to morbidity or to suboptimal patient care.[54] Of note is that 8 of the 58 studies had total discrepancy rates over 25%. The take-home message is that although improved imaging technology may reduce medical error, it won't be eliminated as long as subjective assessment is required.

In a veterinary study involving 105 small animals in which ultrasound diagnoses were compared to abdominal surgical findings, the findings were discrepant in 16% of the cases.[15] Of these discrepancies, 59% were classified as cognitive errors, 29% as inevitable, and 12% as a combination of perception and cognitive errors.

In a paper summarizing several comparisons of manual transrectal palpation with ultrasonography for ovarian structures, Hanzen et al reported that for detection of midcycle corpus luteum, palpation had a sensitivity of 0.92 and a specificity of 0.65 while ultrasonography had 0.95 and 1.0, respectively.[18] For detecting follicular cysts, palpation had a sensitivity of 0.63 and a specificity of 0.46, while ultrasonography had 0.82 and 0.87, and for detecting luteal cysts, palpation had a sensitivity of 0.46 and a specificity of 0.62, while ultrasonography had 0.87 and 0.82. In this synthesis, the accuracy of visualization via ultrasonography exceeded that of manual palpation. As an example of advancing technology that may eventually become practical for field application in commercial herds, Scully et al found that ultrasonography with a 12 MHz linear array probe was 0.98 sensitive and 0.97 specific in day 21 pregnancies.[47] Because this is a high-risk period for early embryo loss, the authors recommended that pregnancies be confirmed by ultrasonography after day 30 and that this exam include visualization of a fetal heartbeat. For an update on ruminant ultrasonography, the reader is referred to the recent *Veterinary Clinics of North America: Food Animal Practice* issue that included a contribution addressing the diagnostic accuracy of ultrasound when used in respiratory disease, an emerging use in bovine practice.[6]

In an extensive review and discussion of pathology diagnostic testing, Rabb and Grzybicki found that the reported frequency of diagnostic error in oncologic pathology ranged from 1% to 15%, depending on definition and detection method, and that approximately 50% of these errors resulted in additional testing or diagnostic delays.[38] In 40 externally reviewed pathology slide sets comprised of a range of tissues, the median discrepancy rate was 19% and ranged from 1% to 80%. Of these discrepancies, a median of 40% were regarded as major. While no studies have been published on the error rate in veterinary diagnostic pathology, Stromberg, a veterinary pathologist, stated that "good research on errors in human radiology is highly relevant to veterinary pathologists reading biopsies" and speculates that "there is every reason to believe the error rate is far higher than any of us want to admit."[49]

What Reduces Medical Error?

Since the public's attention was raised in 2000, considerable effort and money have been expended developing and testing interventions against medical error. To date, many proposed interventions have been found wanting when empirically evaluated in the clinical environment, particularly for diagnostic error. In an extensive review of clinical reasoning error studies, Norman et al concluded that "... the assumption that most errors are a consequence of cognitive biases and could be reduced by training physicians to recognize biases is not borne out by the evidence"[32] Based on a systematic review, Saposnik et al concluded that although specific cognitive biases may be associated with diagnostic inaccuracies, further studies are needed to determine the prevalence of the most common and to identify the most effective strategies, if any, for their mitigation.[46]

Norman et al also concluded that "... knowledge deficits are a significant contributor to diagnostic error, and strategies to induce some reorganization of knowledge appear to have small but consistent benefits."[32] Self-regulating professions such as human and veterinary medicine and dentistry have long relied on self-assessment, attending compulsory amounts of self-selected continuing education, and achieving specialty certification to maintain clinical competency post-graduation. Of note is a systematic review of physician self-assessment studies by David et al in which they concluded that "... physicians have a limited ability to accurately self-assess" and that "studies found the worst accuracy in self-assessment among physicians who were the least skilled and those who were the most confident."[12] In a review of the relationship between self-assessment and continuing professional development (CPD), Redwood et al concluded that "CPD programs are expected to foster self-assessing and self-directed practitioners, but the common structure is reported to be largely ineffectual in modifying behavior" and proposed that practitioners be provided training in self-assessment skills.[44]

Dual process theory, the leading theory of human cognition,[14] is frequently mentioned in the medical error literature as the reason that many cognitive errors are made. McKenzie, a small animal practitioner and past president of the Evidence-based Veterinary Medicine Association (EBVMA), wrote a *JAVMA* commentary in which he discusses dual process theory, lists the major cognitive biases and their origin and suggests ways that clinicians may begin changing processes to mitigate them.[26] Alexander and Stromberg also provide suggestions for error mitigation.[1,49]

