#### Abstracts\*)

# **Diagnostic Error in Medicine**

# 11th International Conference November 4-6, 2018 New Orleans, Louisiana

#### **Planning Committee**

Ruth Ryan, RN, BSN, MSW, CPHRM (Chair) Jeff Brady, MD, MPH Sharon Cusanza, MSN, RN, NEA-BC Robert El-Kareh, MD, MPH, MS Kelly Gleason, RN, PhD Penny Greenberg, RN, MS, CPPS Helen Haskell, MA Janice Kwan, MD, MPH Rebecca Jones, MBA, BSN, RN, CPHRM, CPPS Ashley Meyer, PhD Timothy Mosher, MD Andrew Olson, MD, FACP, FAAP Art Papier, MD Susan Peterson, MD Dana Siegal, RN, CPHRM, CPPS Robert Trowbridge, MD, FACP Laura Zwaan, PhD

#### **Abstract Selection Committee**

Janice Kwan, MD, MPH Ashley Meyer, PhD

<sup>&#</sup>x27;These abstracts have been reproduced directly from the material supplied by the authors, without editorial alteration by the staff of this Journal. Insufficiencies of preparation, grammar, spelling, style, syntax and usage are the authors' responsibility.

# **Oral Plenary Abstract Session**

Tuesday, November 6, 2018 | 10:15 AM - 11:45 AM

#### Impact of Primary Care Medical Home Implementation on Follow-up of Abnormal Diagnostic Test Results for **Gastroenterological Cancer Evaluation**

D. Choi, A. N. Meyer, L. Wei, D. Murphy, H. Singh

Center for Innovations in Quality, Effectiveness and Safety, Michael E. DeBakey Veterans Affairs Medical Center and Baylor College of Medicine, Houston, TX

Background: Delays in cancer diagnosis can result from lack of timely follow-up of positive fecal occult blood tests (FOBT) and iron-deficiency anemia (IDA) for colorectal cancer (CRC) and elevated serum alpha-fetoprotein (AFP) levels for hepatocellular carcinoma (HCC). The VA implemented the patient-centered medical home (PCMH) model in 2010-2011 to provide patient-driven, team-based care with goals of improving healthcare outcomes. We hypothesized that PCMH will improve timely follow up of FOBT, IDA and AFP tests nationally in the VA.

Methods: To identify patients with delayed follow-up after abnormal results, we applied previously validated electronic "trigger" algorithms to VA's national repository of electronic health record data. The trigger included patients with newly abnormal test results, excluding patients for whom follow-up was not required or action had been completed within 60 days of result. Positive predictive values of 57.0%, 55.1% and 82.3% for FOBT, IDA and AFP respectively, were higher than other known measures of diagnostic safety. We applied each trigger to all patients from 130 VA facilities across 18 national VA networks from 2006-2015. We derived yearly counts of trigger-positive patients based on VA facility and VA network to assess annual percent changes beginning in 2006. Negative binomial regression models were applied to assess overall and yearly changes in number of trigger-positive patients while accounting for clustering by VA network, over-dispersion and correlation due to repeated measures. An offset was created using expected number of trigger-positive patients by facility that adjusted for year and number of patients with primary care provider (PCP) visits that year. Final models were adjusted for VA network and PCP visits that year.

Results: After excluding patients not meeting inclusion criteria, 5,887,006 and 37,762,419 patients had tests for FOBT and IDA from 2006-2015, respectively. Of patients who received FOBT tests, 245,776 patients met trigger-positive criteria and of those who received IDA tests, 303,323 met trigger-positive criteria. Similarly, 888,033 patients received AFP tests, with 12,098 patients meeting trigger-positive criteria. The trigger-positive count means for FOBT, IDA and AFP tests increased immediately following PCMH implementation, however the trigger-positive count means in subsequent years fluctuated between and among tests. Variability was also found by VA network, most likely due to facility size and complexity.

Conclusion: Primary care medical home implementation does not appear to improve follow-up of abnormal test results that warrant cancer evaluation. Further contextual evaluation to explore lack of impact from this teamwork-based intervention is warranted.

#### Implementation of a Diagnostics Consultation Program Reduces Costs and Improves Health Outcomes: Case Studies from a Tertiary Academic Medical Center

B. D. Gunsolus

Augusta University Medical Center, Augusta, GA

Statement of problem: A knowledge and communication gap exists between physicians and the clinical laboratory leading to inappropriate utilization of laboratory testing, inappropriate interpretation of laboratory test results, and incorrect patient diagnosis and treatment.

Description of the intervention or program: Doctorate clinical laboratory science (DCLS) is a new advanced laboratory practitioner specializing in providing physician consultation on all aspects of the clinical laboratory. This practitioner is a member of the interdisciplinary healthcare team and provides guidance in appropriate test utilization, appropriate timing of testing, and interpretation of test results within the clinical context of a specific patient. Consultations were provided by the DCLS practitioner in both patient care rounding team and in conjunction with pathologists. The initiation of DCLS practitioner consultations were documented to assess improvements in health outcomes and cost reductions in various healthcare service delivery areas.

Findings to date: The impact of consultations will be demonstrated by the discussion of case examples. Details of four cases referred to the DCLS, an unknown cause of hyperkalemia, a drug-resistant tracheal aspirate culture, a hypoglycemia fasting admission, and a case of possible non-accidental trauma in a newborn, will demonstrate both significant direct cost savings as well as improved patient health outcomes. Data collection over a one-year period of DCLS diagnostic consultation resulted in direct cost savings of \$697,643.12 relating directly from these consultations. This included cancellation of inappropriate tests, cancellation of inappropriate therapy from incorrect test result interpretation, and reduced patient length of stay from decreased time-to-diagnosis and appropriate treatment.

Lessons learned: These findings predict cost savings and patient health outcome improvements are adequate to justify implementing diagnostic consultation services.

#### Patients As Partners for Improved Diagnoses: A 'How to' Guide

K. Litman<sup>1</sup>, M. Hugh<sup>1</sup>, L. J. Fahey<sup>2</sup>, S. Becken<sup>2</sup> <sup>1</sup>Southern California Permanente Medical Group, Pasadena, CA <sup>2</sup>Kaiser Permanente, Southern California, Pasadena, CA

Background: Diagnostic error reduction frequently focuses on improving systems from a practitioner/organizational 'expert' perspective. The Patient and Family Centered Care movement has shown that patients are 'experts' in their own experience and can improve systems to reduce diagnostic errors. However, many diagnostic error reduction researchers are unaware of this great potential resource and how to partner with patients. We will share successful experiences which can be used by researchers and others to bring the patient voice into their diagnostic accuracy work.

Methods: Kaiser Permanente Southern California (KPSC) region and its affiliated Southern California Permanente Medical Group is an integrated healthcare system with 7,000 physicians and 72,000 employees providing healthcare to 4.5 million patients. In July 2017 we began a program to improve diagnostic reliability. Patients were sought for the development of a variety of patient facing tools to reduce the risk of diagnostic errors. The group reached out to leaders of the KPSC Regional Patient Advisory Council, a group of 18 engaged KPSC patients who volunteer to assist our organization in improving care and safety in collaboration with improvement leaders. This partnership led to the successful development of tools to reduce diagnostic error from a patient perspective, including tools to be used before, during, and after a medical appointment to encourage patients to better partner with their providers to improve diagnostic accuracy. The group also assisted in creating a patient video and other materials to empower patients to better partner with providers to increase successful diagnoses.

Results: According to recent surveys Patient Advisory Council are present in the majority of U.S. medical centers. These are an effective tool to bring patient input to work on diagnostic error reduction. Our efforts benefited from over 150 hours of volunteer patient advisor time resulting in a much more patient centered approach that engages patients and families in improving diagnostic reliability. Feedback from patients was an integral part of our work and informed the creation of educational materials for physicians (to allow them to expect and encourage patient questions) and for patients (to know what questions to ask). We have created several useful tools that will be shared with attendees. Conclusion: Our existing Patient Advisory Council greatly informed and improved efforts to reduce diagnostic error by bringing engaged patient perspective to this work, Similar successful partnerships are possible at the majority of U.S. medical centers utilizing their existing Patient Advisory Councils and other types of patient engagement programs.

#### Burden of Serious Misdiagnosis-Related Harms in the United States—Population-Based Estimate Using the "Big Three" (Vascular Events, Infections, and Cancers)

D. E. Newman-Toker<sup>1</sup>, N. Nassery<sup>2</sup>, A. C. Schaffer<sup>3</sup>, C. W. Yu-Moe<sup>4</sup>, A. S. Tehrani<sup>5</sup>, G. D. Clemens<sup>6</sup>, Z. Wang<sup>7</sup>, M. Fanai<sup>8</sup>, D. Siegal<sup>9</sup> <sup>1</sup>Johns Hopkins University School of Medicine, Baltimore, MD

<sup>2</sup>Johns Hopkins University, Baltimore, MD

<sup>3</sup>Brigham and Women's Hospital, Boston, MA

4CRICO, Boston, MA

<sup>5</sup>Massachusetts Eye & Ear Infirmary/Harvard University, Boston, MA

<sup>6</sup>The Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

<sup>7</sup>Johns Hopkins University School of Public Health, Baltimore, MD

<sup>8</sup>Center for Diagnostic Excellence, Armstrong Institute for Patient Safety and Quality, Johns Hopkins University School of Medicine, Baltimore, MD

<sup>9</sup>CRICO Strategies, Boston, MA

**EMBARGOED** 

#### Improvements in Team-Based Diagnostic Performance Following Training in the Identification and Use of 'High Yield' Signs and Symptoms to Diagnose Serial Cue-Based Case Vignettes

F. J. Papa

UNIVERSITY OF NORTH TEXAS, Fort Worth, TX

Purpose/Problem: Research into the cognitive factors contributing to expert/novice performance differences have demonstrated that: 1) experts correctly diagnose a higher percentage of cases than non-experts, despite 2) gathering fewer signs/symptoms (S/S), largely because 3) experts 'know what S/S to gather.' That is, the experts' higher levels of diagnostic accuracy and efficiency are achieved because they know which S/S best differentiate the diseases associated with the problem at hand. Such S/S are called 'high yield' findings. This study sought to determine if teams of year two medical students could be taught to identify and utilize high yield S/S, and thereby outperform a cohort, when diagnosing a panel of Acute Chest Pain case vignettes.

Description of program, assessment, or study: The author utilized Dual Processing Theory (DPT) as a framework for developing an instructional methodology by which students would be trained in how to extract high yield S/S from a matrix of conditional probability estimates representing the strength of associations between nine common/important differentials and fifty S/S useful for diagnosing case presentations of Acute Chest Pain. The treatment cohort was trained via this DPT-based instructional methodology while the control cohort received our institution's traditional approach to the development of team-based diagnostic capabilities. Ten serial cue-based case vignettes were used to assess both cohorts' diagnostic performance along three diagnostic performance measures.

Outcomes: Diagnostic accuracy: The treatment teams correctly diagnosed 18% more cases (M=63.8, SD=12.6, N=13 teams) than control teams (M=53.9, SD=23.3, N=38 teams). Levene's test indicated unequal variance (F=7.64, p=0.008), thius degrees of freedom were adjusted from 49 to 39; t(39) one-tailed, p < 0.031, Cohen's d = 0.53. Efficiency of diagnostic workup: For cases correctly diagnosed, the treatment group collectively gathered 44% fewer S/S per case (M=7.93, SD=2.80, N=83 cases) than the control group (M=11.38, SD=5.47, N=205); t(286), p = 0.001\*, Cohen's d = 0.79. Use of high yield signs/symptoms: For cases correctly diagnosed, the treatment group collectively gathered 15% more high yield S/S per case (M = 63.55, SD = 15.65, N = 83 cases) than the control group (M = 55.01, SD = 14.53, N = 205); t(286), p = .001\*, Cohen's d = 0.57 \* Bonferroni adjusted p values; alpha of p = 0.05 was divided by 2 (number of family related trials), with significance established as p = 0.025.

Discussion: A DPT based instructional activity enabled teams of year two medical students to exhibit three performance outcomes commonly associated with the development of expertise.

Significance of findings: DPT represents a robust framework for designing and launching instructional activities intended to improve diagnostic capabilities.

#### User Control of Electronic Health Record Design and Diagnostic Reasoning

Y. Senathirajah<sup>1</sup>, E. Borycki<sup>2</sup>, A. Kushniruk<sup>2</sup>, K. Cato<sup>3</sup> <sup>1</sup>University of Pittsburgh, Pittsburgh, PA <sup>2</sup>University of Victoria, Victoria, BC, Canada <sup>3</sup>Columbia University School of Nursing, New York, NY

Background: Diagnostic errors and clinical reasoning are known to be affected by the manner of information presentation, including presentation in the electronic health record (EHR). A known cause of cognitive load and potential error is 'display fragmentation' - the common problem that the data in an EHR is typically presented by information type, across many different screens, requiring navigation by the user and usually, integration by the user in his/her working memory. Display fragmentation results in related information not being presented together, inducing cognitive load and possibly making it more difficult for the user to make essential connections. EHR design has also been shown to affect the structure of note writing, with users changing from a full narrative to more disjoint point-form notes reflecting the EHR structure.

Methods: We utilized EHR navigation mapping methods, interviews with six clinicians as they use their usual EHR with eye tracking, and study of 16 users utilizing a novel user-composable interface, to understand the effects of EHR structure on clinical reasoning processes. With the novel composable interface users' ability to juxtapose information elements in order to understand relationships between elements, share created interfaces, and highlight or rearrange elements as their thinking progresses, may affect clinical reasoning (including the possibility of anchoring or premature closure) and have implications for redesign.

Results: Findings to date: Problems with the structure of the EHR, including substantial fragmentation in all 3 major commercial EHRs were identified by respondents and from task recordings. Fragmentation may necessitate the user traversing dozens of pages and over 100 clicks for routine tasks such as admission. Study of the effects of user interaction on their forming connections between data elements and thus clinical reasoning, and study of the effects of user-created interface sharing and error/omission detection patterns will be reported.

Conclusion: EHR design can have an impact on clinical reasoning and errors via display fragmentation, information omission, and lack of cognitive support for pattern recognition as well as workflow aids. Permitting users to compose interfaces appropriate to the task may have advantages for clinical reasoning. Further study of the exact mechanisms occurring during difficult diagnostic tasks is warranted.

#### Learning to diagnose X-rays: Seeing many examples is more effective than receiving extensive feedback on few examples.

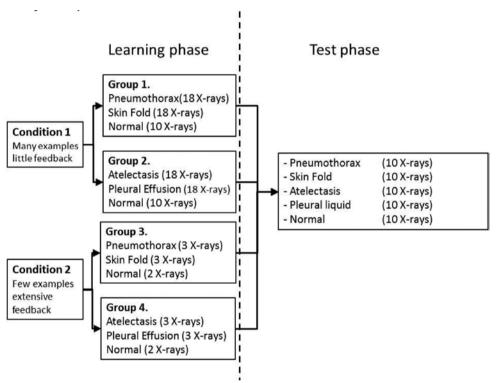
L. Zwaan<sup>1</sup>, B. Hussain<sup>1</sup>, A. Devos<sup>1</sup>, W. W. Van den Broek<sup>1</sup>, S. Mamede<sup>1</sup>, H. Schmidt<sup>2</sup> <sup>1</sup>Erasmus MC, Rotterdam, Netherlands <sup>2</sup>Erasmus University, Rotterdam, Netherlands

Background: Improving the diagnostic skills of medical students is considered as an effective way to reduce diagnostic errors. Many educational strategies have been studied to improve students' ability to distinguish one disease from another. However, the one strategy that distinguishes experienced and inexperienced physicians -seeing many patients- has not been applied in medical education. In the undergraduate curriculum, generally only few typical examples of diseases are discussed. This study examines whether it would be more effective to practice with many cases while receiving little feedback, compared to few cases with extensive feedback.

Methods: We set up a two-phased experimental study. To study the diagnostic process in a realistic setting while allowing to diagnose many cases, we studied X-ray diagnosis. Two hundred twenty-one first-year medical students were randomly assigned to one of the two conditions (i.e. practice with 46 images with little feedback or 8 images with extensive feedback). See Figure 1. After an introduction, participants practiced with diagnosing two similar looking abnormalities (i.e. either atelectasis versus pleural effusion, or pneumothorax versus skin folds) and images without abnormalities. One week later all participants diagnosed 50 images of abnormalities they practiced with and abnormalities they did not practice with. Additionally, the amount of time they needed to diagnose the images was recorded.

Results: The students who practiced with many images with little feedback performed better in the test phase than the participants who had diagnosed few images with extensive feedback (p<0.05). Additionally, the students who practiced with many cases were significantly faster to diagnose the images in the test phase than the participants who practiced with few cases (p < 0.001). There was an overall learning effect regardless of experimental condition, thus students performed better on the abnormalities they practiced with compared to the abnormalities that had not been exposed to in the learning phase (p < 0.001).

Conclusion: Being exposed to many cases and therefore many different presentations of abnormalities leads to better performance on new cases of the same abnormality as well as a faster diagnostic process. Exposing medical students to more cases early in medical school may be an effective way to improve their diagnostic skills. Further research is required to show the effect at different levels of expertise and for a larger variety of diseases in the educational module.



#### Oral Abstracts of Current Research and Educational Innovations

Monday, November 5, 2018 | 3:15 PM - 4:15 PM

Big Data Dashboards to Measure Misdiagnosis-Related Harms - Visual Analytics to Identify Hospital Crossovers within and across Health Systems in Regional Health Information Exchange Datasets

A. Hassoon<sup>1</sup>, M. Austin<sup>2</sup>, M. Fanai<sup>3</sup>, H. Rupani<sup>4</sup>, A. L. Liberman<sup>5</sup>, Z. Wang<sup>6</sup>, D. E. Newman-Toker<sup>3,7</sup>

<sup>1</sup>Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

<sup>2</sup>Johns Hopkins School of Medicine, Baltimore, MD

<sup>3</sup>Center for Diagnostic Excellence, Armstrong Institute for Patient Safety and Quality, Johns Hopkins University School of Medicine,

<sup>4</sup>Johns Hopkins Hospital, Baltimore, MD

<sup>5</sup>Albert Einstein College of Medicine, New York, NY

<sup>6</sup>Johns Hopkins University School of Public Health, Baltimore, MD

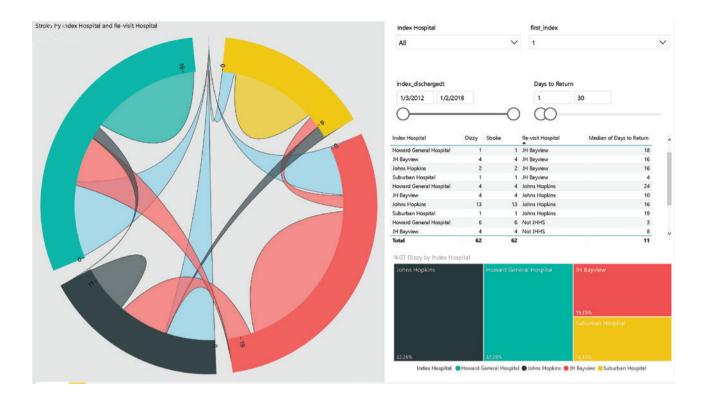
<sup>7</sup>Johns Hopkins University School of Medicine, Baltimore, MD

Background: Diagnostic error is a leading cause of harm in health care. In both malpractice claims and epidemiologic data, missed stroke is the leading cause of harms from missed vascular events. It has been shown that stroke misdiagnosis can be measured operationally using a method known as SPADE (Symptom Disease Pair Analysis of Diagnostic Error). SPADE leverages readily available discharge and hospitalization data to track differences between observed and expected rates of hospitalization (e.g., for stroke) following treat-and-release discharges (e.g., for dizziness/vertigo). Excess hospitalizations reflect misdiagnosis-related harms and could soon be used for publicly reported measures. However, missing data from out-of-network follow-up for outcome events (e.g., stroke hospitalization at a competing hospital/health system) risks underestimation of misdiagnosis-related harms. Relatively little is known about the frequency and nature of such crossovers between hospitals or health systems.