Dr. James Reason, a psychologist who spent his early career working in aviation and his later career working on safety in other industries including health care,[34] developed the widely applied 'Swiss cheese' model of system accidents and is largely responsible for the shift from the person approach to the system approach for accident prevention.[21,40] His goal is to achieve a fair, reporting and learning culture, but not a blame-free one. In other domains, he has found that 90% of the errors are honest and due to a flawed system, while 10% are culpable and blameworthy. His 1995 paper on human factors describes and classifies the various types of human errors, providing a good starting point for those interested in the subject.[39] Several of his other papers providing further guidance for understanding his approach that are available online are listed in the references.[41-43] Oxtoby et al discuss the application of Reason's Swiss cheese model to veterinary medicine.[33]

The use of process checklists, which are critical in aviation, have been shown to reduce process errors in human surgery and anesthesia.[27] In pre-post observational veterinary studies, Hofmeister et al found that the use of a simple anesthesia process checklist was associated with the reduction of small animal anesthesia incidents from 4% to 1%,[20] and Bergstrom et al found that the use of a surgical process checklist was associated with the reduction of small animal surgical complications from 17% to 7%.[2] A caveat is that controlling bias due to the Hawthorne effect is difficult in such studies.

Of particular note is that while general checklists designed to improve clinical diagnostic reasoning have not proven beneficial, specialized problem-specific checklists designed for use after complex diagnoses have been shown to improve the diagnostic accuracy of even seasoned clinicians.[24,32] In a review of post-diagnostic reflection, Mamede and Schmidt concluded that "reflective reasoning positively affected diagnostic performance when conceived as a process of examining the grounds of initial diagnoses generated through intuitive judgement", and that "the benefits of reflection were particularly substantial when physicians are provided with specific reasoning instructions that led them to be confronted with the evidence from the case."[24]

To reduce the introduction of medical error by subjective scoring and testing systems, before being adopted for clinical use they should have been shown to have sufficient repeatability within individuals (low intraobserver variation) and reasonable reproducibility between individuals (low interobserver variation) under typical usage conditions.[7] To reduce the likelihood of adopting an unreliable scoring system or diagnostic test, clinicians may wish to apply the Standards for Reporting Diagnostic Accuracy (STARD), which lists the components of diagnostic accuracy studies that clinicians should expect to see in study reports.[3]

A developing area involving error detection and mitigation is the veterinary clinic audit, which is required of Royal College of Veterinary Surgeons accredited practices. In their systematic review of 21 papers on veterinary clinic audits, Rose et al include a number of references to this literature.[45] Another technique from industry that is being applied to human health care is root cause analysis (RCA), which is a structured technique for analyzing and preventing industrial accidents.[10,51] An investigation is triggered by a sentinel event and the Toyota '5 Whys' iterative technique is used to uncover the root cause against which an intervention can be developed. As yet, no materials applying RCA to veterinary medicine have appeared online or in the indexed scientific literature, but it would likely be a useful technique to apply to problems involving livestock systems.

Conclusions

Medical error is a complex and sensitive topic, is likely more frequent than many clinicians recognize and is often difficult to reduce, particularly if significant behavioral change is required. Of note is the sense of frustration running through the human medical literature about the intractable nature of medical error, how difficult developing effective solutions is despite almost 2 decades of work, and how slowly the ones shown to be efficacious replace those of unknown efficacy or that are not. Following the lead of high-reliability organizations functioning in fault-intolerant industries, the focus of error reduction is shifting from behavioral change to system overhaul. Where the processes of veterinary medicine are similar to those of human medicine, the frequencies of veterinary medical errors are likely similar to those of human medicine. When empirical studies show that particular error mitigation strategies are efficacious in human medicine, these strategies will also likely be efficacious in veterinary medicine if they are feasible. Evidence suggests that both human and veterinary health care have a long way to go to catch up with the best, as represented by the airline and the nuclear industries.

Endnotes

The following trade press books provide general background on human cognition and human error:

Casner S. *Careful: A user's guide to our injury-prone minds*. New York: Riverhead Books, 2017.

Feldman Barrett L. *How emotions are made: The secret life of the brain*. New York: Houghton Mifflin Harcourt Publishing Co., 2017.

Gawande A. *The checklist manifesto: How to get things right*. New York: Henry Holt and Co., 2010.

Kahneman D. *Thinking, fast and slow*. New York: Farrar, Straus and Giroux, 2011.

Markum A. *Smart thinking: Three essential keys to solve problems, innovate and get things done*. New York: Penguin Group (USA) Inc., 2012.

Schulz K. *Being wrong: Adventures in the margin of error*. New York: HarperCollins Publishers, 2010.