Methods: We sought to measure misdiagnoses and characterize crossovers within and beyond the Johns Hopkins Health System (JHHS) by constructing a data visualization dashboard using data from the CRISP Regional Health Information Exchange in Maryland (2012-2018). We used SPADE methods to identify harms using the dizziness-stroke dyad. ED index visits were those with a "benign" dizziness discharge diagnosis from a Maryland JHHS hospital (n=4). Outcome visits were stroke hospitalizations to a JHHS or non-JHHS hospital within the Maryland-wide CRISP database. We counted only first-incident dizziness discharges and first-incident strokes after discharge. Observed stroke hospitalizations within 30 days were considered misdiagnosis-related harms if above the expected rate (measured using the 30-day average sampled from the base-rate period of 91 to 180 days post discharge). We plotted the data on chord diagram to visualize index and outcome hospital visits using visual analytics tools that allow real-time data exploration using simple interactive click-and-drag or keyboard functions. Results: Over a 7-year period there were 62 return stroke hospitalizations from JHHS (observed, Figure), with a base rate of 29 per 30 days (expected, not shown). In total, short-term revisits were intra-system hospital crossovers in 75.8% (n=47/62) and extra-system crossovers (outside JHHS) in 24.2% (n = 15/62). Base-rate revisits (91-180 days) were intra-system hospital crossovers in 74.1% (n = 43/58) and extra-system crossovers in 25.9 (n = 15/58).

Conclusion: Nearly quarter of our short-term hospitalizations for stroke were hospital crossovers, but many were within the same health system. Nevertheless, high rates of intra-system and smaller extra-system crossovers suggests that cross-institutional data sharing may be necessary to ensure accurate estimates of misdiagnosis-related harms for publicly reported measures.

**DE GRUYTER** 



# Real-World Virtual Patients: Case-Based Simulation Using Actual History and Physical Exam Data to Improve Diagnosis

S. Kotwal<sup>1</sup>, M. Fanai<sup>1</sup>, A. Bery<sup>2</sup>, R. Omron<sup>1</sup>, B. Garibaldi<sup>1</sup>, Z. Wang<sup>3</sup>, D. E. Newman-Toker<sup>1</sup>

<sup>1</sup>Center for Diagnostic Excellence, Armstrong Institute for Patient Safety and Quality, Johns Hopkins University School of Medicine, Baltimore, MD

<sup>2</sup>University of Ottawa, Ottawa, ON, Canada

<sup>3</sup>Johns Hopkins University School of Public Health, Baltimore, MD

**Purpose/Problem:** Improving diagnosis is a pressing public health and medical education concern. Lack of meaningful feedback on clinical diagnostic accuracy hinders development of expertise. Lack of expertise contributes to diagnostic error, and addressing related knowledge/skill gaps improves diagnosis. The National Academy of Medicine recommends simulation training with early exposure to both typical and atypical cases to improve diagnostic performance. The clinical authenticity of standard vignette-based diagnostic simulation is limited, raising concerns about learning transfer to actual clinical practice. Incorporation of real-world patient data into a library of virtual patients could potentially overcome this concern. We will combine available resources to develop an interactive, case-based virtual patient curriculum to improve diagnosis of stroke in patients with acute dizziness and vertigo, who are frequently misdiagnosed.

Description of program, assessment, or study: SCIDD (Simulation-based Curriculum to Improve Dizziness Diagnosis) will use real-world patient data to improve diagnostic expertise among internal medicine residents. Virtual patients are being created in an interactive, screen-based microsimulation software environment (Virtual Interactive Case System, University of Toronto). The software simulates an uncued bedside encounter, leveraging gaming approaches to incentivize a parsimonious, "least-moves" strategy to diagnosis. Structured history and videotaped physical exams from an ongoing, NIH-sponsored clinical trial of acute dizziness diagnosis (AVERT-ClinicalTrials.gov #NCT02483429) are incorporated into the case for maximal case fidelity. Additional links will be built to a mobile phone-based partial task trainer to teach critical psychomotor skills in a second screen-based simulation environment (aVOR app, University of Sydney). This mixed simulation environment closely mirrors real-world experience (focused history-taking and a targeted physical examination). Individualized feedback will enable deliberate practice, and the case library will include a mix of progressively more atypical/challenging diagnostic cases as learners progress. Proposed is a quasi-experimental pre-test/post-test nonequivalent comparison group study. The experimental group (medical interns) will be exposed to traditional lectures plus SCIDD and the control group (graduating senior medical residents) will be exposed to lectures alone. Asymmetric intervention groups allow us to compare SCIDD to the cumulative clinical educational experience for dizziness/vertigo.

**Outcomes:** Our primary outcome will be posttest scores reflecting diagnostic accuracy (experimental vs. control). Secondary outcomes will be change in scores pre- vs. post-test, usability, and user satisfaction.

Discussion: We hypothesize that SCIDD will enhance the knowledge and skills of resident physicians in their approach to patients with dizziness, enabling interns to outperform graduating residents.

Significance of findings: If successful, this condensed educational experience might be applied across a broad range of common symptoms.

#### An Inquiry into the Early Careers of Master Clinicians

V. K. Murthy<sup>1</sup>, G. Dhaliwal<sup>2</sup>, B. O'Brien<sup>3</sup> <sup>1</sup>Johns Hopkins Hospital, Baltimore, MD

<sup>2</sup>San Francisco VA Medical Center and University of California San Francisco, San Francisco, CA

<sup>3</sup>University of Calfornia San Francisco

Purpose/Problem: Residents and fellows often seek to emulate master clinician role models who are skilled diagnosticians, compassionate communicators, and revered teachers. These role models of clinical excellence remind trainees of why they devoted themselves to the study and practice of clinical medicine. However, the methods by which these clinicians reached their elite station of practice are unknown. This gap leaves trainees who aspire to become outstanding clinicians with a destination but no map to guide their early career activities.

**Description of program, assessment, or study:** We performed a retrospective, qualitative interview study of 17 members of the University of California San Francisco Department of Medicine Council of Master Clinicians. Between March 1, 2016 and May 31, 2016, we interviewed participants using a semi-structured interview guide surveying their early career clinical experiences and learning habits. Interviews were audio-recorded and transcribed. We used a general inductive approach to code transcripts and identify consistent themes.

Outcomes: Participants included 12 men and 5 women with an average of 7 years in clinical practice (range 13 to 50 years). Six participants were general internists, and 11 were internal medicine subspecialists. From thematic analysis of interview transcripts, 4 themes of clinical development emerged: 1) consistent learning efforts (reading, teaching, tracking patient outcomes); 2) rigorous skill development (listening and communication, physical exam skills, clinical reasoning skills); 3) cultivating habits of mind (humanism, finding joy in work, humility, rigorous case analysis); 4) clinically rich environments (high clinical volume, practicing outside of their comfort zone, learning from peers, role models). Discussion: Based on strategies described by peer-defined master clinicians, our study suggests 4 categories of purposeful learning activities that trainees can consider adopting in their early careers. Four other descriptive studies of peer-nominated master clinicians at academic medical centers also generated themes of mid- and late-career practice that align closely with our findings: depth of knowledge, self-directed learning, humanism, enthusiasm for patient care and teaching, effective interpersonal skills, and clinical judgment refined through self-analysis. The congruence between the results of these studies and our findings suggests that expert physicians adopt specific learning behaviors during the earliest stages of their careers and maintain those practices.

Significance of findings: Our study represents a first step towards understanding the early-career learning strategies and clinical experiences of master clinicians. Our findings suggest clinical and learning behaviors that may be helpful for graduate medical trainees who set clinical excellence as an important career goal.

Table I Extrapolation of master clinician sub-themes into action steps

Sub-themes	Proposed action			
Reading	Establish a routine of reading about current patient cases for 15 minutes daily			
Teaching	Seek teaching roles			
Tracking outcomes	Establish a patient tracking system			
Listening and	Enroll in training courses on communication skills			
communication	Reflect on communication feedback from patients			
Physical exam skills	Deliberately practice specific exam skills (e.g. estimate jugular venous pressure and			
	compare to ultrasound visualization of IVC; track accuracy over time)			
Clinical reasoning skills	Rehearse clinical reasoning by reading case reports (pause at fixed intervals to reason			
	through clinical data and compare conclusions to those of the care team or clinical			
	discussant)			
Humanism	Develop the habit of asking patients about their illness experiences			
Finding joy in work	Reflect on one rewarding aspect of clinical work each day			
Humility	Adopt a growth mindset by seeking the margin of knowledge or skills			
Rigorous case analysis	Deepen knowledge by routinely asking "why" rather than "what" when clinical			
	decisions are formulated			
High clinical volume	Seek clinical roles that maximize patient volume or fall outside assigned duties			
Practicing outside of their	Seek clinical roles involving unfamiliar patients and environments			
comfort zone				
Learning from peers	Establish a network of colleagues to discuss clinical cases			
Role models	Establish relationships with clinically excellent colleagues and senior faculty to			
	emulate their clinical/teaching skills and discuss diagnostic/management challenges			

#### What Misdiagnosed Patients Can Tell Us about Types and Origins of Diagnostic Errors

U. Shahid, A. N. Meyer, T. Giardina, H. Singh

Center for Innovations in Quality, Effectiveness and Safety, Michael E. DeBakey Veterans Affairs Medical Center and Baylor College of Medicine, Houston, TX

Background: Research priorities related to diagnostic errors have increasingly focused on systems and provider-related factors. Meanwhile, patients are critical observers of their individual care, but their views of diagnostic error remain largely unexplored. Using secondary data, we evaluated perspectives of patients who experienced diagnostic errors to determine types of errors and contributory factors related to patients' interactions with their physicians.

Methods: We conducted a secondary data analysis of data collected in an online survey of patients to assess their outcomes and experiences of using web-based patient-facing 'Isabel Symptom Checker'. The first 385 respondents from a pool of 4000 registered U.S. users of the symptom-checker participated, of which 329 responses were complete and usable. Within this sample, we used thematic content analysis to analyze data from patients who self-reported diagnostic errors and wrote narratives of their experiences (n = 108).

Results: Our sample consisted mostly of females (n = 89, 82.4%) and Caucasians/whites (n = 99, 91.7%), with mean age of 48.8 years (SD = 16.6). Patients reported several types of clinical conditions where they experienced diagnostic errors: rheumatology/immunology (n=35, 32.4%; most commonly Fibromyalgia), infectious (n=11, 10.2%), gastroenterology (n=8, 7.4%), reproductive (n=8, 7.4%), cancer (n=7, 6.5%), endocrine (n=5, 4.6%; most commonly hypothyroidism), cardiovascular (n=3, 2.8%), respiratory (n=3, 2.8%), and mental health-related conditions (n=3, 2.8%). Twelve additional patients incorrectly labelled with mental health issues were labelled as such after they became frustrated, anxious, or depressed after a prolonged wait to obtain a correct diagnosis and multiple referrals. Patients reported several contributory factors including 1) physicians' inability to manage diagnostic uncertainty (n = 33, 30.6%), 2) physicians making multiple unnecessary referrals to others when faced with challenging diagnoses (n=22, 20.4%) rather than thoughtfully connecting the dots between diagnostic data themselves, 3) physicians prioritizing financial gains versus patient benefit (n=19, 17.6%), 4) physicians unfairly labelling patients (e.g., drug or attention seekers, "drama-queens", having symptoms 'all in [their] head[s]', or not 'look[ing] sick': n=16, 14.8%), and 5) physicians not taking time to listen to patients (n = 13, 12.0%). Several patients reported harm, including long-term health consequences from errors, such as disability or life-threatening experiences (n = 73, 67.6%).

Conclusion: Patients' perspectives can reveal useful information about the types and origins of diagnostic errors. Indeed, because patientreported data confirms the vast heterogeneity of conditions that can be misdiagnosed, disease-focused approaches might be limited in their impact to address diagnostic errors on a large scale. Conversely, common contributory factors identified in this analysis are ripe for additional exploration and intervention.

# **Oral Abstracts of Practice Improvement Strategies**

Monday, November 5, 2018 | 4:15 PM - 5:15 PM

#### A Novel Tele-Dizzy Consultation Program in the Emergency Department Using Portable Video-Oculography to Improve Peripheral Vestibular and Stroke Diagnosis

D. Gold¹, R. Tourkevich², A. Brune⁴, S. Peterson⁵, B. Maliszewski⁶, J. Bosley⁶, J. Otero-Millan⁶, M. Fanaiȝ, D. Roberts⁶, D. Zee⁷, D. E. Newman-Tokerȝ <sup>1</sup>Johns Hopkins Medicine, Baltimore, MD

<sup>2</sup>Johns Hopkins, Baltimore, MD

<sup>3</sup>Center for Diagnostic Excellence, Armstrong Institute for Patient Safety and Quality, Johns Hopkins University School of Medicine, Baltimore, MD

<sup>4</sup>University of Michigan

<sup>5</sup>Johns Hopkins University, Baltimore, MD

<sup>6</sup>Johns Hopkins

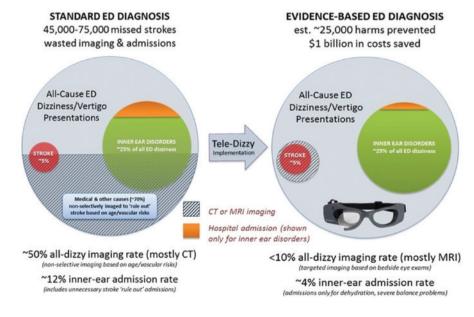
<sup>7</sup>Johns Hopkins School of Medicine, Baltimore, MD

Statement of problem: Missed stroke in the emergency department (ED) is a leading cause of misdiagnosis-related harms. The symptom most tightly associated with missed stroke is dizziness/vertigo. Diagnostic errors in acute dizziness and vertigo are frequent for both vestibular disorders (~80%) and stroke (~35%), due to the complexity of bedside evaluations, which rely heavily on subtle eye movement assessments. Lack of access to subspecialty expertise in real time is a major barrier. We sought to deploy a tele-dizzy consultation service to improve diagnosis of acute dizziness and vertigo in the ED.

**Description of the intervention or program:** This is a preliminary case study of a systems-level quality improvement intervention. The pilot program included (1) defining a new care pathway; (2) securing leadership buy-in; (3) modeling quality and cost benefits; (4) conducting training and implementing technology; and (5) identifying barriers and lessons learned.

**Findings to date:** Over the course of two years we accomplished the following objectives. (1) Defined a care pathway for evaluating ED patients with dizziness/vertigo of suspected neurologic or peripheral vestibular etiology using portable video-oculography (VOG) in the ED, with urgent clinic referrals for unclear cases. (2) Secured buy-in of ED Department Directors at 5 health system hospitals. (3) Modeled health system reduction of 50 missed strokes and ~\$1 million per year saved on unnecessary imaging and admissions. (4) Conducted training for ED technicians and implemented secure data platforms for eve movement recordings to be electronically transferred for review by clinical faculty using 'store-forward' telemedicine approach. (5) Identified need for culture change through local champions; method to sustain 'critical mass' of workforce trained; faculty and fellow staffing to sustain service availability on evenings/weekends; and novel billing mechanisms in hybrid payment system (mixed population-based budgeted and fee-for-service). In the first year at the pilot hospital site, we have assessed 128 teledizzy consultations resulting in 65 specific vestibular diagnoses, roughly twice the base rate of specific diagnoses given. Most consultations are accomplished within one hour. Rates of advanced neuroimaging recommended are less than half the base rate for advanced neuroimaging in the ED. The response to the new clinical service by ED providers has been overwhelmingly positive.

Lessons learned: Bringing subspecialty expertise via tele-dizzy consultation using VOG-based rapid triage to EDs is feasible. The next step is to test the hypotheses of higher quality (greater diagnostic accuracy and fewer misdiagnosis-related harms) and lower costs, as outlined in the enclosed Figure.



#### **Does Limiting Unnecessary Testing Improve the Diagnosis of UTI?**

J. Howard-Anderson<sup>1</sup>, S. Ashraf<sup>2</sup>, E. Overton<sup>2</sup>, J. T. Jacob<sup>3</sup>

<sup>1</sup>Emory University, Atlanta, GA

<sup>2</sup>Emory Healthcare

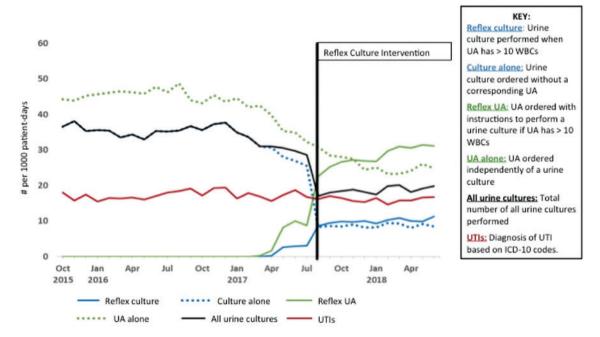
3Emory University

Statement of problem: Diagnosing urinary tract infections (UTIs) or catheter-associated UTIs (CAUTIs) in hospitalized patients is challenging, in part because the diagnosis is supported by, but not exclusively based on, laboratory criteria. Asymptomatic bacteriuria is common in elderly, community-dwelling adults and patients with urinary catheters. The Infectious Diseases Society of America recommends against treating patients without associated signs or symptoms of an infection. However, in hospitalized patients, symptoms are difficult to assess and urine cultures (UCs) are often ordered indiscriminately, resulting in inappropriate UC orders, diagnostic errors, and unnecessary antibiotic use in patients with asymptomatic bacteriuria.