Sloman S, P Fernbach. *The knowledge illusion: Why we never think alone*. New York: Riverhead Books, 2017.

[a]PubMed https://www.ncbi.nlm.nih.gov/pubmed search examples: "medical errors"[MeSH Terms] AND (("2015/01/01"[PDAT]: "2015/12/31"[PDAT]) AND veterinary[sb]) free full text reviews: "medical errors"[MeSH Terms] AND (Review[ptyp] AND "loattrfree full text"[sb])

Acknowledgement

The author declares no conflict of interest.

References

(URL's accessed June, 2017)

1. Alexander K. Reducing error in radiographic interpretation. *Can Vet J* 2010;51:533-536. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2857438/.
2. Bergstrom A, Dimopoulou M, Eldh M. Reduction of surgical complications in dogs and cats by the use of a surgical safety checklist. *Vet Surg* 2016;45:571-576.
3. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, Lijmer JG, Moher D, Rennie D, de Vet HCW, Kressel HY, Rifai N, Golub RM, Altman DG, Hooft L, Korevaar DA, Cohen JF. STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. *BMJ* 2015;351. http://www.equator-network.org/reporting-guidelines/stard/. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4623764/.
4. Brady AP. Error and discrepancy in radiology: Inevitable or avoidable? *Insights Imaging* 2017;8:171-182. DOI:10.1007/s13244-016-0534-1. Available at: https://link.springer.com/article/10.1007/s13244-016-0534-1.
5. Bruno MA, Walker EA, Abujudeh HH. Understanding and confronting our mistakes: The epidemiology of error in radiology and strategies for error reduction. *RadioGraphics* 2015;35:1668-1676. DOI: 10.1148/rg.2015150023. Available at: http://pubs.rsna.org/doi/abs/10.1148/rg.2015150023.

6. Buczinski S, ed. Update on ruminant ultrasound. *Vet Clin North Am Food Anim Pract* 2016;32:i. https://doi.org/10.1016/S0749-0720(16)00020-7.
7. Buczinski S, O'Connor AM. Specific challenges in conducting and reporting studies on the diagnostic accuracy of ultrasonography in bovine medicine. *Vet Clin North Am Food Anim Pract* 2016;32:1-18.
8. Cambridge B, Cina S. The accuracy of death certificate completion in a suburban community. *Am J Forensic Med Pathol* 2010;31:232-235. DOI: 10.1097/PAF.0b013e3181e5e0e2
9. CDC National Center for Health Statistics. Health, United States, 2015: With special feature on racial and ethnic health disparities. Hyattsville, MD. 2016. Available at: https://www.cdc.gov/nchs/hus/index.htm.
10. Charles R, Hood B, Derosier JM, Gosbee JW, Li Y, Caird MS, Biermann JS, Hake ME. How to perform a root cause analysis for workup and future prevention of medical errors: a review. *Patient Saf Surg* 2016;10:20. DOI: 10.1186/s13037-016-0107-8. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5031337/.
11. Dank G, Segev G, Moshe D, Kent MS. Follow-up study comparing necropsy rates and discrepancies between clinical and pathological diagnoses at a veterinary teaching hospital: 2009 versus 1989 and 1999. *J Small Anim Pract* 2012;53:679-683.
12. Davis DA, Mazmanian PF, Fordis M, Van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: A systematic review. *JAMA* 2006;296:1094-1102. DOI:10.1001/jama.296.9.1094. Available at: http://jamanetwork.com/journals/jama/fullarticle/203258.
13. Elder NC, Pallerla H, Regan S. What do family physicians consider an error? A comparison of definitions and physician perception. *BMC Fam Pract* 2006;7:73. Available at: https://bmcfampract.biomedcentral.com/articles/10.1186/1471-2296-7-73.
14. Evans J St BT, Stanovich KE. Dual-process theories of higher cognition: Advancing the debate. *Perspect Psychol Sci* 2013;8:223-241. DOI: 10.1177/1745691612460685.
15. Garcia DAA, Froes TR. Errors in abdominal ultrasonography in dogs and cats. *J Small Anim Pract* 2012;53:514-519. DOI: 10.1111/j.1748-5827.2012.01249.x Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1748-5827.2012.01249.x/abstract.
16. Grober ED, Bohnen JMA. Defining medical error. *Can J Surg* 2005;48:39-44.
17. Hansen FD. Human error: A concept analysis. *J Air Transportation* 2006;11:61-77. Available at: NASA Technical Reports Server. https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20070022530.pdf.
18. Hanzen CH, Pieterse M, Scenczi O, Drost M. Relative accuracy of the identification of ovarian structures in the cow by ultrasonography and palpation *per rectum*. *Vet J* 2000;192:161-170.
19. Heriot GS, Pitman AG, Gonzales M, McKelvie P. The four horsemen: Clinicopathological correlation in 407 hospital autopsies. *Intern Med J* 2010;40:626-632.
20. Hofmeister EH, Quant J, Braun C, Shepherd M. Development, implementation and impact of simple patient safety interventions in a university teaching hospital. *Vet Anaesth Analg* 2014;41:243-248.
21. Kim KO. A first step toward understanding patient safety. *Korean J Anesthesiol* 2016;69:429-434. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5047977/.
22. Kim YW, Mansfield LT. Fool me twice: Delayed diagnoses in radiology with emphasis on perpetuated errors. *Am J Roentgenol* 2014;202:465-470. DOI:10.2214/AJR.13.11493. Available at: http://www.ajronline.org/doi/abs/10.2214/AJR.13.11493.
23. Makary M, Daniel M. Medical error - The third leading cause of death in the US. *BMJ* 2016;353:i2139. Available at: http://www.bmj.com/content/353/bmj.i2139.
24. Mamede S, Schmidt HG. Reflection in medical diagnosis: A literature review. *Health Professions Education* 2017;3:15-25. Available at: http://dx.doi.org/10.1016/j.hpe.2017.01.003.
25. McConnel CS, Garry FB, Lombard JE, Kidd JA, Hill AE, Gould DH. A necropsy-based descriptive study of dairy cow deaths on a Colorado dairy. *J Dairy Sci* 2009;92:1954-1962.
26. McKenzie BA. Veterinary clinical decision-making: cognitive biases, external constraints, and strategies for improvement. *JAVMA* 2014;244:217-276. https://doi.org/10.2460/javma.244.3.271 Available at: http://avmajournals.avma.org/doi/pdf/10.2460/javma.244.3.271.
27. Merry AF, Mitchell SJ. Advancing patient safety through the use of cognitive aids. *BMJ Qual Saf* 2016;25:733-735. Available at: http://qualitysafety.bmj.com/content/early/2016/01/04/bmjqs-2015-004984.full.pdf.
28. Morin P-A, Chorfi Y, Dubuc J, Roy J-P, Santschi D, Dufour S. Short communication: An observational study investigating inter-observer agreement for variation over time of body condition score in dairy cows. *J Dairy Sci* 2017;100:3086-3090. DOI: 10.3168/jds.2016-11872.
29. National Academies of Sciences, Engineering, and Medicine. *To Err is Human: Building a safer health system.* Washington, DC: The National Academies Press. 2000. Available at: https://www.nap.edu/read/9728/.
30. National Academies of Sciences, Engineering, and Medicine. *Improving diagnosis in health care.* Washington, DC: The National Academies Press. 2015. Available at https://www.nap.edu/read/21794/.
31. National Library of Medicine National Center for Biotechnology Information. Medical Errors. Available at: https://www.ncbi.nlm.nih.gov/mesh?term=Medical%20Errors.
32. Norman GR, Monteiro SD, Sherbino J, Ilgen JS, Schmidt HG, Mamede S. The causes of errors in clinical reasoning: Cognitive biases, knowledge deficits, and dual process thinking. *Acad Med* 2017;92:23-30. DOI: 10.1097/ACM.0000000000001421 Available at: http://journals.lww.com/academicmedicine/Citation/2017/01000/The_Causes_of_Errors_in_Clinical_Reasoning__.13.aspx.
33. Oxtoby C, Ferguson E, White K, Mossop L. We need to talk about error: causes and types of error in veterinary practice. *Vet Rec* 2005;177:438. DOI: 10.1136/vr.103331 Available at: http://veterinaryrecord.bmj.com/content/177/17/438.
34. Peltomaa K. James Reason: Patient safety, human error, and Swiss cheese. *Q Manage Health Care* 2012;21:59-63.
35. Perkins IU. Error disclosure in pathology and laboratory medicine: A review of the literature. *AMA J Ethics* 2016;18:809-816. Available at: http://journalofethics.ama-assn.org/2016/08/nlit1-1608.html.
36. Porta M, ed. A Dictionary of Epidemiology, 5th ed. Oxford University Press, 2008.
37. Popovic ZB, Thomas JD. Assessing observer variability: A user's guide. *Cardiovasc Diagn Ther* 2017;7:317-324. Available at: http://cdt.amegroups.com/article/view/14271/15027.
38. Raab SS, Gryzybicki DM. Quality in cancer diagnosis. *CA Cancer J Clin* 2010;60:139-165. DOI: 10.3322/caac.20068. Available at: http://onlinelibrary.wiley.com/doi/10.3322/caac.20068/full.
39. Reason J. Understanding adverse events: human factors. *Qual Health Care* 1995;4:80-89. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1055294/.
40. Reason J. Human error: models and management. *BMJ* 2000;320:768-770. Available at: http://www.bmj.com/content/320/7237/768.long.
41. Reason J. Combating omission errors through task analysis and good reminders. *Qual Saf Health Care* 2002;11:40-44. Available at: http://qualitysafety.bmj.com/content/11/1/40.full.pdf.
42. Reason J. Beyond the organisational accident: the need for "error wisdom" on the frontline. *Qual Saf Health Care* 2004;13(Suppl II):ii28–ii33. DOI: 10.1136/qshc.2003.009548. Available at: http://qualitysafety.bmj.com/content/13/suppl_2/ii28.full.pdf.
43. Reason JT, Carthey J, de Leval MR. Diagnosing "vulnerable system syndrome": an essential prerequisite to effective risk management. *Qual Health Care* 2001;10(Suppl II):ii21–ii25. Available at: http://qualitysafety.bmj.com/content/10/suppl_2/ii21.full.pdf.
44. Redwood C, Winning T, Townsend G. The missing link: self-assessment and continuing professional development. *Aust Dent J* 2010;55:15-19. Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1834-7819.2009.01177.x/epdf.
45. Rose N, Toews L, Pang DSJ. A systematic review of clinical audit in companion animal veterinary medicine. *BMC Vet Res* 2016;12:40. DOI: 10.1186/s12917-016-0661-4. Available at: https://bmcvetres.biomedcentral.com/track/pdf/10.1186/s12917-016-0661-4.