**Description of the intervention or program:** Ordering a urinalysis (UA) with microscopy prior to performing a UC can limit the number of UCs performed and improve diagnostic accuracy. The presence of pyuria may suggest, but not necessarily diagnose, a UTI. However, the absence of pyuria has a high negative predictive value for infection. In 8/2017, three 500-bed hospitals implemented a "reflex" UC intervention, where the default inpatient order for UC was replaced with an order-set for UA with microscopy with "reflex" to UC only if the UA had > 10 white blood cells per high-power field. Clinicians were still able to order a UC without a UA for a limited set of conditions such as

Findings to date: From 10/2015 to 6/2018, we assessed rates of UA and UC ordered at all 3 hospitals (Figure 1). The mean rate of UCs ordered without a UA per 1000 patient-days decreased from 34 before the intervention to 9 after the intervention (p < 0.001). During this period, UC utilization per 1000 patient-days also decreased (34 v. 19, p < 0.001) and the mean proportion of UCs with clinically significant bacterial growth increased (22% v. 31%, p < 0.001). UTIs diagnosed per 1000 patient-days using ICD-10 billing codes decreased (17.2 v 16.0, p = 0.004). Lessons learned: By implementing a process that guides clinicians to order a UA that reflexes to UC only if the UA is abnormal, we decreased the number of unnecessary UCs ordered and improved the diagnostic yield of UCs. We postulate that the observed decrease in UTIs billed reflects more accurate diagnoses, since patients without pyuria would no longer be falsely diagnosed with a UTI.

Figure 1: Rates of urine testing and diagnosis before and after reflex urine culture testing



#### Direct-to-Patient Technology-Enabled Communication Improves Adherence to Diagnostic Laboratory Testing **Regimens for High-Risk Medications**

M. A. Raebel<sup>1</sup>, S. M. Shetterly<sup>2</sup>, E. B. Schroeder<sup>2</sup>, P. Epner<sup>3</sup>, D. H. Smith<sup>4</sup>, J. A. Chorny<sup>2</sup>, B. Bhardwaja<sup>2</sup>, <u>I. M. Lubin</u><sup>5</sup>

<sup>1</sup>Kaiser Permanente Colorado, Aurora, CO

<sup>2</sup>Kaiser Permanente Colorado

<sup>3</sup>Paul Epner LLC, Evanston, IL

<sup>4</sup>Kaiser Permanente Northwest

<sup>5</sup>Centers for Disease Control and Prevention, Atlanta, GA

Statement of problem: Disease-modifying anti-rheumatic drugs (DMARDs) are effective medications that carry risks of organ system adverse events ranging from liver damage to bone marrow toxicity. For patients newly-starting or chronically taking DMARDs such as methotrexate and leflunomide, published guidelines recommend laboratory testing every 4 and 12 weeks, respectively, to diagnose emerging adverse events. If, for example, liver function laboratory testing reveals alanine aminotransferase (ALT) has risen to >2x the upper limit of normal (ULN), the DMARD dosage is reduced or the medication is stopped. At Kaiser Permanente Colorado (KPCO), an integrated healthcare delivery system with about 600,000 members in the Denver-Boulder area, we identified a gap in guideline concordance (timeliness) of patients completing recommended laboratory testing.

Description of the intervention or program: A multi-disciplinary team developed a laboratory-led quality improvement intervention that utilized data in the KPCO medical data warehouse to identify patients taking DMARDs who were due or overdue for laboratory testing. The direct-to-patient outreach we implemented utilized an interactive voice response (IVR) system to text or call patients in English or Spanish to remind them they were due/overdue for testing. We analyzed proportions of patients completing testing and with >2x ULN ALT results during baseline and intervention periods.

Findings to date: The cohort included 3763 patients; 1332 (35.4%) of whom newly-started a DMARD during baseline (9/6/16-9/5/17) or intervention (9/6/17-6/5/18) periods. Fully 2829 (75.2%) of the cohort had DMARD use during both periods. Methotrexate (n=2380 [63.2%]) and leflunomide (n = 698 [18.5%]) were the most commonly used DMARDs. Among patients newly-starting methotrexate or leflunomide during baseline, 51.3% had a gap in timeliness of ALT testing, while during intervention, 42.4% had a gap (P < 0.001). ALT testing among patients newly-starting methotrexate or leflunomide identified 0.5% of results >2x ULN during baseline and 3.9% during intervention. Among patients chronically taking methotrexate or leflunomide during baseline, 55.0% had a timeliness gap in ALT testing, while during intervention, 43.7% had a gap (P=0.003). ALT testing among patients chronically taking methotrexate or leflunomide identified 0.8% of results >2x ULN during baseline and 1.6% during intervention. Discussions are underway to expand the outreach to additional patient populations.

Lessons learned: Communication technology outreach to remind patients to obtain laboratory testing improved the timeliness of availability of recommended test results and revealed an increased proportion of abnormal findings, enabling more proactive evaluation for iatrogenic harm. A medical data warehouse and IVR system facilitated patient adherence to the diagnostic regimen.

#### Early Intervention to Reduce Need for Mechanical Ventilation in Sepsis Patients with Respiratory Failure in the **Emergency Department**

H. Soleimani<sup>1</sup>, K. E. Henry<sup>1</sup>, A. Zhan<sup>1</sup>, Y. Ahmad<sup>1</sup>, T. Wang<sup>2</sup>, N. Rawat<sup>3</sup>, D. Hager<sup>2</sup>, D. R. Thiemann<sup>2</sup>, A. Markowski<sup>4</sup>, A. Sridharan<sup>5</sup>, S. Saria<sup>6</sup>

<sup>1</sup>Johns Hopkins University, Baltimore, MD

<sup>2</sup>The Johns Hopkins University, Baltimore, MD

<sup>3</sup>Armstrong Institute for Patient Safety and Quality

<sup>4</sup>The Johns Hopkins Medical Institutions, Bethesda, MD

<sup>5</sup>The Johns Hopkins Medical Institutions, Columbia, MD

<sup>6</sup>Johns Hopkins School of Public Health

Statement of problem: Sepsis is one of the leading cause of complications in patients with respiratory failure (RF). Studies have shown that early treatment of sepsis and acute RF associated with sepsis improves outcomes such as hospital length of stay and need for invasive mechanical ventilation (MV). Diagnosis of sepsis is often delayed in large due to lack of reliable objective diagnostic criteria; most existing identification tools are also rule-based criteria with high false alarm rate and low adoption. We investigate the impact of a highly specific system for timely identification of sepsis in the Emergency Department (ED) on patients with acute RF.

Description of the intervention or program: TREWS (Targeted Real-time Early Warning System) is a machine-learning based system that continuously monitors patients' routinely collected EHR data and sends a notification when it identifies a patient is at high risk for sepsis; providers can either dismiss the alert or confirm sepsis and initiate treatment. We developed TREWS using a statistical algorithm that uses historical data to learn patterns for early identification of sepsis. At a sensitivity of 80%, TREWS has a specificity of 92% and positive predictive value of 47% (i.e. approximately 1 in every 2 alerts is correct). TREWS has been piloted in the ED of a community hospital since November 2017. Findings to date: The intervention population consists of patients with signs of RF in the ED and with TREWS alerts that were confirmed as sepsis by providers between November 15, 2017- May 15, 2018. The control (retrospective) population consists of patients with retrospective TREWS alerts that identified as having sepsis and had signs of RF in the ED during Nov. 15, 2016 - May 31, 2017. In the intervention population

11% vs. 20% in the retrospective period required MV. Table 1 also reports improvement in hospital LOS, ICU admission, and rate of developing new organ failures.

Lessons learned: This study shows that early identification is associated with improved clinical outcomes in sepsis patients with RF in the ED. The improvements can be attributed to multiple factors: we see a 56% increase in the rate of non-invasive oxygen therapy in the ED, which has been shown in other studies to reduce LOS and need for MV. The time from ED admission to initiating antibiotics has also decreased by 19%. We plan to deploy TREWS to other hospitals (ED and inpatient units) and study the impact on other patient populations in a prospective randomized trial.

Table 1: Comparison of clinical outcomes between the intervention and control periods.

Variable		Intervention Period (11/15/2017 to 05/31/2018)	Control Period (11/15/2016 to 05/31/2017)	
Number of patients included in the population		87	99	
	Male	42	50	
Population	Average Age	66.5	62.4	
Characteristics	Asthma	14	17	
	Emphysema	2	6	
	Chronic Airway Obstruction	25	22	
Fraction requiring invasive mechanical ventilation during their hospital stay		11%	20%	
ICU admission rate		63%	78%	
Median (IQR) ICU length of stay		3.01 days (IQR = [1.48, 6.27])	3.36 days (IQR = [2.09, 6.91])	
Median hospita	al length of stay	6.0 days (IQR = [3.25, 10.0])	7.0 days (IQR = [4.0, 13.0])	
	eloping new organ failure EWS alert	16.7%	24.2%	
Fraction receiving non-invasive oxygen therapy in the ED		39%	25%	
Median time from ED a		1.70 hours (IQR = [0.65, 4.5])	2.1 hours (IQR = [1, 3.85])	

<sup>\*</sup>Respiratory failure is defined as patients who require ventilation (NIV or MV) at some point during their hospital stay who experience at least one the following conditions in the ED: PaO2 < 60, or SpO2 < 91, or (PaCO2 > 50 and arterial pH < 7.35), or PaO2/FiO2 < 300.

## **Oral Presentation of Clinical Vignette**

Tuesday, November 6, 2018 | 10:15 AM - 11:45 AM

#### A Plague of Rats

S. Godil1, Z. Qamar2, J. Akhtar3 <sup>1</sup>UPMC Mercy, Pittsburgh, PA <sup>2</sup>Dow University of Health Sciences, Karachi, Pakistan <sup>3</sup>UPMC, Pittsburgh, PA

Learning objectives: This bias describes a physician's unwillingness to consider a diagnosis because it is very obscure, even if it is correct. In this context it means backing away from a rare diagnosis.

Case information: A 25 year-old man presented for with three days of nausea and vomiting, and mild abdominal pain. The work-up for abdominal pain was negative and patient was discharged with diagnosis of gastroenteritis. He returned two days later with nausea, vomiting and upper quadrant abdominal pain. CT scan abdomen was concerning for acute cholecystitis and he was taken to the operating room for cholecystectomy. Post operative course was complicated by desaturation and bloody frothy sputum requiring reintubation. Chest X-ray showed diffuse alveolar hemorrhage (DAH). The working diagnosis was sepsis syndrome. Infectious diseases was consulted, and obtained the history of him being exposed to sewage when he was putting a new sewer tap. This history and clinical symptoms of DAH/ARDS, and pancytopenia with cholecystitis picture led to the diagnosis of leptospirosis and patient was treated for it. Patient did well and was discharged home. His leptospira DNA was later reported as positive.

Discussion: Zebra retreat is the process of rejecting uncommon presentations or rare presentations of common presentations. Zebra retreat occurs when a physician retreats from a serious consideration of an unusual diagnosis, perhaps because he is unfamiliar with the condition or does not want to over utilize resources.

### **Poster Session 1: Scientific Abstracts**

Sunday, November 4, 2018 | 4:15 PM - 5:45 PM

#### Diagnostic Pitfalls: A New Construct to Understand Diagnosis Errors

S. D. Agarwal<sup>1</sup>, E. Ruan<sup>2</sup>, H. Reyes Nieva<sup>1</sup>, G. D. Schiff<sup>1</sup> <sup>1</sup>Brigham and Women's Hospital, Boston, MA <sup>2</sup>Tufts University, Boston, MA

Background: Diagnosis errors are common in the medical field and can lead to irrevocable harm. Although mitigating these errors are a critical part of the patient safety movement, they remain poorly understood. To improve our understanding of diagnosis errors, we developed a new construct known as "diagnostic pitfalls," that we defined as vulnerabilities in the diagnostic process that predispose to diagnosis errors. Methods: We examined multiple data sources to identify diagnostic error cases. Sources included: (1) narrative literature review, (2) an online repository of morbidity and mortality (M&M) cases maintained by the Agency for Healthcare Research and Quality (AHRQ) as well as institutional M&M cases with a focus on ambulatory care, (3) event reports related to patient safety and risk management, (4) focus groups with specialist physicians, and (5) closed diagnosis-related primary care malpractice claims from two leading malpractice insurers in Massachusetts. From these cases, we identified clinical situations with patterns of vulnerabilities where we could delineate potential diagnostic pitfalls. All cases were then categorized by diagnosis. For the ten most common diagnoses, we then used our curated list of pitfalls to determine how often they occurred among those cases.

Results: From the five data sources, we identified 991 cases. Eighteen themes, or diagnostic pitfalls, emerged from our analysis (Figure). These included disease A repeatedly misdiagnosed as disease B; failure to appreciate limitations of test results or exams; diagnoses thrown off by atypical presentation; presumption that chronic diseases account for new symptoms; overlooking drug or other environmental causes; and failure to monitor evolving symptoms. The top 10 most common diagnoses represented among the 991 cases included cancer (colon, lung, prostate, breast, thyroid and melanoma) as well as myocardial infarction, sepsis, stroke, and pulmonary embolism. These diagnoses accounted for approximately a quarter of all cases. In cases related to colon cancer, for example, the top pitfalls were that counter-diagnosis cues being overlooked (46%), failure to follow-up (22%), risk factors not adequately appreciated (22%), disease A repeatedly mistaken for disease B (20%), communication failures between primary care physician and specialist (15%), and limitations of a test not appreciated (11%). Conclusion: We developed a new construct known as diagnostic pitfalls to help identify vulnerabilities faced by front-line physicians. Cancer diagnoses are among the most common cases of diagnosis error, and we identified several pitfalls common among these cases. Initiatives to reduce diagnosis errors should focus on finding ways to mitigate these vulnerabilities.

Diagnostic Pitfalls Among the Top 10 Diagnoses Represented in Cases of Diagnosis Error												
Diagnostic Pitfalls		/	//	//	My St Garcer IN 34,	/ /	Show 161 "1M" 291	Pulme 16)	/ /	Mela Mela	Overall M = 241	/
Failure to follow-up	22%	33%	47%	42%	21%	-	-	-	57%	67%	28%	
Limitations of a test or exam finding not appreciated	11%	28%	24%	42%	41%	-	25%	18%	43%	33%	26%	
Disease A repeatedly mistaken for disease B	20%	26%	15%	-	41%	38%	50%	27%	-	67%	24%	
Risk factors not adequately appreciated	22%	-	44%	12%	24%	-	-	27%	-	-	16%	
Atypical presentation	-	-	12%	45%	28%	19%	44%	-	-	-	15%	
Counter-diagnosis cues overlooked (e.g. red flags)	46%		-	-	-	25%		-	-	-	10%	
Communication failures between primary care physician and specialist	15%	9%	-	12%	10%	-	-	-	-	-	7%	
Issues surrounding referral	-	-	9%	-	14%	-	-	-	-	-	3%	
Urgency not fully appreciated	-		-	-	-	25%	13%	-	-	-	2%	
Chronic disease presumed to account for new symptoms	-	-	-	-	-	19%	6%	-	-	-	2%	
Miscommunication related to lab ordering	-	-	-	-	-	19%	-	-	-	-	1%	
Evolving symptoms not monitored	-	5%	-	-	-	-	-	-	-	-	1%	
Drug or other environmental cause overlooked	-	-	-	-	-	-	-	-	-	-	0%	
Empiric treatment delayed recognition (by masking symptoms or signs)		-	-	-	-	-	-	-	-	-	0%	
Diagnosis is rare or unfamiliar	-		-		-	-		-		-	0%	
Symptoms are non-specific or vague	-	-	-	-	-	-	-	-	-	-	0%	
Symptoms are intermittent	-		-		-	-	-	-	-	-	0%	
Language-related communication failures	-	-	-	-	-	-	-	-	-	-	0%	

#### Emergency Department Diagnostic Accuracy in ACUTE Dizziness and Vertigo: Results from the Avert Logistical PILOT Trial

A. Batazzi<sup>1</sup>, T. P. Chang<sup>2</sup>, N. Tevzadze<sup>1</sup>, V. Eslami<sup>1</sup>, M. Fanai<sup>3</sup>, J. Kattah<sup>4</sup>, K. Kerber<sup>5</sup>, W. Meurer<sup>6</sup>, D. Hanley<sup>1</sup>, A. Blitz<sup>7</sup>, J. Carey<sup>1</sup>, D. Zee<sup>1</sup>, D. E. Newman-Toker<sup>3</sup>

<sup>1</sup>Johns Hopkins School of Medicine, Baltimore, MD

<sup>2</sup>Taichung Tzu Chi Hospital, Buddhist Tzu Chi Medical Foundation, Taichung, Taiwan

<sup>3</sup>Center for Diagnostic Excellence, Armstrong Institute for Patient Safety and Quality, Johns Hopkins University School of Medicine, Baltimore, MD

<sup>4</sup>University of Illinois College of Medicine at Peoria, Illinois Neurological Institute, Peoria, IL

<sup>5</sup>University of Michigan, Ann Arbor, MI

<sup>6</sup>University of Michigan

<sup>7</sup>Johns Hopkins School of Medicine

Background: Retrospective studies of acute dizziness and vertigo suggest high rates of diagnostic error in patients with vestibular disorders, but prospective studies are lacking. Acute Video-oculography for Vertigo in Emergency Rooms for Rapid Triage (AVERT, NCT02483429) is a multicenter, phase II randomized clinical trial comparing the impact of video-oculography (VOG)-guided rapid triage to standard Emergency Department (ED) diagnostic care for patients with a chief complaint of acute dizziness or vertigo with evidence of pathologic nystagmus on screening bedside VOG examination. We report baseline ED diagnostic accuracy from the two-year logistical pilot phase.

Methods: From November 2015 to November 2017, we pre-screened 1,799 and identified 291 ED patients with acute dizziness or vertigo at three centers. We conducted a standardized screening battery of symptom history, VOG tests using the ICS Impulse- (Otometrics), and neurologic exams. A single neuro-otologist screened the first 100 complete VOG results for pathologic nystagmus. For these suspected vestibular cases, two independent neuro-otologists, masked to the ED diagnosis, reviewed eye movement videos, quantitative VOG recordings, and other clinical information to determine the underlying cause of the dizziness or vertigo. Patients were classified according to three hierarchical diagnostic groupings of different granularity (12-category; 6-category [primary outcome]; 3-category). Differences were adjudicated by a third neuro-otologist rater, also masked to ED diagnosis. Recorded ED diagnoses were compared to expert diagnoses using the same categories.

**Results:** Nystagmus was present in 30% (n = 30/100) of screened records. Final diagnoses were peripheral vestibular disorders (73%, n = 22/30), central vestibular disorders (20%, n = 6/30), and unknown (7%, n = 2/30). Inter-rater agreement between experts on final diagnosis was high (12-category  $\kappa$  = 0.82; 6-category  $\kappa$  = 0.87; 3-category  $\kappa$  = 0.87). ED diagnostic accuracy was low for all levels of granularity: 12-category (17%, n=5/30); 6-category (20%, n=6/30) (Figure); 3-category (33%, n=10/30). The ED team recorded pathologic nystagmus in chart notes in just 43% (n = 13/30), and noted this was found by the AVERT trial coordinator in three of these. Among peripheral cases, the ED team obtained unnecessary magnetic resonance imaging (MRI) in 68% (n=15/22) and non-indicated hospital admission in 64% (n=14/22). Among central cases, 67% (n = 4/6) were inappropriately discharged from the ED.

**Conclusion:** ED diagnosis of vestibular patients with acute dizziness or vertigo is highly inaccurate, operating at nearly chance performance. This resulted in both over- and under-utilization, including inappropriate final disposition in 60%. Interventions to address this problem are needed. The impact of device-based tele-consultation and diagnostic decision support is being assessed in the AVERT trial.

Same	diagnosis	Central - Periph	VOG-guided diagnosis					
Perip	h - Periph	Central- Equiv		Peripheral		Equi	vocal	Central
Perip	h - Central	Equiv - Equiv	P1	P2	P3	01	O2	C1
	al	P1	2	1				1
is	Peripheral	P2	=					
diagnosis	Pe	P3	2	-	2			
) dia		01		1		5	1	
ED	Equivocal	O2	2	5	4		1	4
	Central	C1	1		2			1

Figure 1: Individual diagnoses made by the ED physicians (vertical axis) vs VOG-guided reviews (horizontal axis) following the AVERT 6 category diagnosis schema.

#### Understanding Diagnostic Practices of Physicians Treating Non-Specific Abdominal Pain

V. Bauer<sup>1</sup>, A. Solomonides<sup>2</sup>, K. Kirley<sup>3</sup>, P. Epner<sup>4</sup>, G. Rao<sup>5</sup>

<sup>1</sup>NorthShore University HealthSystem, Evanston, IL

<sup>2</sup>NorthShore University Health System, Evanston, IL

<sup>3</sup>American Medical Association, Chicago, IL

<sup>4</sup>Paul Epner LLC, Evanston, IL

<sup>5</sup>Case Western Reserve University & University Hospitals of Cleveland, Cleveland, OH

Background: Abdominal pain is among the most challenging complaints to evaluate, and one which is prone to serious diagnostic errors. Diagnosis is complex, uncertain, and error-prone. A diagnostic path, as defined by the authors, consists of diagnostic steps taken from initial presentation until a diagnosis is obtained or the evaluation ends for other reasons. In many cases, diagnostic processes followed by different physicians for the same presenting complaint are highly variable. Our objective was to interview physicians to gain insight into diagnostic practices related to abdominal pain in order to better understand how diagnostic processes are determined.

Methods: From a large suburban integrated healthcare system, we invited 30 primary care physicians for semi-structured interviews, 7 agreed to participate. Each physician was told to describe their preferred diagnostic practices for a fictional case of a patient with abdominal pain, and what they found challenging. Using grounded theory, transcripts were scoured for notable comments and patterns until a mutuallyagreed upon list of themes was developed.

Results: Physicians emphasized patient history and individual characteristics when it came time to evaluate potential diagnostic paths and providing a diagnosis. Physicians were also very concerned about missing serious diagnosis, stressing the identification of "red flags" during clinical evaluations. Additionally, laboratory and other testing supported the clinical evaluation, rather than being used as a primary tool to establish a diagnosis. And consultants, including gastroenterologists are primarily used to carry out specific procedures rather than for their clinical acumen in making diagnoses.

Conclusion: Physicians underscored the importance of history taking over other types of evaluation. It may be the case that clinicians, well aware that a diagnosis may not be reached in a large proportion of patients, may actually order diagnostic tests only in patients whom a specific diagnosis is more likely. Further analysis on a large scale of Electronic Health Record (EHR) data can widen our understanding of clinician decision making in diagnostic paths and minimize problems such as attribution or the absence of free text notes, since large numbers of physicians with patients with broadly similar characteristics can be analyzed. Future directions include providing visualizations of diagnostic paths identified from the EHR data to physicians and interviewing them about their own patterns in comparison to their colleagues for the same clinical presentations.

#### Frequency of Diagnostic Errors and Associated Risk Factors in the Micu - a Retrospective Cohort Study

P. A. Bergl<sup>1</sup>, A. Taneja<sup>1</sup>, R. El-Kareh<sup>2</sup>, H. Singh<sup>3</sup>, R. S. Nanchal<sup>1</sup>

<sup>1</sup>Medical College of Wisconsin, Milwaukee, WI

<sup>2</sup>University of California, San Diego, La Jolla, CA

<sup>3</sup>Center for Innovations in Quality, Effectiveness and Safety, Michael E. DeBakey Veterans Affairs Medical Center and Baylor College of Medicine, Houston, TX

Background: Diagnostic error rates in critically ill patients historically have been estimated from autopsy series; these studies, however, are limited by referral bias and the exclusion of survivors. We sought to determine the frequency of diagnostic errors in unplanned medical ICU admissions and risk factors contributing to these errors.

Methods: We conducted a retrospective record review of medical ICU admissions at a single academic hospital. We randomly selected 436 records from a 12-month study period. After excluding planned admissions, interhospital transfers, and patients cared for by study team members, we analyzed 256 unplanned ICU admissions. The primary investigator reviewed records using an abbreviated form of the Safer Dx instrument consisting of 6 of 11 items which had previously been found to correlate most strongly with diagnostic errors. We collected demographic and physiologic data from which severity of illness scores (e.g. qSOFA, APACHE2) were computed. Secondary record reviews were performed using a tiered approach, and reviewers met periodically to adjudicate disagreements. For each case, the team determined whether a diagnostic error had occurred prior to and/or within 24 hours of ICU admission. We performed multivariable logistic regression analyses to identify risk factors for diagnostic errors. We also performed regression analyses to examine relationships between adjudicated diagnostic errors and individual items on the Safer Dx instrument.

Results: Of 256 unplanned medical ICU admissions, 18 (7.0%) were associated with diagnostic errors prior to or within 24 hours of transfer. Only 6 of these errors were recognized by the ICU team within the first 24 hours. Delayed recognition of sepsis or septic source was the most common error (11/18, 61%). Female sex was independently associated with risk of diagnostic error both at admission (OR 3.97, 95%CI 1.30-12.03) and at 24 hours (OR 5.59, 95%CI 1.40-22.28). A qSOFA screening score of 2 or greater at admission was also associated with diagnostic error (OR 5.240, 95%CI 1.767-15.537). When dichotomized by simple agreement, three Safer Dx instrument questions independently predicted identification of a diagnostic error: (1) whether the history suggested alternative diagnoses (OR 6.76, 95%CI 1.00-45.56), (2) whether alarm symptoms were acted upon early enough (OR 21.17, 95%CI 2.46-181.74), and (3) whether the final diagnosis was an evolution of the initial presumed diagnosis (OR 0.032, 95%CI 0.004-0.278).

**Conclusion:** Diagnostic errors occurred in 7% of unplanned ICU admissions, and most were not recognized within 24 hours of admission. Female patients and those with qSOFA scores >/=2 had significantly elevated risk.

#### Stroke Risk after Diagnosis of "Benign Vertigo" Is Lower in Specialty Care Than General Practice

A. Bery<sup>1</sup>, T. P. Chang<sup>2</sup>, H. Y. Chuang<sup>2</sup>, Z. Wang<sup>3</sup>, D. E. Newman-Toker<sup>4</sup>

<sup>1</sup>University of Ottawa, Ottawa, ON, Canada

<sup>2</sup>Taichung Tzu Chi Hospital, Buddhist Tzu Chi Medical Foundation, Taichung, Taiwan

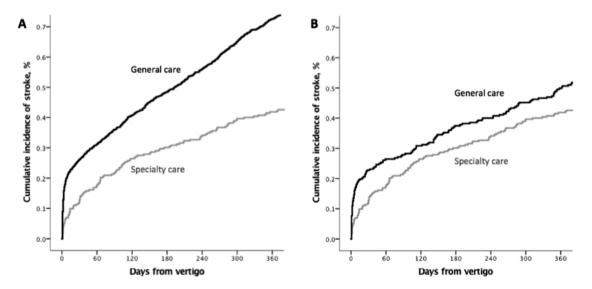
<sup>3</sup>Johns Hopkins University School of Public Health, Baltimore, MD

<sup>4</sup>Center for Diagnostic Excellence, Armstrong Institute for Patient Safety and Quality, Johns Hopkins University School of Medicine, Baltimore, MD

**Background:** Stroke presenting with dizziness/vertigo is often misdiagnosed as benign, causing potentially preventable, misdiagnosis-related harms. Harms may be measured using short-term risk of stroke hospitalization after outpatient discharge with "benign vertigo." These short-term risks vary across specialties, implying misdiagnosis, but baseline stroke risk may vary because of referral bias, confounding the association. To differentiate misdiagnosis-related harms from referral bias, we used propensity-score matching to compare short-term stroke risk across settings.

**Methods:** Patients with first incident dizziness/vertigo (ICD-9-CM 386.x, 780.4) in outpatient departments were included from the National Health Insurance Research Database of Taiwan (2002-2009). We excluded patients with emergency department referral or hospitalization at the index visit and patients initially diagnosed with stroke-related or "central" vertigo. Patients were classified into two groups based on the medical discipline of the clinic provider—general care (internal medicine, family medicine) or specialty care (neurology, otolaryngology). Each patient in specialty care was matched to one patient in general care with similar propensity score (calculated based on 27 covariates). The relative risk of stroke at 30, 90, and 360 days after "benign vertigo" diagnosis were compared between general care and specialty care before and after propensity-score matching.

Results: We analyzed 178,981 patients diagnosed with "benign vertigo" (n = 117,117 diagnosed in general care; n = 27,238 diagnosed in specialty care). Before propensity-score matching, the relative risk of stroke (ratio specialty to general care) at 30, 90, and 360 days were 0.49 (95% CI, 0.35-0.70), 0.60 (95% CI, 0.46-0.79), and 0.58 (95% CI, 0.47-0.70), respectively. A subset of general care patients (n = 27,238) were matched to the specialty care patients by propensity-score methods. After matching, the relative risk of stroke (ratio specialty to general care) at 30, 90, and 360 days were 0.57 (95% CI, 0.37-0.85), 0.78 (95% CI, 0.55-1.09), and 0.84 (95% CI, 0.65-1.08), respectively. Cumulative stroke incidence curves without and with propensity-score matching demonstrate that higher short-term risk persists after adjusting for baseline differences (Figure). Conclusion: Risk of delayed stroke hospitalization after "benign" dizziness/vertigo discharge differs between general and specialty care. Long-term risk is higher in general care, but this difference is largely eliminated by propensity-score matching, indicating higher baseline stroke risk in general care and greater long-term stroke incidence unrelated to misdiagnosis. By contrast, the statistically-significant difference in short-term risk remains, despite matching, and being diagnosed in general care carries twice the risk of misdiagnosis-related harms. Solutions such as improved training or diagnostic decision support should be considered to address this care gap.



**Figure.** Cumulative stroke incidence before (A) and after (B) propensity-score matching. Long-term risk is higher in general care, but this difference is largely eliminated after matching. By contrast, short-term risk (<30 days) remains, and being diagnosed in general care confers a nearly two-fold higher risk of misdiagnosis-related harms.

#### A Skill-Based Clinical Reasoning Curriculum for Clerkship-Level Medical Students: A Randomized and Controlled Study

E. Bonifacino<sup>1</sup>, W. Follansbee<sup>2</sup>, A. H. Farkas<sup>3</sup>, M. McNeil<sup>2</sup>, D. DiNardo<sup>2</sup> <sup>1</sup>University of Pittsburgh Medical Center, Pittsburgh, PA <sup>2</sup>University of Pittsburgh School of Medicine, Pittsburgh, PA 3Medical College of Wisconsin, Milwaukee,, WI

Background: A 2015 IOM report highlighted the urgent need for better training in medical decision-making. Despite this, a recent national survey reported that up to 57% of medical schools lack formal education dedicated to clinical reasoning. We sought to determine if formalized instruction in clinical reasoning through use of an innovative multi-modal curriculum is an effective means of improving clinical reasoning skills in third year medical students.

Methods: We conducted a randomized and controlled experiment to evaluate the impact of this curriculum on clinical reasoning knowledge and skills in third year medical students at a single allopathic U.S. medical school. Students in the intervention group completed interactive online modules focused on clinical reasoning concepts and participated in a skills-based workshop. Students in the control group participated in standard educational experiences. We assessed the impact of the curriculum on clinical reasoning knowledge, skills, and perceptions of students' own education according to the following metrics: 1. Performance on a 20-item quiz about clinical reasoning concepts at the conclusion of the 4-week intervention period; 2. Clinical reasoning skills through blinded scoring of weekly student hospital admission notes using a validated clinical reasoning assessment tool; 3. Perception of students' own education using a Likert scale. Comparisons between control and intervention group were conducted using a t-test for independent samples.

Results: In total, 67 students participated from whom we received a total of 250 hospital admission notes. Students in the intervention group demonstrated superior performance on the 20-item clinical reasoning quiz (67% correct vs. 54%, p < 0.05). Admission notes from the intervention group received significantly higher average ratings for data synthesis (2.3 vs. 2.0 on a 3-point scale, p < 0.05) and for diagnostic reasoning (2.2 vs. 1.9, p < 0.05), while average scores for data gathering, reporting, and decision-making did not differ. Students in the intervention group reported more frequent use of clinical reasoning terminology and more explicit discussion of clinical reasoning by their attending physicians. Conclusion: Exposure to our innovative multi-modal curriculum was associated with improved knowledge regarding clinical reasoning concepts and superior written demonstration of clinical reasoning skills in 3rd year medical students. Thus, interactive online modules paired with opportunity for skills practice is an effective and efficient means of delivering clinical reasoning education. Our student survey responses suggest that participation in such experiences may prime students to identify demonstration of clinical reasoning by clinical role models, thus providing a scaffolding for development of medical decision-making skills in the clinical setting.

#### Marked Variation in Symptom Combinations at the Individual Level in Women and Men with Acute Myocardial Infarction.

J. Brush<sup>1</sup>, R. Dreyer<sup>2</sup>, G. Norman<sup>3</sup>, R. Lamichhane<sup>4</sup>, B. McMichael<sup>5</sup>, H. Krumholz<sup>2</sup> <sup>1</sup>Eastern Virginia Medical School and Sentara Healthcare, Norfolk, VA <sup>2</sup>Yale University 3McMaster University, Hamilton, ON, Canada <sup>4</sup>Eastern Virginia Medical School <sup>5</sup>Sentara Healthcare

Background: The diagnosis of acute myocardial infarction (AMI) is under-recognized in women, as compared with men. The population frequencies of some AMI symptoms differ in women and men, however, distributions of symptoms at the population level do not necessarily determine how symptoms cluster in individuals. Because clinicians recognize diagnoses in individuals and remember individual examples, symptom combinations in individuals (phenotypes) and gender differences in phenotypes may be more important for recognizing a diagnosis. Methods: We utilized data from the VIRGO Study, a large NIH-sponsored multi-center registry of 3,501 young AMI women and men aged 18-55 years. VIRGO collected symptom information (e.g. chest pain, nausea) using direct interviews by study coordinators (10 interview symptoms) and from the abstracted medical record (9 abstracted symptoms). We analyzed symptom phenotypes for AMI patients to assess overall variation and compared between genders to determine whether symptom phenotypes are different in women versus men. We also compared phenotypes using symptoms from the direct interviews to those abstracted from the recorded medical record. Differences in the distribution of phenotypes were analyzed using Chi Squared analysis.

Results: In total, there were 488 different phenotypes of AMI using the interview symptoms. Over half of the patients had one of 38 phenotypes, and the remaining half were spread over 450 different phenotypes. The distribution of phenotypes was statistically different in women and men (p<0.001). There were 426 phenotypes in women and 280 phenotypes in men. The top 25 phenotypes included 40% of the women, with a mean of 4.1 symptoms, and 46% of the men, with a mean of 3.7 symptoms (p < 0.001). Using the abstracted symptoms, there were 272 phenotypes, substantially fewer than the 488 phenotypes using the interview symptoms. The distribution of phenotypes using the abstracted symptoms was also statistically different in women (245 phenotypes) and men (151 phenotypes, p < 0.001). The average number of abstracted symptoms was not different in women and men (3.0 versus 2.8).

Conclusion: We found enormous variation in individual-level symptom combinations in patients with AMI, a finding that is obscured by population-level symptom frequencies. There were substantially more symptom phenotypes in women than men. There were fewer phenotypes using the abstracted symptoms, suggesting that clinicians heuristically lump symptom combinations when collecting and recording symptom histories. Knowledge about phenotypic variation and differences in women and men could have important implications for teaching and assuring diagnostic accuracy and reliability and could reduce gender disparities in AMI diagnosis.

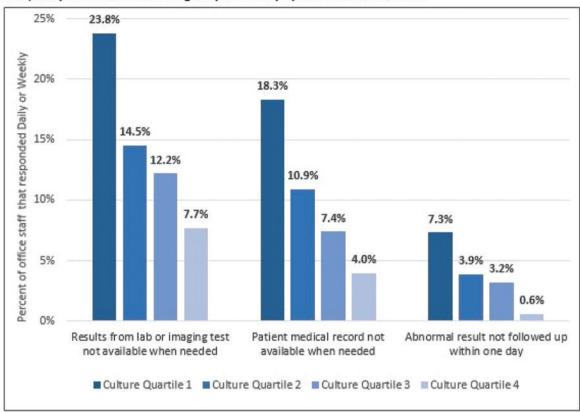
#### Patient Safety Culture, Health It Implementation, and Medical Office Problems That Could Lead to Diagnostic Error.