46. Saposnik G, Redelmeier D, Ruff CC, Tobler PN. Cognitive biases associated with medical decisions: a systematic review. *BMC Med Inform Decis Mak* 2016;16:138. Available at: https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-016-0377-1.
47. Scully S, Butler ST, Kelly AK, Evans ACO, Lonergan P, Crowe MA,. Early pregnancy diagnosis on days 18 to 21 postinsemination using high-resolution imaging in lactating dairy cows. *J Dairy Sci* 2014;97:3542-3557.
48. Shojania KG, Dixon-Woods M. Estimating deaths due to medical error: The ongoing controversy and why it matters. *BMJ Qual Saf* 2017;26:423-428. DOI:10.1136/bmjqs-2016-006144. http://qualitysafety.bmj.com/content/qhc/26/5/423.full.pdf.
49. Stromberg PC. Cognitive errors in veterinary diagnostic pathology. Pages 5-7 in: Proceedings of the 29th Meeting of the European Society of Veterinary Pathology & the European College of Veterinary Pathologists, Sept. 7-10, 2011, Uppsala, Sweden. Available at: https://www.esvp.eu/images/downloads/pdf/AbstractBook2011.pdf.
50. Thomsen PT, Dahl-Pedersen K, Jensen HE. Necropsy as a means to gain additional information about causes of dairy cow deaths. *J Dairy Sci* 2012;95:5798-5803.
51. Trbovich P, Shojania KG. Root-cause analysis: swatting at mosquitoes versus draining the swamp. *BMJ Qual Saf* 2017;26:350-353. DOI: 10.1136/bmjqs-2016-006229. Available at: http://qualitysafety.bmj.com/content/26/5/350.long.
52. Wäsle K, Pospischil A, Hässig M, Gerspach C, Hilbe M. The postmortem examination in ruminants and its possible benefit to ruminant clinical medicine. *J Comp Path* 2017;156:202-216. doi.org/10.1016/j.jcpa.2017.01.003. Available at: http://www.sciencedirect.com/science/article/pii/S0021997517300129.
53. Wittschieber D, Klauschen F, Kimmritz AC, von Winterfeld M, Kamphues C, Scholman HJ, Erbersdobler A, Pfeiffer H, Denkert C, Dietel M, Weichert W, Budczies J, Stenzinger A. Who is at risk for diagnostic discrepancies? Comparison of pre- and postmortal diagnoses in 1800 patients of 3 medical decades in East and West Berlin. *PLoS One* 2012;7:e37460. DOI: 10.1371/journal.pone.0037460. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3358345/.
54. Wu MZ, McInnes MDF, Macdonald DB, Kielar AZ, Duigenan S. CT in adults: Systematic review and meta-analysis of interpretation discrepancy results. *Radiology* 2014;270:717-735. doi.org/10.1148/radiol.13131114. Available at: http://pubs.rsna.org/doi/full/10.1148/radiol.13131114.