J. R. Campione<sup>1</sup>, R. E. Mardon<sup>2</sup>, K. M. McDonald<sup>3</sup> <sup>1</sup>Westat, Durham, NC <sup>2</sup>Westat, Rockville, MD 3Stanford University, Stanford, CA

**Background:** Diagnostic errors may cause harm to patients by preventing or delaying appropriate treatment, or by leading to unnecessary or harmful treatment. This study investigates the relationship between patient safety culture, health IT implementation, and the frequency of problems that could lead to diagnostic errors in the medical office setting. Methods: We used survey data from 925 medical offices nationwide that voluntarily submitted results to the 2012 Agency for Healthcare Research and Quality Medical Office Surveys on Patient Safety Culture- database. The following three problems were analyzed: 1) patient medical chart not available when needed; 2) testing results not available when needed; and 3) abnormal results not followed up within one day. For each problem, frequency was defined as the percent of survey takers that responded that the problem happened daily or weekly over the past twelve months. At the office level, we ran a multivariate regression model to estimate the effect of culture on problem frequency while controlling for office-reported implementation levels of health IT; office characteristics such as the number of locations; and survey characteristics such as the percent of respondents that were physicians.

Results: The most frequent problem was "results from a lab or imaging test were not available when needed"; across 925 offices, the average was 15% reporting that it happened daily or weekly. Higher overall culture scores were significantly associated with fewer occurrences of each problem assessed. Compared to offices with completed health IT implementation, offices in the process of health IT implementation had higher frequency of problems. An increase in the culture dimension titled Patient Care Tracking and Follow-up was associated with lower Percent Daily or Weekly for both Results Not Available (p<.001) and Untimely Follow-up of Abnormal Result (p<.001). An increase in positive Work Pressure and Pace dimension scores was associated with a decrease in the frequency of Results Not Available (p = .002) and Patient Record Not Available (p=.035).

Conclusion: This study offers insight into how patient safety culture and health IT implementation in medical offices can influence the frequency of breakdowns in processes of care, thereby identifying potential vulnerabilities that can increase diagnostic errors. As measured by the AHRQ Medical Office Survey on Patient Safety Culture, lower overall culture scores and the implementation phase of health IT was associated with more frequent office problems.



#### Frequency of Problems Occurring Daily or Weekly by Office Culture Quartile\*

Note: Per problem, the office problem mean percent for Quartile 1 and for Quartile 4 are significantly different from the other quartiles' mean percent (p < 0.05).

#### Missed Diagnosis of Pediatric Sepsis: Analysis of a Large Population-Based Sample

T. L. Cifra<sup>1</sup>, E. Westlund<sup>2</sup>, P. Ten Eyck<sup>2</sup>, M. M. Ward<sup>2</sup>, N. M. Mohr<sup>2</sup>, D. A. Katz<sup>2</sup> <sup>1</sup>University of Iowa, Iowa City, IA <sup>2</sup>University of Iowa

Background: Missed diagnosis of pediatric sepsis contributes significantly to poor outcomes. There are currently no population-based estimates of the risk of missed sepsis in children. Our objective is to estimate the risk of missed pediatric sepsis in US emergency departments (EDs) and determine associated patient/institutional factors.

Methods: We conducted a retrospective analysis of linked inpatient and ED records from Healthcare Cost and Utilization Project (HCUP) databases in 2010-2011 from four states (CA, FL, MA, NY). Using ICD-9 codes, we identified pediatric patients (0-19 years) admitted with severe sepsis (995.92) and/or septic shock (785.52) who had at least one ED treat-and-release visit in the 7 days prior to admission. A three-member expert panel of pediatricians then rated how likely each ED visit was related to the patient's subsequent sepsis admission, given their age and HCUP Clinical Classification System diagnosis codes assigned for the ED visit. Those with a prior ED visit rated by at least two reviewers as likely to be related to their sepsis admission were categorized to have had a probable missed diagnosis of sepsis. Generalized linear mixed models were constructed to determine multivariable associations between patient/institutional characteristics and probable missed sepsis. Results: 159 of 1,945 patients (8%) admitted with severe sepsis/septic shock had probable missed sepsis in the ED. The most common reasons for ED treat-and-release visits followed by a sepsis admission were fever of unknown origin (29%), other lower respiratory disease (12%), and viral infection (12%). Preliminary results show that the odds of probable missed sepsis were significantly higher among black patients compared to white patients (OR 2.9 [1.0, 8.0]), lower for each additional comorbid chronic condition (OR 0.3 [0.2, 0.4]), and lower among patients seen in a facility with a pediatric intensive care unit (PICU) (OR 0.3 [0.1, 0.8]). The patient's gender/age, primary payer, type of hospital ownership (private vs. public), and day of admission (weekday vs. weekend) were not significantly associated with probable missed sepsis.

<sup>\*</sup> The 925 offices were ranked and categorized into relative equal-size quartiles of overall culture score, with Quartile 4 having top culture.

Conclusion: We estimate that 8% of children admitted for severe sepsis/septic shock had a missed diagnosis of sepsis in a prior ED visit. Preliminary analysis suggests that a high index of suspicion for sepsis may prevent misdiagnosis among black children and in children without comorbid conditions. Institutions without dedicated pediatric acute care services such as a PICU may be at higher risk of misdiagnosing pediatric sepsis in the ED and should consider interventions to reduce diagnostic error in this population.

#### Piloting Methods to Determine Factors Associated with Diagnostic Error on Admission to a Pediatric Intensive Care **Unit (PICU)**

T. L. Cifra<sup>1</sup>, P. Ten Eyck<sup>2</sup>, J. D. Dawson<sup>2</sup>, V. M. Steelman<sup>2</sup>, H. S. Reisinger<sup>3</sup>, H. Singh<sup>4</sup>, L. A. Herwaldt<sup>3</sup>

<sup>1</sup>University of Iowa, Iowa City, IA

<sup>2</sup>University of Iowa

<sup>3</sup>University of Iowa, IOWA CITY, IA

<sup>4</sup>Center for Innovations in Quality, Effectiveness and Safety, Michael E. DeBakey Veterans Affairs Medical Center and Baylor College of Medicine, Houston, TX

Background: Diagnostic errors (DEs) cause significant harm to critically ill children. However, little is known about factors that affect the occurrence of error. Our objective is to conduct a feasibility study to determine methods and sample size required to identify patient, provider, and work system factors associated with DE occurring in the first 12 hours after non-elective PICU admission.

Methods: We conducted a pilot retrospective cohort study of 50 consecutive non-elective admissions of patients 0.17 years old to a tertiary referral academic PICU. A single physician reviewer conducted a structured chart review using the Safer Dx Instrument, a validated medical record review tool (71% sensitivity, 90% specificity), to detect DEs occurring during the first 12 hours after PICU admission. A non-physician reviewer independently collected data on patient, provider, and work system factors. We performed descriptive data analysis, grouping factors by DE status. We calculated the required sample size to determine significant associations of various factors with DE (80% power at a = 0.05) accounting for potential clustering of DE by provider. Preliminary analysis with the t test/Wilcoxon rank-sum test and Fisher's exact test was performed to compare the occurrence of DE with respect to potential associated factors.

Results: We found 4/50 (8%) patients who had DEs occurring within the first 12 hours of PICU admission. One had a missed diagnosis of chronic ear infection, two had delayed diagnoses of increased intracranial pressure (one with ischemic stroke and another with subdural hemorrhage), and one had delayed diagnosis of cat-scratch disease. Three out of 4 cases presented with neurologic symptoms in the PICU. Unadjusted analysis did not identify statistically significant differences between patients with vs. without DE, except for discrepancies between PICU admission diagnoses and PICU discharge diagnoses. We calculated that we need to review 610 PICU admissions to achieve 80% power to detect significant associations with DE.

Conclusions: We successfully conducted a pilot of methods to determine factors associated with DE on PICU admission. Preliminary results show an 8% rate of DEs. The Safer Dx instrument facilitated chart review, made assessments more objective, and was easy to apply in the ICU setting. We will now investigate associations between patient, provider, and work system factors and DE on admission with a larger sample size for adequate power. We need to determine factors associated with DE in the PICU as a necessary step to design future interventions to prevent DEs in this setting.

#### Interruptions to Clinical Reasoning during in-Situ Observations of Medical Students

C. Clancy

Hospital of the University of Pennsylvania, Philadelphia, PA Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA

Background: Clinical reasoning is a core skill in medicine, but because of its interior nature, is often relegated to the hidden curriculum. Current understanding of the clinical reasoning process is derived predominantly from low and high fidelity simulated cases, which do not reflect the numerous outside influences on reasoning.

Methods: Students on clinical internal medicine sub-internship rotations in a VA hospital setting were observed for the course of a complete patient admission, defined as starting at the time the resident notified the student of the incoming patient, through initial preliminary data collection, handoff from the emergency room physician, history and physical exam, initial note writing, initial order entry, and discussion of the case with the supervising resident. Observations were performed by a single observer, over the course of 4 months. Transcripts of the observations were analyzed, and interruptions were identified as any disruption in the student's work flow or thought process. Interruptions were categorized into domains according to their source and content. Supervising residents were surveyed separately to assess their entrustment of the student, limited to the domains in the History and Physical (H&P) and Differential Diagnosis entrusbable professional activity.

Results: The mean number of interruptions observed per observation was 27.8, with a range of 17-34. Observations ran an average of 2.5 hours and were evenly distributed between afternoon and evening hours. Supervising residents rated all students as entrusbable for H&P and Differential Diagnosis activities. The largest source of interruption to the students was the supervising resident (26 percent). Other major categories included social talk (14%) and phone calls to other students and house staff in the same room (13%). 8% of interruptions were generated by in-person visits from nurses or consultants and 6% of interruptions were due to phone calls to the student. See Figure 1.

Conclusion: Interruptions and distractions are omnipresent in the workflow of sub intern-level medical students. The supervising resident, while likely well-intentioned, represents the largest portion of interruptions observed, and may be limiting opportunities for the student to practice and develop their clinical reasoning acumen. Limitations of this study include a small sample size in a single clinical environment with a relatively low workload, with a single observer. Importantly, we are unable to determine which interruptions or distractions are necessary, and whether interruptions are helpful or harmful to the clinical reasoning process. Ongoing work on the content, structure and accuracy of the students' clinical reasoning is in progress.

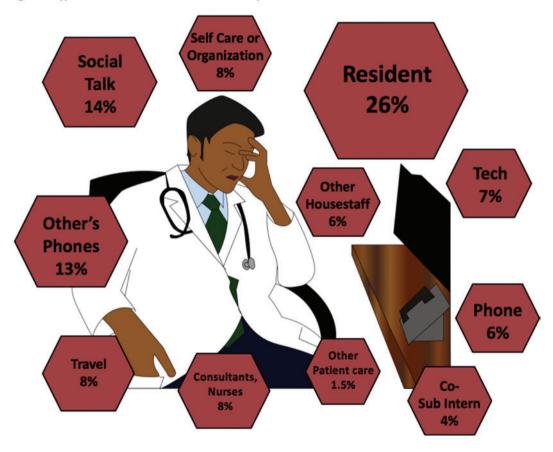


Figure 2: Types and Relative Contributions of Interruptions

#### Actual Diagnoses and Patient Outcomes in the Misdiagnosis of Inpatient Opioid Overdose

J. B. Cohen, D. Bruscino, D. Craig, K. Gonzalvo, P. Sahadeo, A. Sandberg, C. Vonnes, D. Wolf Moffitt Cancer Center, Tampa, FL

Background: Naloxone is often administered in situations of diagnostic uncertainty, even when the patient's clinical picture is not consistent with opioid overdose (i.e. pinpoint pupils, bradypnea, and lethargy). A chart review was conducted to determine the incidence, actual diagnoses, and outcomes of inpatient naloxone administration at an NCCN/NCI designated cancer center.

Methods: Over 36 months, cases were identified by automated dispensing cabinet reports indicating that naloxone was removed for administration. Charts were reviewed to confirm administration and to gather clinical details surrounding the event, along with interviews of staff members when appropriate. Naloxone was considered appropriate for administration if the patient met criteria set forth by the Society for Hospital Medicine, had immediate and sustained response to naloxone without the need for further diagnostic workup or treatment, had known opioid overdose, or had hypercarbia combined with opioid administration. If none of these criteria were met, the administration was determined to be inappropriate and the chart was reviewed to determine the actual diagnosis. All charts were reviewed to determine patient disposition at discharge.

Results: Naloxone use was determined to be inappropriate 52% of the time. The most common triggers for inappropriate administration were altered mental status (76%) and hypoxia (19%). The most frequent actual diagnoses for patients with altered mental status were SIRS, metabolic derangements, hypotension / hypovolemic shock, and polypharmacy. The most frequent actual diagnoses for patients with hypoxia

were pleural effusion and pulmonary embolism. Patients who received naloxone appropriately were more likely to be discharged to home and had a lower hospital mortality rate than those who received naloxone inappropriately (p <0.05, using the z-score test).

Conclusion: Often, the administration of naloxone is caused by uncertainty and the perception that it can be used as a benign diagnostic tool. Naloxone administration is not without the potential for harm. Its side effects can be life-threatening and its administration may result in diagnostic confusion, resulting in delays in establishing the actual diagnosis and initiating appropriate treatment. Understanding the common triggers for incorrect naloxone administration will assist in the development of diagnostic algorithms and clinical education to help guide providers to the correct diagnoses and prevent treatment delays. This is especially important, since the inappropriate use of naloxone is associated with a lower likelihood of discharge to home and a higher inpatient mortality rate.

#### "Mere Exposure" Bias: Comparing Cognitive Biases Reported in DEM Conferences and the Published Literature

S. Constantine<sup>1,2</sup>, M. Saijo<sup>1</sup>, U. Nakagawa<sup>1</sup>, M. Imamura<sup>1</sup> <sup>1</sup>Sapporo Tokushukai Hospital, Sapporo, Japan <sup>2</sup>Guam Regional Medical City Hospital, Dededo, Guam

Background: The mere exposure bias, also called familiarity principle, is "a psychological phenomenon by which people tend to develop a preference for things merely because they are familiar with them. We endevour to describe the types of cognitive biases reported in clinical vignetes during the Diagnostic Error in Medicine conferences and compare them with the type of cognitive biases vignetes published in the medical literature. Our hypothesis is that certain cognitive errors (anchoring, availability, etc) are more frequently reported and therefore, we are more "familiar" with them making them more easry to identify. Because of this "mere-exposure" bias, we might be missing the opportunity to find clinical examples of other more rare cognitive biases.

Methods: We will review all the case vignetes submitted to the past four Diagnostic Error in Medicine (DEM) conferences held in 2014, 2015, 2016 and 2017. We will review these cases and record the type of cognitive biases reported. In addition, we will perform a literature search of unique case reports of cognitive biases in Medline from 1/2015 to 12/2017 and will record the type of biases reported. We will finally compare the types of biases reported in both venues. We will not re-define the cognitive bias. We will use the type of bias as reported by the author. In case one case reports more than one type of cognitive bias, each different type will be recorded.

Results: We will present frequency tables of types of cognitive biases reported in DEM and in Medline.

Conclusion: Our conclusion will depend on our results. We hope this work will help people interested in "diagnostic errors" to identify other kinds of cognitive errors that might not be reported so frequently.

#### Discrete Choice Experiment to Assess Variation in Sepsis Definition

D. Woods<sup>1</sup>, J. B. Corboy<sup>2</sup>, D. Pinto<sup>3</sup> <sup>1</sup>Northwestern University, Feinberg School of Medicine, Chicago, IL <sup>2</sup>Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL

<sup>3</sup>Northwestern University Feinberg School of Medicine, Chicago, IL

Background: More than 75,000 children are treated in US hospitals each year for severe sepsis and sepsis-related complications. In 2012, The American Academy of Pediatrics (AAP) Septic Shock Collaborative developed a Triage Trigger Tool aimed at early recognition of pediatric sepsis which included several evidence-based clinical criteria for early sepsis recognition, yet the sepsis diagnosis continues to be missed leading to excess morbidity and continued mortality of 20-30%. We hypothesized that in the clinical setting, clinicians were attaching different levels of importance to these criteria that led to bias in the diagnostic process. To assess the relative importance of the AAP criteria as applied by pediatric emergency medicine physicians in the same institution, we conducted a study asking these clinicians to rank sets of these criteria for the sepsis diagnosis.

Methods: Discrete Choice Experiment model. Participants of different clinical stages were asked to rank combinations of criteria to assess an estimation of the relative importance of each of the AAP clinical sepsis criteria compared with each other

Results: Significant variation in the relative importance of clinical criteria was noted among the participating pediatric emergency medicine physicians.

- 50% of participants ranked capillary refill as the most important clinical factor determining the presence of sepsis
- 50% of participants attached no importance to growth and development (age of patient at presentation and developmental capabilities) while 25% of participants ranked this 5thof 9 positions and the remaining 25% ranked this 6thof 9 positions
- One physician ranked skin appearance and pulse rate as most important while capillary refill and temperature were the two most important criteria for another
  - When we review the mean criterion weights for this sample, the three most important criteria did NOT match with AAP findings and were:
- Mental status: 0.155
- Source of infection: 0.145
- Skin appearance/color: 0.141

Conclusion: Significant variation existed even in the same institution and the same service

Conclusion: Although these results are from the survey development phase, it illustrates discordance in clinicians' assessment of the diagnostic importance of these criteria. The results of this preliminary study uncover the variation in the estimation of the importance of each of these criteria for the diagnosis of sepsis. This variation can lead to bias in the diagnostic process where different pediatric emergency medicine physicians will focus on some criteria over others, where others may focus on different criteria and potentially miss a diagnosis of sepsis.

# Figure 1. DCE Spider Diagram

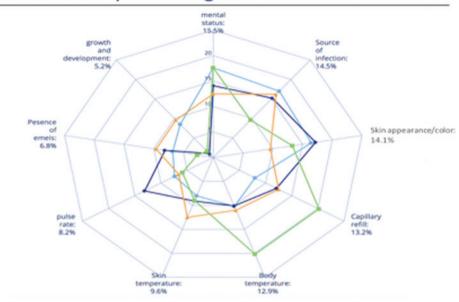


Figure 1 Provider-level preferences for criteria to diagnose sepsis

#### The Importance of Diagnostics in Cause of Death Reporting and Mortality Statistics: Alzheimer Disease As a Case Study

L. A. Flagg, R. Anderson National Center for Health Statistics, Hyattsville, MD

Background: Mortality data has many important uses, including tracking trends in life expectancy, death rates, and leading causes of death. However, the quality of mortality data is variable as it is subject to improvements in diagnostics and the level of training of medical certifiers. The change in the diagnostic guidelines for Alzheimer disease in 2011 is used as an illustration.

Methods: The 2011 and 2016 mortality data from the National Vital Statistics System were analyzed to determine age-adjusted mortality rates for Alzheimer disease, unspecified dementia, and all degenerative diseases of the nervous system.

Results: The age-adjusted mortality rate of Alzheimer disease was 24.73 per 100,000 in 2011 and increased to 30.29 per 100,000 in 2016, while unspecified dementia decreased from 32.31 per 100,000 in 2011 to 26.14 per 100,000 in 2016. The age-adjusted mortality rate of all degenerative diseases of the nervous system had a slight increase from 2011 to 2016 (61.75 per 100,000 and 64.94 per 100,000, respectively).

Conclusion: While it is possible that some of the increase in Alzheimer disease mortality was due to actual increases in the burden of this disease, the data suggest that these changes could be due to a shift from reporting unspecified dementia to Alzheimer disease as an underlying cause of death. These differences in cause of death reporting are likely related to the changes in the diagnostic criteria for Alzheimer disease. Resources and training in cause of death reporting newly available to medical certifiers are discussed.

#### Utility of Point-of-Care Ultrasound As a Tool for Medical Decision Making in a Resource-Limited Setting

M. Fleshner<sup>1</sup>, A. W. Fujita<sup>1</sup>, D. Bhamidipati<sup>1</sup>, S. Tilstra<sup>2</sup>, T. Bui<sup>2</sup> <sup>1</sup>University of Pittsburgh Medical Center, Pittsburgh, PA <sup>2</sup>University of Pittsburgh School of Medicine, Pittsburgh, PA

Background: Point-of-care ultrasound (PoCUS) is a quick, inexpensive, and noninvasive tool for bedside diagnosis that is most effective when used to answer a specific clinical question. Brief, intensive training is often the predominant educational model for PoCUS training.

Few studies have looked at the effect of PoCUS on clinical diagnosis in resource-limited settings. This study investigated the ability of PoCUS to assist with medical decision making when used by internal medicine trainees in Malawi.

Methods: This was a retrospective analysis of PoCUS studies conducted at Kamuzu Central Hospital in Lilongwe, Malawi through the University of Pittsburgh Medical Center (UPMC) Internal Medicine Residency program. Residents underwent a week-long intensive ultrasound training course using a SonoSite Nanomaxx cardiac probe, then used PoCUS as part of their routine clinical care of patients. Residents recorded each study as part of their training. Prior to performing the scan, residents documented their "pre-PoCUS" diagnosis and specific clinical question. Afterwards, they recorded their "post-PoCUS" diagnosis and the ability of PoCUS to answer the clinical question and prioritize the differential diagnosis. Change in management was a secondary outcome. Data were analyzed using Microsoft Excel and SAS.

Results: A total of 58 studies were performed by three residents over four weeks. Twenty-five (43%) were abdominal ultrasounds, 24 (41%) were cardiac, and 9 (16%) were pleural. Residents were able to obtain adequate images in 39 (68%) studies, incomplete images in 16 (28%) studies, and inadequate images in 3 (5%) studies. Ultrasound scans answered residents' clinical question in 46 (79%) studies, helped to prioritize the differential diagnosis in 51 (88%) studies, and changed the diagnosis in 21 (36%) studies. Notably, eight of the scans (67%) that did not answer the clinical question were still helpful in prioritizing the differential diagnosis. A higher percentage of cardiac (83%) and pleural (100%) scans answered the clinical question than abdominal scans (68%). Of the scans that changed the diagnosis, 12 (57%) were cardiac, 7 (33%) were abdominal, and 2 (9%) were pleural. Finally, eighteen (31%) ultrasound scans changed management, mostly by adding or removing medications.

Conclusion: PoCUS is a grossly underutilized imaging modality that has the potential to be an extremely useful tool for medical decision making in areas with limited resources. Cardiac and pleural scans contributed the most to the diagnostic process. PoCUS curricula for residents should focus on answering patient-specific clinical questions and working to improve diagnostic accuracy.

#### Measuring Diagnostic Errors in Psychiatry through Electronic Medical Record Reviews

T. Fletcher<sup>1</sup>, M. Stanley<sup>2</sup>, M. Kunik<sup>1</sup>, H. Singh<sup>3</sup>, A. Helm<sup>1</sup>, V. Vaghani<sup>3</sup>

<sup>1</sup>Michael E. DeBakey VA Medical Center, Houston, TX

<sup>2</sup>Baylor College of Medicine, Houston, TX

<sup>3</sup>Center for Innovations in Quality, Effectiveness and Safety, Michael E. DeBakey Veterans Affairs Medical Center and Baylor College of Medicine, Houston, TX

Background: Diagnostic errors in psychiatry are understudied and difficult to measure, particularly due to absence of objective diagnostic data such as labs and imaging. We adapted the Safer Dx Instrument, a structured tool to review electronic medical records for errors in medical diagnoses, to measure psychiatric diagnostic errors. Our specific aim was to adapt and test the Safer Dx Instrument to evaluate errors in anxiety diagnoses in a sample of Veterans diagnosed in primary care-mental health integration (PCMHI) settings as a first step to improve measurement of diagnostic errors in psychiatry. We focused on anxiety diagnostic error because unspecified anxiety disorder is the most common anxietyrelated diagnosis in VA PCMHI settings but <2% of Veterans diagnosed with unspecified anxiety actually meet diagnostic criteria for this disorder. Methods: The iterative adaption process included initial refinement of the Safer Dx Instrument questions by the research team, review by mental health providers to ensure content and face validity, and pilot testing of the revised instrument on a sample of 150 patient records. Patient records with an unspecified anxiety diagnosis (ICD-10 code F41.9) associated with a VA PCMHI appointment in FY16 were eligible for chart review. Two trained chart reviewers independently reviewed the first 30 records to determine the presence or absence of anxiety diagnostic errors at the index PCMHI visit and subsequent visits in the following year. Chart reviewers discussed the first 30 cases with the research team to develop a rule book for defining anxiety diagnostic errors and to calibrate procedures. Following calibration, one chart reviewer independently reviewed the remaining 120 records. Records with a coexistent specific anxiety diagnosis were excluded (n = 47), leaving a final sample of 103 records. The second chart reviewer independently reviewed 20% (n = 16) of cases included following calibration (n = 80) for interrater reliability assessment.

Results: Of 103 records that met inclusion criteria, 62 were determined to involve a diagnostic error (42 errors due to use of unspecified anxiety diagnosis when a specific anxiety diagnosis was warranted; 20 errors due to use of unspecified anxiety diagnosis when anxiety symptoms were either not documented, or were documented but not severe enough to warrant diagnosis. Reviewer agreement on presence/absence of errors was 88% ( $\kappa = 0.71$ ) following calibration.

Conclusion: The revised 'Safer Dx - Mental Health Instrument' has high reliability for detecting anxiety-related diagnostic errors. Future studies should test this instrument in additional psychiatric populations and clinical settings.

#### **Emergency Medicine Physician Perceptions of Feedback on the Diagnostic Process**

J. W. Fox, S. Cohen, R. Ruf, A. Naughton Duke University Health System, Durham, NC

Background: Everyone in the United States will experience a diagnostic error during their lifetime. Despite the fundamental importance of accurate diagnoses in health care and clear opportunities for improvement, a systematic process to provide physicians with feedback on their diagnostic accuracy is lacking. Little is known about emergency medicine (EM) physician perceptions of their diagnostic accuracy. We surveyed EM residents and faculty to improve our understanding of current practices to assess personal diagnostic accuracy and perceptions on implementing a systematic process to provide this feedback.

Methods: EM residents and pediatric emergency department attending physicians were asked to complete an online Qualtrics® survey. Participation was solicited by in-person and email invitation and was voluntary. Descriptive statistics were used for analysis.

Results: Twenty-two EM residents (73% response) and 23 attendings (77% response) completed the survey. Nearly all EM residents (91%) and attending physicians (96%) question their diagnostic accuracy at least monthly, with 55% and 30%, respectively, expressing concern multiple times per week. Alarmingly, 36% of residents had never received diagnostic feedback and an additional 55% indicated they received this type of feedback only a few times per year. All respondents reported reviewing patient records after providing care. However, the frequency of this effort did not match the frequency of their concern for diagnostic accuracy (data not shown). When physicians did review patient records to determine the ultimate diagnosis, they were frequently surprised (EM residents = 30% [range 5-68%], attendings 27% [range 9-100%]). The vast majority of resident (95%) and attending physicians (96%) reported that they would like to have a system in place that provides feedback on diagnostic accuracy. One respondent remarked, "this is incredibly important for our development as physicians and I can't believe this isn't already standard practice." The primary barrier cited to use of a new feedback system relate to incorporation into an already busy workflow. Importantly, 18% of resident and 25% of attending physicians noted that they would be uncomfortable giving or receiving feedback related to the diagnostic process.

Conclusion: Physicians who care for acutely ill and injured children frequently question their diagnostic accuracy, yet infrequently receive important feedback to calibrate their diagnostic impressions. Participants would welcome a systematic process to provide feedback on diagnostic accuracy but worry about workflow processes and discomfort around communication of this feedback. Based on these results, our team is currently developing a system to address this educational gap while minding perceived barriers.

#### Close the Loop: A Pilot Study of Diagnostic Feedback Among Pediatric Physicians

J. W. Fox, S. Cohen, R. Ruf, A. Naughton Duke University Health System, Durham, NC

Background: Closed-loop communication among clinical care teams is integral to patient safety. This type of communication occurs routinely when critical lab results are found and when adjustments to radiology interpretations are made. Surprisingly, a systematic process to provide closed-loop communication among providers regarding the diagnostic process is lacking. This feedback is critical to improve physicians' diagnostic calibration, thereby reducing diagnostic errors. Here we report on our pilot project to develop closed-loop communication among emergency department (ED) and inpatient pediatric providers.

Methods: Duke Health Technology Services developed a data element that physicians could use to "tag" patient charts by inserting the element into any patient note. We encouraged providers to use this tool when physicians faced a high degree of diagnostic uncertainty, a change in diagnosis occured, or patients' clinical courses were substantially different than expected. Charts tagged with our data element were identified by a customized report, and then reviewed by our research team. Secure emails were sent to notify providers of these educational "Close the Loop" (CTL) cases and encourage them to review the record for a learning opportunity. This pilot focused on patients admitted from the Pediatric ED to General Pediatric inpatient teams. As a quality control measure, all such admissions were reviewed to identify cases that met criteria for CTL but were not tagged with the data element.

Results: In 6 months, 95 patients were tagged for CTL feedback. Of these, 88% were identified by an ED provider. Our research team identified an additional 69 cases that met CTL criteria. The most common presentations tagged for CTL feedback were respiratory distress or failure (18%) and fever without identified source (17%). Focal neurologic deficits were more commonly identified as appropriate for CTL feedback on chart review (9%) than by the clinical team (2%). The final diagnosis of ~50% of cases identified by each method was infectious in nature. Neurologic diagnoses were also common, representing about 20% of the cases tagged for CTL feedback or identified by chart review.

Conclusion: ED physicians utilized our diagnostic feedback system far more than inpatient pediatric providers, likely due to the desire for resolution of diagnostic uncertainty. Respiratory symptoms, infectious diagnoses, and neurologic conditions were the most common reasons for CTL feedback. Future research will seek to discover barriers to utilization of this feedback process by inpatient physicians and the educational benefit to CTL users.

#### Learning from Patient Experiences Related to Diagnostic Errors Is Essential for Improving Patient Safety

T. Giardina<sup>1</sup>, H. Haskell<sup>2</sup>, S. Menon<sup>1</sup>, J. Hallisy<sup>3</sup>, F. Southwick<sup>4</sup>, U. Sarkar<sup>5</sup>, K. Royse<sup>1</sup>, H. Singh<sup>1</sup>

<sup>1</sup>Center for Innovations in Quality, Effectiveness and Safety, Michael E. DeBakey Veterans Affairs Medical Center and Baylor College of Medicine, Houston, TX

<sup>2</sup>Mothers Against Medical Error, Columbia, SC

<sup>3</sup>The Empowered Patient Coalition, San Francisco, CA

<sup>4</sup>Univeristy of Florida Health, Gainesville, FL

<sup>5</sup>University of California San Francisco, San Francisco, CA

Background: Diagnostic error research has largely focused on clinicians' individual decision-making and system redesign while overlooking potentially more complex behavioral and interpersonal factors. We conducted a secondary data analysis of a unique new data source of patient- and family-reported error narratives to explore behavioral and interpersonal factors that contribute to diagnostic errors.

Methods: Our study sample included 887 patients- and family-reported adverse medical events submitted to a survey-based national online reporting platform housed by the Empowered Patient Coalition between January 2010 and February 2016. We selected self-reported diagnostic errors where respondents included written narratives of their experiences (n = 465). Two authors reviewed these narratives to determine presence/absence of diagnostic error using the definition proposed by the National Academy of Medicine. Using a thematic analysis approach, narratives were coded independently, merged, and discussed to identify major themes. Major themes were then further refined by two additional authors, both patient advocates with expertise in patient safety.

Results: We included 184 unique patient narratives of diagnostic error; majority were female respondents (65.8%) and average age was 52.4 years (range 2-98 years, SD19.7). Most (80.4%) involved the hospital setting (inpatient or emergency room). Nearly three-fourths (n=134, 72.8%) included discussion of problematic clinician behavior, (i.e., behaviors not consistent with patient-centered care) related to the diagnostic process. Of these we identified 224 such instances involving all types of clinicians: physicians, nurses, and physician assistants. Behaviors were grouped into four categories: ignoring patient knowledge (n=92), disrespecting patients (n=63), failing to communicate (n=54), and manipulation or deception (n=15).

Conclusion: Using a unique new data source of patient- and family-reported diagnostic errors, we found that unprofessional clinician behavior appears to contribute to diagnostic errors much more commonly than previously considered. While some of this information could be gleaned from passively looking through patient-generated complaints, health systems should use more proactive approaches to identify unprofessional clinician behavior that harms patients. Findings call for better understanding of the association between clinician behaviors and diagnostic errors and development of strategies for mitigation. Health systems should develop and implement formal programs to collect patient experiences with the diagnostic process, including interpersonal and behavioral issues, and use this data to promote organizational learning and a culture of effective communication and safe diagnosis. These programs could facilitate more equitable relationships and positive patient-centered communication behaviors that engage patients, families, and caregivers in the diagnostic process and help prevent diagnostic errors.

Table 1	1 · Case	examples	for each	theme
rable	i: Case	examples	ior each	meme.

Table 1. Case examples for	tach titum.
Ignoring patient knowledge	A patient went to see his doctor for rectal bleeding and was diagnosed with an internal hemorrhoid. Feeling that his concern was not taken seriously and not being appropriately diagnosed, he shared his symptoms with a physician in an unrelated specialty. Concerned, the physician contacted the original physician but was also disregarded, according to the patient. "I went to the [clinic] and told them I had rectal bleeding and they blew me off! Found out what a colonoscopy was and demanded the [clinic] give me one" (P 132) He was diagnosed with stage 3 colon cancer. In the end, this participant believed, due to the way he was treated, that the clinic "hoped that I would die." While this may have been an extreme case, expression of loss of trust in the healthcare system was common among patients and family members who experienced a diagnostic error and attributed it to being ignored.
Disrespecting patients	Age related bias, though less common in this sample, was particularly painful for family members. Participants articulated a common refrain: they felt that their family member's life was devalued because the patient was considered old and likely to die anyway. "Total feeling of failure due to inability to make the physicians listen to my father's screams of pain and need to obtain help in finding out the cause of the pain instead of writing him off due to age Yes, he was 86 years old, yes, he was on dialysis - but he was penalized because of his age and was misdiagnosed because of it." (P 204)
Failure to communicate	A patient underwent a hysterectomy, and was experiencing severe pain following the surgery. She contacted her surgeon with complaints of "pressure, urinary problems, pain, stomach distention. All complaints were reasoned away." Following complaint of severe flank pain, an ultrasound revealed fluid in her stomach. "For an additional nine days, I was in excruciating pain.  Narcotics kept being ordered. My surgeon never saw me or called me during the nine days of hell and numerous calls day and night to their office. I was admitted for pain on a Friday night." (P 190) Following hospitalization, it was determined that her ureter had been cauterized during surgery and had gone undiagnosed for over a week.
Manipulation	Telling a family member that a consultant was too busy to see the patient when the consult had simply not been placed in a timely manner. "I asked on day 29 why had there not been a PT [physical therapy] consult. MD told me they were too busy there. Day thirty PT and OT [occupational therapy] come in, saying they had just got the consult, and asked could she do a physical exam and she started at my mom's head, got to her left shoulder a couple minutes later asked, 'how long has this shoulder been dislocated?" (P 213)

#### Exploring the Use of Patient Experience Surveys As Indicators of Diagnostic Safety

H. Singh¹, T. Giardina¹, D. K. Upadhyay², S. Korukonda², J. Hunt², G. F. Burke², D. Torretti²

<sup>1</sup>Center for Innovations in Quality, Effectiveness and Safety, Michael E. DeBakey Veterans Affairs Medical Center and Baylor College of Medicine, Houston, TX

<sup>2</sup>Geisinger Medical Center, Danville, PA

Background: Patients' feedback is increasingly used to identify quality of care but the potential of using patient experience surveys to monitor and improve diagnostic errors remains unexplored. We tested if routinely collected patient experience data can provide signals suggestive of diagnostic error in primary care settings.

Methods: At a large tertiary care health system, we analyzed patient responses to surveys after primary care visits during July 2013-2017. We hypothesized that certain signals in patient-generated data can help health systems identify diagnostic safety issues and develop methods akin to that of diagnostic "trigger" tools, i.e., algorithms that scan thousands of electronic health records to flag those with clues suggestive of diagnostic errors and care delays. A multidisciplinary team identified six questions relevant to the diagnostic process from Press Ganey's Medical Practice survey. We identified a trigger positive cohort as one where there were low responses in two or more survey questions of relevance (low response was defined as worst possible outcomes for the question, e.g. 1 and 2 on a scale of 1-5). We conducted medical record reviews of all associated clinic and phone visit records and identified missed opportunities in making a timely and correct diagnosis in the corresponding patient-physician encounters of interest. Reviewers also identified other types of management errors which led to patient dissatisfaction. We calculated positive predictive values (PPV) for detection of diagnostic errors and management errors.

Results: Of 512,198 surveys returned for office visits in the study period for primary care, 3982 were trigger-positive. Of these 615 had consistently low responses for all 6 questions of relevance. Of 3982 trigger-positive records, we thus far manually reviewed a random sample of 200 and identified a total of 18 (PPV = 9%) records with errors: 8 diagnostic errors (PPV = 4%) and 10 management errors (PPV = 5%). Four of eight diagnostic errors involved missed infections, including Lyme disease, otitis media, pneumonia and cellulitis. Management errors were related to wrong treatment or medication.

**Conclusion:** Use of existing patient experience surveys provides a modest signal to identify both diagnostic errors and management errors. However, combining this signal with additional diagnostic data could increase signal intensity. For example, data related to the patient visit of interest (such as diagnostic visit codes for possible infection) could help identify specific types of errors (such as missed infections). With further refinement and validation, health systems could leverage these data for patient safety measurement and improvement.

#### Comparison of Pediatrician Comfort Discussing System Versus Diagnostic Errors in Morbidity and Mortality Conference

J. A. Grubenhoff<sup>1,2</sup>, S. I. Ziniel<sup>1,2</sup>, G. Singhal<sup>3,4</sup>, T. L. Cifra<sup>5</sup>, R. E. McClead Jr.<sup>6</sup>, H. Singh<sup>7</sup>

<sup>1</sup>University of Colorado School of Medicine, Aurora, CO

<sup>2</sup>Children's Hospital Colorado, Aurora, CO

<sup>3</sup>Baylor College of Medicine, Houston, TX

<sup>4</sup>Texas Children's Hospital, Houston, TX

<sup>5</sup>University of Iowa, Iowa City, IA

<sup>6</sup>Nationwide Children's Hospital, Columbus, OH

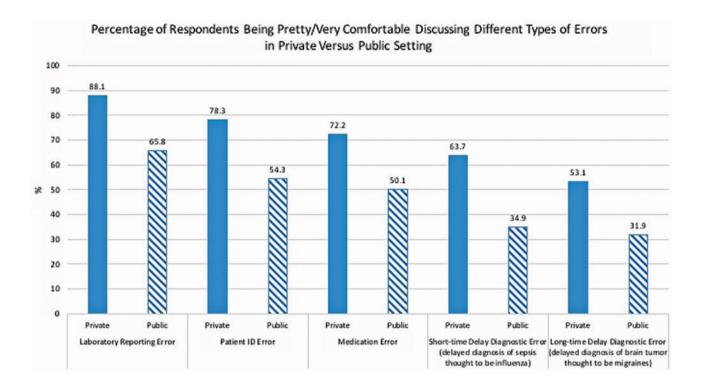
<sup>7</sup>Center for Innovations in Quality, Effectiveness and Safety, Michael E. DeBakey Veterans Affairs Medical Center and Baylor College of Medicine, Houston, TX

Background: Improving Diagnosis in Healthcare suggests promoting "a non-punitive culture that values open discussion and feedback on diagnostic performance". A "just culture" could help balance competing demands of acknowledging personal accountability against creating a safe venue to discuss and learn from diagnostic errors. Little is known about pediatricians' perceptions of error discussion culture within and outside of quality improvement settings such as Morbidity and Mortality Conferences (M&M). We sought to explore pediatricians' comfort in discussing different types of errors, including diagnostic errors, in private and public forums as well as barriers to participating in those discussions.

Methods: We developed and administered a survey to pediatricians, advanced practice providers (APPs) and graduate trainees in general pediatrics and pediatric subspecialties at four U.S. university-affiliated children's hospitals. Survey questions assessed their comfort in discussing situations of system-attributable errors (related to laboratory reporting and patient misidentification), medication errors and primarily cognition-attributable diagnostic errors in two venues respectively: private (with trusted colleagues) and public (M&M). Respondents identified the most significant barriers to discussing their own diagnostic errors in such venues.

Results: 838 participants completed the survey (22% response rate; range 10-35%), of which 57% were physicians, 22% APPs, 13% trainees and 8% declined to identify. As expected, significantly fewer respondents indicated being "pretty comfortable" or "very comfortable" discussing different error type in public compared to private settings (Figure). However, there were significant institutional differences in comfort level regarding discussion of all types of error in public settings. Furthermore, in both venues, comfort level for discussing diagnostic errors was lower than all other error types. s cited poor perception of oneself as a clinician (63%), peer judgment of decision-making (51%), knowledge-base (47%) and loss of reputation (47%) as most significant barriers to discussing one's own diagnostic errors. Concerns about job loss (10%), malpractice risk (11%) and peer perception of work ethic (5%) were cited infrequently. Only 15% reported no barriers to discussing diagnostic errors.

Conclusion: Pediatric providers are far less comfortable discussing cognition-attributable diagnostic errors compared to other system-attributable errors. Building and promoting a local institutional climate supporting open discussion and learning could help reduce preventable diagnostic harm.



#### **Understanding Diagnostic Reasoning Using a Case-Based Approach**

A. Gupta<sup>1</sup>, M. Quinn<sup>2</sup>, R. Lewis<sup>2</sup>, S. Winter<sup>2</sup>, S. Saint<sup>1</sup>, V. Chopra<sup>1</sup> VA Ann Arbor Healthcare System, Ann Arbor, MI <sup>2</sup>University of Michigan, Ann Arbor, MI

**Background:** While understanding the cognitive contributions to diagnosis has become increasingly important, few studies have aimed to understand physician diagnostic reasoning in real-time. Using a case-based approach, we sought to: (a) characterize differences in diagnostic reasoning among hospitalists; and (b) identify factors associated with diagnostic accuracy.

**Methods:** A total of 10 hospital medicine attending physicians from an academic medical center took part in the case-based interview. Hospitalists were provided pertinent information regarding a patient's presentation in a step-wise fashion (**Figure**), and asked to formulate a differential diagnosis following each new piece of key data. Following presentation of basic labs, participants were able to request results for additional studies, which were provided if performed. An updated clinical course was also provided to participants regularly. All sessions were audio-recorded and transcribed verbatim. Synthesis of quantitative and qualitative data using mixed methods were used to identify key findings by two reviewers (AG, VC).

**Results:** Ten hospitalists (5 male and 5 female) participated in interviews. The average years post-residency was 7.2 (range 2-16 years. The correct diagnosis (shingles) was included in the final differential of 4 of the 10 participants. Hospitalists that identified the correct answer differed from those that did not in several ways. First, hospitalists that included the correct diagnosis included more diagnoses in their initial differentials (mean 9.75 [range 8-13, SD 1.18] vs 6.0 [range 4.0-10.0, SD 0.89], p=0.03). Second, those that included the correct diagnosis also used a specific organizational structure to generate candidate diagnoses, including *systems-based* (i.e., cardiovascular, pulmonary) and *anatomical* (i.e., chest, abdomen) approaches. Third, while all hospitalists prioritized life-threatening diagnoses, those whose initial differential included only those causes were less likely to arrive at the correct diagnosis and more likely to demonstrate anchoring further into the case. Finally, those correctly identifying the diagnosis frequently re-examined data regarding the patient's initial presentation or physical examination.

**Conclusion:** Although limited to a small sample size from one academic medical center, particular behaviors in diagnostic evaluation appear to limit anchoring and improve diagnostic accuracy. Frequent and intentional re-evaluation of data in the context of the patient's initial presentation are likely important in refining differential diagnoses.

Figure: Flow diagram of case-scenario progression (received by all participants)

Chief Complaint: 66yoM with history of tobacco abuse presents with chest pain **History of Present Illness:** Burning over L chest starting last evening, developing into dull ache this AM prior to ED (+) radiation to R arm, constant with intermittent flares, dyspnea associated with flare, pleuritic discomfort. (-) nausea, palpitations, lightheadedness, dizziness. Patient has not been active, unclear exertional component. Past Medical/Surgical History: Hypertension, tobacco abuse Medications: Metoprolol 25mg twice daily Family History: Father with MI at age 55, mother with diabetes, hypertension Social History: Lives with wife and 2 daughters; smokes 1 pack/day x 35 years, Denies alcohol or illicit drug use; retired autoworker Physical Exam (Pertinent Exam Findings) Vitals: Temperature 100.1F Pulse 109 beats per minute Blood Pressure 151/95 Respiratory Rate 18 Pulse Oximetry 95% room air Chest: mild tenderness to palpation over left chest Cardiovascular: tachycardic, regular, II/VI systolic murmur best at left upper sternal border Extremities: trace lower extremity edema bilaterally Complete blood count: WBC 9.0 x10^3 g/microliter, Hgb 14.5 g/dL, Platelet 376 x10^3 Basic Metabolic Panel: Sodium 135 mmol/L, potassium 3.6 mmol/L, chloride 110 mmol/L bicarbonate 24 mmol/L, BUN 14 mg/dL, creatinine 1.1 mg/dL, glucose 147 mg/dL Liver Function Tests: Protein 8.6, Albumin 3.9, AST 16 units/L, ALT 22 units/L, Alkaline phosphatase 90 units/L, total bilirubin 0.5 mg/dL

#### **Understanding Systems-Based and Cognitive Contributions to Diagnosis**

A. Gupta<sup>1</sup>, M. Harrod<sup>1</sup>, M. Quinn<sup>2</sup>, K. Fowler<sup>1</sup>, S. Saint<sup>1</sup>, V. Chopra<sup>1</sup> <sup>1</sup>VA Ann Arbor Healthcare System, Ann Arbor, MI <sup>2</sup>University of Michigan, Ann Arbor, MI

Background: Diagnostic error in medicine is increasingly recognized as "the next frontier for patient safety". Current research has focused on diagnostic errors in two unique dimensions: as systems-based or cognitive-based etiologies. One study, however, suggests that nearly half of all errors stem from both domains. Using a focused ethnographic approach, we sought to understand how systems-based factors contribute to the cognitive missteps underlying diagnostic error.

Methods: We conducted a focused ethnography of inpatient medicine teaching teams at two affiliated academic medical centers. Teams (consisting of 1 attending, 1 senior resident, 2-3 interns, and 1-4 medical students) were observed on rounds and during post-round work. Field notes related to the diagnostic process as well as the work system were recorded and collated. Observations were followed by focus groups with interns and residents and interviews with attending physicians, to better understand perceptions around diagnosis and factors contributing to error. Inductive analysis was used to organize findings into themes, which were then categorized according to a previously described taxonomy into systems-factors contributing to the cognitive domains of faulty data gathering, faulty information processing, and faulty data verification. An additional category describing work environment factors encompassing all of these cognitive domains was included.

Results: Between January 2016 and April 2017, 10 teams were observed over the course of 286 hours. All 10 attendings and 31 interns and residents participated in interviews or focus groups. The following key themes that bridged system and cognitive factors were identified. First, challenges with interdisciplinary communication, particularly in-person communication, as well as closed-loop communication within the electronic medical record (EMR) were felt to contribute to faulty data gathering. Second, information contained within the EMR was often considered unreliable and inaccessible due to poor integration and interface. Third, organizational structures such as the operation of consulting services in silos and the reliance on inexperienced overnight radiology interpretations were felt to contribute to faulty information processing. Fourth, frequent care handoffs, between inpatient providers as well as the emergency department and the inpatient ward, were noted to provide opportunities for faulty data verification. Finally, interruptions, time constraints, and a cluttered physical environment were found to adversely impact all cognitive domains.

Conclusion: Systems-based and cognitive factors likely occur simultaneously in diagnostic errors. Interventions aimed at addressing cognitive contributions should be coupled with systems-based interventions in order to maximize effectiveness.

#### Changes in Symptom Sensitivity and Specificity When Other Symptoms Are Already Known: Exceptions to the Simple Application of Baves' Theorem.

J. Van Den Ende<sup>1</sup>, E. Bottieau<sup>1</sup>, R. M. Hamm<sup>2</sup> <sup>1</sup>Institute of Tropical Medicine, Antwerp, Belgium <sup>2</sup>University of Oklahoma Health Sciences Center, Oklahoma City, OK

**Background:** The Bayesian ideal diagnostic process is often described as the adjustment of disease probabilities when additional findings are observed. Adjustment, one finding at a time, is based on the findings' known sensitivities and specificities. However, the sensitivity and specificity of clinical signs and symptoms may change depending on findings previously observed. It is straightforward to calculate changes in specificity: as accumulating findings change the relative probabilities of the "other" diseases, the overall probability of a negative observation of the symptom drifts. But there is no formula for calculating changing sensitivity – empirical data are needed. Using a large patient data base, we inspected how the diagnostic characteristics of one symptom change depending on what is known about another symptom.

Methods: A previously assembled data base had clinical symptoms and final diagnoses for 2249 patients with fever arriving in Antwerp from tropical travel. For 4 diseases (malaria, dengue, pneumonia, Rickettsia Africae) we selected pairs of symptoms (from 6 to 10 pairs per disease). For each symptom in each pair, we calculated sensitivity and specificity if the other symptom was known positive, known negative, or not yet known. We also observed how each symptom's likelihood ratios (LR+, LR-) changed depending on the knowledge of the other symptom. Results: The table illustrates, for malaria, how the diagnostic characteristics of reported vomiting vary, as a function of whether patient has jaundice. It can be seen that the sensitivity of vomiting increases (waxes), while its specificity wanes, when jaundice is positive, and the opposite when jaundice is negative. The LRs change, as well.

Vomiting	Sensitivity	Specificity	LR+	LR-
Jaundice unknown	.34	.86	2.43	0.77
Jaundice positive	.60	.41	1.03	0.96
Jaundice negative	.30	.87	2.29	0.80

Generally, this pattern of waxing sensitivity and waning specificity is observed when symptoms are positively correlated, and is larger when the correlation is larger. However, the impact on the likelihood ratios varies. Waning specificity can compensate for waxing sensitivity, leaving the LR somewhat stable; but the direction of change is unique for each symptom pair.

Conclusion: Clinicians and analysts should keep in mind the potential impact of dynamic sensitivity and specificity. While the changes in the basic test characteristics were relatively small, when ratios are made (LRs) the effects can be large. That a finding's sensitivity and (1 - specificity) move in the same direction as other symptoms are observed does not guarantee that the LRs will remain stable.

#### Automatically Identifying Sepsis in Retrospective Data with a Clinical Phenotyping Algorithm

K. E. Henry<sup>1</sup>, H. Soleimani<sup>1</sup>, N. Rawat<sup>2</sup>, S. Saria<sup>3,4</sup> <sup>1</sup>Johns Hopkins University, Baltimore, MD <sup>2</sup>Armstrong Institute for Patient Safety and Quality <sup>3</sup>Johns Hopkins University <sup>4</sup>Johns Hopkins School of Public Health

Background: Sepsis is an out-of-control infectious response leading to organ dysfunction and has high mortality and morbidity. An enduring challenge in sepsis research is how to reliably identify sepsis from retrospective data. This is critical for assessing the burden of diagnostic error in sepsis and developing tools to improve identification. Billing codes or the presence of systemic inflammatory response syndrome (SIRS) and signs of organ dysfunction are commonly used to identify sepsis. However, both approaches have severe limitations. Billing codes have been repeatedly shown to have low sensitivity. SIRS is often criticized for having low specificity and some studies suggest SIRS also misses true sepsis cases. To reliably identify sepsis from retrospective data, we automatically learn sepsis phenotypes from retrospective electronic health record (EHR) data.

Methods: Data were collected from 170,472 adult patients admitted through the emergency department at a community hospital between April 2014 - April 2017. We used a machine learning methodology to develop a clinical phenotyping algorithm that identifies sepsis using laboratory values, vital signs, medications, procedures, and patient demographics and history. The phenotyping algorithm was applied to new data collected from the same hospital between November 15, 2017 - May 15, 2018. Physicians randomly evaluated 469 of the identified sepsis cases and were asked to evaluate presence of infection and sepsis based on the patient's EHR. Patients with chronic pancreatitis were excluded due to known challenges automatically identifying sepsis in this population.

Results: Of the cases evaluated, 81.7% were confirmed to be sepsis. An additional 2.8% were confirmed to have infection, but did not present sufficient signs of organ dysfunction. Physicians indicated that 15.6% had no infection. Of the confirmed cases, 25.8% had a billing code for severe sepsis, 37.3% did not have SIRS at or prior to the time of sepsis onset, and 10.7% of them never developed SIRS at any point in their hospital stay (Table 1).

Conclusion: A clinical phenotyping algorithm that takes into account diverse clinical data sources identifies sepsis with higher reliability. Despite being known for low specificity, we found that over a third of cases identified by the phenotyping algorithm did not meet SIRS criteria. Consistent with previous work, we found that fewer than half of confirmed sepsis cases had a sepsis-related code. Studies investigating the impact of sepsis on diagnostic error or proposing new sepsis-identification methodologies should not rely solely on billing codes or SIRSdependent methods to identify sepsis.

	Number of encounters (%)
Labeled severe sepsis cases evaluated by provider	469
No infection	73 (15.6%)
Infection, but no related organ dysfunction	13 (2.8%)
Confirmed to be severe sepsis	383 (81.7%)
Of the 383 provider-confirmed sepsis cases	s:
Have a severe sepsis ICD-9/10 code	99 (25.8%)
Have a sepsis (without organ dysfunction) ICD-9/10 code	117 (30.5%)
Have no sepsis-related ICD-9/10 code	224 (58.5%)
Never have SIRS	41 (10.7%)
Never have SIRS at or prior to sepsis onset	229 (37.3%)

Table 1: Evaluation of sepsis labels generated by clinical phenotyping algorithm

#### Influence of Comorbid Depression on Diagnostic Accuracy in Physical Illness: A Randomized Experiment

L. M. Isbell<sup>1</sup>, M. Graber<sup>2</sup>, D. R. Rovenpor<sup>3</sup>, G. Liu<sup>1</sup>, D. J. Pallin<sup>4</sup> <sup>1</sup>University of Massachusetts Amherst, Amherst, MA <sup>2</sup>RTI International, Research Triangle Park, NC

Background: Evidence suggests that everyone will experience at least one diagnostic error in their lifetime (IOM, 2015), with some likely to experience significantly more. One population at particular risk are those with a history of mental illness, a highly-stigmatized group that routinely evokes negative reactions (Hinshaw & Stier, 2008). Patients with mental illness are less likely to receive the same physical health care as those without mental illness, and are less likely to be treated in accordance with clinically-established guidelines (McGinty et al., 2015). These disparities have been identified largely on the basis of retrospective chart reviews. Consequently, the causes of these disparities remain unclear. The current investigation employed a randomized experiment to investigate the causal influence of co-morbid mental illness upon diagnosis. Methods: We presented physicians with a vignette describing a patient with a complex presentation due to pernicious anemia. Physicians were randomized to one of two conditions: the patient had co-morbid depression with a remote psychiatric hospitalization, or the patient had no co-morbid psychiatric illness. All other clinical information was the same. After reviewing as much case information as they desired, physicians recorded a differential diagnosis and indicated what additional diagnostic tests they would order, if any. Physicians also indicated the extent to which interacting with the patient would lead them to rate her as likeable, happy, and depressed. Results: Sixty-one physicians completed the experiment. The patient with depression was rated as significantly less likeable, less happy, and more depressed. Physicians ordered a similar number of diagnostic tests in both conditions and arrived at a correct diagnosis in approximately half of the cases. Co-morbid depression was associated with a statistically non-significant *tendency* toward less accurate diagnosis (p = .15). However, the number of tests ordered interacted with patient co-morbid depression to significantly predict diagnostic accuracy (p=.03). For

Ohio State University

<sup>&</sup>lt;sup>4</sup>Brigham and Women's Hospital

the non-depressed patient, there was no relationship between test ordering and diagnostic accuracy (p>.80), but for the depressed patient, there was a significant positive relationship (p=.01). The difference between the depressed and non-depressed patients was particularly large among physicians who ordered few tests. Conclusion: We found that the extent of the diagnostic workup was associated with diagnostic accuracy for the patient with depression, but not for the non-depressed patient. This might be explained by a tendency of some physicians to ascribe symptoms to psychiatric disease, leading them to order fewer tests and think less carefully in the setting of co-morbid depression.

#### Exploring How Internal Medicine Residents and Attending Physicians Perceive and Behave When Faced with Uncertainty in the Gme Learning Environment at a Single US Academic Medical Center

M. W. Johnson<sup>1,2</sup>, G. Gheihman<sup>2</sup>, H. Thomas<sup>3</sup>, A. Simpkin<sup>4</sup>, G. D. Schiff<sup>5</sup>, A. P. J. Olson<sup>6</sup>

<sup>1</sup>Mount Auburn Hospital, Cambridge, MA

<sup>2</sup>Harvard Medical School, Boston, MA

<sup>3</sup>Kaiser Permanente Oakland, Oakland, CA

<sup>4</sup>Massachusettes General Hospital, Boston, MA

<sup>5</sup>Brigham and Women's Hospital, Boston, MA

<sup>6</sup>University of Minnesota, Minneapolis, MN

Background: William Osler is frequently quoted describing medical practice as "a science of uncertainty, and an art of probability." In today's practice environments this remains true, yet the breadth and depth of medical science has immensely increased; there is even public perception that uncertainty may be eliminated. However, physicians must be able to navigate diagnostic and therapeutic decision-making in the face of ubiquitous uncertainties. Prior research correlates uncertainty to pitfalls of ineffective communication strategies, cognitive bias, and inappropriate resource use. A rigorous understanding of how uncertainty is manifest may improve the current model through which patients and physicians acknowledge, manage, and cope with uncertainty in healthcare and disease processes. Methods: The present analysis is a mixed-method study investigating perceptions and manifestations of uncertainty in the clinical environment within the department of medicine at a single US academic medical center. Three separate methodologies, one quantitative (survey) and two qualitative (observations and semi-structured interviews), were used to rigorously triangulate the nature of uncertainty in this clinical-educational setting. Results: Comparison of responses (two-tailed t-test) of validated survey instruments demonstrated resident physicians (n = 35) have more perceived stress from uncertainty than attending physicians (n = 14) (15.6 [5.0] vs. 19.0 [3.6], p < 0.05), and significantly higher symptoms of burnout (7.1 [2.9] vs. 4.9 [2.0], p < 0.05). Regression analyses suggest that perceiving the learning environment as more competitive and stressful strongly correlates with an increased reluctance to disclose uncertainty (r = -0.44, p < 0.01). This remains true for residents and attending physicians (r = -0.41, p = 0.01 and r = -0.67, p = 0.01, respectively). Qualitative analysis of semi-structured interviews demonstrates 1) a desire to communicate uncertainty to patients, 2) the influence of the authority, 3) role modeling as an educational tool, and that 4) embracing uncertainty is not a part of the formal curriculum. **Conclusion**: Uncertainty is an inherent part of medicine not adequately addressed. In the practice of hospital medicine uncertainty is common and the ways in which it is emotionally perceived differ by the role and level of training of the provider. Observations of clinician behavior demonstrate the importance of the attending physician in introducing, or role modeling, language for uncertainty especially in the patient-physician encounter. It seems essential for the medical education system to develop educational strategies to explicitly reference and embrace uncertainty to benefit physician wellness and promote high quality therapeutic relationships with patients.

#### **Measuring Diagnostic Error: A Review of Patient Complaints**

M. H. Kanter<sup>1,2,3</sup>, L. D. Lurvey<sup>3</sup>, K. Litman<sup>3</sup>, R. Loo<sup>3</sup>, W. Strull<sup>4</sup>, M. Hugh<sup>3</sup> <sup>1</sup>The Permanente Federation, Pasadena, CA <sup>2</sup>Kaiser Permanente School of Medicine, Pasadena, CA <sup>3</sup>Southern California Permanente Medical Group, Pasadena, CA <sup>4</sup>The Permanente Federation, Oakland, CA

Background: Kaiser Permanente Southern California (KPSC), a multi-specialty healthcare organization with 4.5 million members, routinely collects patient complaints pertaining to multiple issues. Non-clinical intake staff summarize the complaint and code the complaint based on a standardized coding system. Among the 500 issues coded, 29 may relate to diagnostic error using the Institute of Medicine (IOM), definition of diagnostic error - accuracy, timeliness or communication of the health problem. Are patient complaints a reliable and specific source for measuring diagnostic error?

Methods: The complaints for issues related to diagnostic error for a 7-month period in 2017 were reviewed. Of these, issues with the highest volume were selected for detailed review -- access to obtaining procedure, test, or referrals, test results delay, communication, diagnosis delayed/missed/incorrect. A total of 158 records were evaluated to determine if the patient's perspective indicated a diagnostic error, i.e., were any of the three elements of the definition of diagnostic error present in the summary? The patient's medical record was reviewed by a physician to determine if from the physician's perspective, the three elements of diagnostic error were present in medical record.

Results: From the patient's perspective, two issues, Diagnosis Delayed/Missed/Incorrect and Test Results Delay, yielded higher proportions of diagnostic error, over 80%. Other issues either had low volumes or did not meet the criteria for diagnostic error. From a physician's perspective, diagnostic error was identified at much lower rates based on documented evidence in the medical record of accurate and timely diagnosis, and communication to the patient. We believe the patient's perspective should be the gold standard regarding whether communication occurred and the diagnosis was timely because communication may be documented in the medical record but the patient failed to understand what was communicated. Because the IOM definition does not define timely, the patient's perspective should be used as this is not subject to the inter-observer variation of physician reviewers and is more patient centric. For the two issues, Diagnosis Delayed/Missed/ Incorrect and Test Results Delay, the specificity of the patient perspective was high (93% and 81%), and therefore, a reliable source to measure and monitor rates of diagnostic error.

Conclusion: Identifying diagnostic errors from the patient perspective may indicate areas for early intervention. Further review of the differences between the patient perception and physician's review of the medical record is needed to identify the ramifications of these differences, i.e., close communication gaps and assure patients' understanding of their diagnosis.

#### Diagnostic Errors, Harms, and Waste in Evaluating Dizziness and Vertigo in Ambulatory Care Settings across the **United States**

M. Keita<sup>1,2</sup>, N. Nassery<sup>3</sup>, K. Sebestyen<sup>2</sup>, Z. Wang<sup>4</sup>, W. Padula<sup>5</sup>, D. E. Newman-Toker<sup>6</sup>

The Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD

<sup>2</sup>The Johns Hopkins University School of Medicine, Baltimore, MD

<sup>3</sup>Johns Hopkins University, Baltimore, MD

<sup>4</sup>Johns Hopkins University School of Public Health, Baltimore, MD

<sup>5</sup>Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

Background: Dizziness and vertigo are among the most common symptoms seen in ambulatory care settings, including primary care and the emergency department (ED). The plurality of patients have a benign peripheral (inner ear) vestibular cause, but a small percentage (3-5%) have a dangerous central (brain) vestibular disorder, such as acute ischemic stroke. Although much has been written about failures in diagnosis of dizziness and vertigo, overall national estimates of their impact are lacking. We sought to combine multiple data sources to estimate diagnostic errors, misdiagnosis-related harms, and waste in diagnosis nationally to identify the potential benefits of implementing currently available solutions. Methods: Observational study combining US population-based estimates of dizziness visits and imaging utilization (CDC/National Center for Health Statistics); clinical trial-based diagnostic error rates (AVERT-ClinicalTrials.gov #NCT02483429); and rate of hospitalization after release with missed stroke (Kaiser Permanente Mid-Atlantic States). We analyzed the National Ambulatory Medical Care Survey (NAMCS, 2011-2015) for ambulatory office visits/utilization and the National Hospital Ambulatory Medical Care Survey (NHAMCS, 2011-2015) for ED visits/utilization. We calculated for office, ED, and total: (1) dizziness visits; (2) "benign" dizziness diagnoses (ICD-9-CM 780.4 or 386.x); (3) estimated misdiagnosed peripheral vestibular disorders; (4) stroke-related harms; and (5) neuroimaging waste.

Results: From 2011-2015 there were 3,381 sampled office visits (NAMCS) and 4,558 sampled ED visits (NHAMCS), representing a weighted total of 13.2M office and 4.5M ED visits per year (Table). Among these, 50% were given benign dizziness diagnoses, an estimated 80% of which were likely incorrect (based on prospective clinical trial data). Stroke-related harms (30-day stroke rehospitalizations after treat-and-release visit minus the expected long-term base rate of stroke revisits) were 0.19% in office care and 1.34% in ED care (based on a prior observational study of over 50,000 patients). CTs were obtained in 2.8% of office and 35.5% of ED visits, totaling 2.0M CT scans annually. Based on prior research, ~95% of CT scans for dizziness represent inappropriate use, yielding waste of \$1.38B.

Conclusion: There are more than 17.7M ambulatory care visits for dizziness and vertigo annually. Best evidence suggests that of 8.9M presumed benign vestibular cases, there are 7.1M misdiagnosed and likely ~70,000 harmed by missed stroke, while \$1.4B is wasted on CT scans alone. Current diagnosis for dizziness and vertigo patients is low-value, producing poor outcomes at high cost. Improving diagnosis of patients with dizziness and vertigo could produce major gains in quality while reducing costs.

Table Assual HC actimates for d	lizziness veleted diesmesse	middiagnages barres	and wasts in ambulatous saus sattings
Table, Abbilat US estimates for o	11221Ness-relaten niaonoses	. Misaigonases, narms	, and waste in ambulatory care settings.

Parameter (Annual US)	NAMCS Ambulatory Office Visits (95% CI)	NHAMCS Emergency Department Visits (95% CI)	Total Ambulatory (Office & Emergency) (95% CI)
Dizziness/Vertigo Visits (n)	13,230,790 (10,269,630 – 15,873,807)		
Vestibular Diagnoses (n)		1,817,684 (1,380,204 – 2,256,804)	<b>8,926,944</b> (6,638,045 – 11,249,549)
Vestibular Misdiagnoses (n)	5,687,408 (4,206,273 – 7,194,196)	1,454,147 (1,104,163 – 1,805,443)	<b>7,141,555</b> (5,310,436 – 8,999,639)
Stroke-Related Harms (n)	17,729 (14,244 – 21,271)	52,603 (41,952 – 63,294)	<b>70,333</b> (56,196 – 84,565)
CT Neuroimaging Waste (\$)	261,531,441 (33,408,065 – 806,101,765)	1,115,834,285 (274,703,562 – 2,496,466,581)	<b>1,377,365,726</b> (308,111,627 – 3,302,568,345)

<sup>&</sup>lt;sup>6</sup>Center for Diagnostic Excellence, Armstrong Institute for Patient Safety and Quality, Johns Hopkins University School of Medicine, Baltimore, MD

#### Factors of Unexpected Early Readmissions within 7 Days after Prior Discharge at an Acute Care Hospital in Japan

I. Kitagawa<sup>1</sup>, N. Harada<sup>2</sup>, R. Hara<sup>1</sup>, T. Kumagae<sup>1</sup>, J. Branch<sup>1</sup> <sup>1</sup>Shonan Kamakura General Hospital, Kamakura, Japan <sup>2</sup>Shonan Kamakura General Hospital, Kmakura, Japan

Background: Data on unexpected readmissions to hospital within six weeks after prior discharge is available in Japan. However, data on unexpected readmissions within a 7-day window period is very rare, and there have been scarce reports analyzing factors associated with those readmissions in Japan. The rate of early readmission in the department of general internal medicine (GIM) in an acute-care Japanese hospital was measured. Factors related to readmission were also investigated. We considered whether such unexpected early readmissions could have been detectable before the first discharges in such patients.

Methods: This was a retrospective cohort study at an acute-care hospital with over 600 beds. The occurrence ratios of unexpected readmission within 7 days from the year 2010 to 2017 on an annual basis, as well as the monthly trends in the year 2017 were analyzed. For data in the year 2017, the ratios in GIM and other specialties were compared. Analysis was made on the data of 59 readmitted patients in GIM with specific focus on factors such as diseases and complications.

Results: Readmission rates from the year 2010 to 2017 did not change significantly. The average age of readmitted patients was 84 years old; male to female percentage was 56% and 44% respectively, in 2017. The same data indicated a tendency of higher readmission rates in the winter compared to the summer season. The readmission rate in GIM was 3.24 percent whereas in other departments it was 1.64 percent. Infection related diseases were the major causes of readmissions.

Conclusion: This retrospective study identified that GIM at this Japanese single acute care center, has somewhat higher unplanned readmission rates. This hospital receives an unselected population of patients, unlike other hospitals in Japan, and therefore this may account for the higher rates of readmission. Patients in our GIM department are generally more elderly and are commonly diagnosed with aspiration pneumonia and urinary tract infections, respectively. Thus, this is reflected in the higher rates of readmission.

#### Leveraging Medical Malpractice Claims to Evaluate Diagnostic Error in Ischemic Stroke Care

A. L. Liberman<sup>1</sup>, J. Skillings<sup>2</sup>, P. Greenberg<sup>2</sup>, D. Siegal<sup>2</sup> <sup>1</sup>Albert Einstein College of Medicine, New York, NY <sup>2</sup>CRICO, Boston, MA

Background: Misdiagnosis of acute ischemic stroke is associated with significant patient morbidity. Cerebrovascular diseases are consistently among the top diagnoses in national medical malpractice claims. We used malpractice claims to further quantify the impact of ischemic stroke misdiagnosis on patient outcomes and identify disease specific breakdowns in the diagnostic progress.

Methods: We conducted a retrospective cohort study of closed malpractice claims housed in the Controlled Risk Insurance Company (CRICO) Strategies Comparative Benchmarking System (CBS) database from 2006 to 2015. We identified all adult ischemic stroke cases excluding those lacking claim allegation type. All claims with diagnostic-related allegations were considered instances of diagnostic error (DE). Patient outcome was defined using the National Association of Insurance Commissioners severity score where a score of 6-9 corresponds to death or permanent major disability. Diagnostic process breakdowns were categorized into five previously developed process dimensions: (1) the provider – patient encounter, (2) performance and interpretation of diagnostic tests, (3) follow-up of diagnostic test results, (4) subspecialty consultation, and (5) patients seeking care and adhering to their instruction/appointments. All study variables were abstracted from the dataset and/or via review of case summaries.

Results: A total of 478 ischemic stroke cases were identified; 235 (49%) alleged DE. Among DE claimants, half were male (50%), most were older than 45 years of age (70%), and many initially presented with both traditional and non-traditional stroke symptoms (40%). DE claims originated from the emergency department (46%) more often than the inpatient (25%) or outpatient (28%) settings. The average clinical severity among those with a DE allegation was 6.7. A total of 207 (88.0%) of DE cases had at least one contributing factor recorded that mapped to a diagnostic process dimension. The most common dimension implicated in DE claims was the provider-patient encounter. Breakdowns in the provider-patient encounter were seen more often than breakdowns in any other process dimension regardless of claimant age, sex, care setting, or stroke symptom type.

Conclusions: Allegations of DE accounted for nearly half of ischemic stroke related malpractice claims in our large database and were associated with substantial patient harm. The provider-patient encounter was most commonly implicated as a source of error. Future interventions to improve stroke diagnostic accuracy should target breakdowns in the provider-patient encounter.

#### Triggers to Identify Opportunities to Improve Diagnosis in Emergency Departments: A Consensus Study

P. Mahajan<sup>1</sup>, E. Alpern<sup>2</sup>, J. Chamberlain<sup>3</sup>, K. Cosby<sup>4</sup>, R. El-Kareh<sup>5</sup>, H. Epstein<sup>6</sup>, T. Giardina<sup>7</sup>, M. Graber<sup>8</sup>, L. Medford-Davis<sup>9</sup>, R. Medlin<sup>10</sup>, C. Mollen<sup>11</sup>, K. Shaw<sup>12</sup>, D. K. Upadhyay<sup>13</sup>, H. Singh<sup>7</sup>

<sup>1</sup>University of Michigan, Ann Arbor, MI

<sup>2</sup>Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL

3Children's National Health System, Washington, DC

<sup>4</sup>Cook County Health and Hospitals System, Chicago, IL

<sup>5</sup>University of California, San Diego, La Jolla, CA

<sup>6</sup>Brightpoint Health, Brooklyn, NY

Center for Innovations in Quality, Effectiveness and Safety, Michael E. DeBakey Veterans Affairs Medical Center and Baylor College of Medicine, Houston, TX

<sup>8</sup>RTI International, Research Triangle Park, NC

<sup>9</sup>Baylor College of Medicine, Houston, TX

10 University of Michgian, Ann Arbor, MI

<sup>11</sup>Children's Hospital of Philadelphia, Philadelphia, PA

<sup>12</sup>The Children's Hospital of Philadelphia, Philadelphia, PA

<sup>13</sup>Geisinger Medical Center, Danville, PA

Background: Triggers, such as 72-hour emergency department (ED) return visit or transfer to a higher level of care after inpatient admission from the ED, are used to identify cases at risk for diagnostic error. Evidence for the effectiveness of these triggers is sparse. We used a modified Nominal Group Technique and Delphi Process to develop consensus on triggers to identify opportunities to improve diagnosis during ED visits. Methods: We conducted a literature review of existing triggers used in ED settings and surveyed pediatric emergency medicine (PEM) division chiefs to understand current processes used in their institutions to trigger cases for review. We then developed a list of potential triggers and convened an expert panel (PEM clinicians, clinical informaticians, safety researchers, and patient representatives) to refine and reach consensus on triggers most likely to identify missed opportunities. Experts also proposed additional triggers that could be useful. The panel was divided into 2 moderated sub-panels to rank trigger categories in order of most important (1) to least important (8). The entire panel then compared rankings and came to consensus on a set of proposed triggers.

Results: We obtained consensus on six triggers to detect ED-related opportunities to improve diagnosis (Table 1). While there was general consensus around using ED return visits and care escalation, the panel recommended modifiers, such as return visits that resulted in admission or another diagnostic or therapeutic intervention, as more likely to represent a missed opportunity in diagnosis. The panel also recommended the use of symptom-disease dyads (e.g. abdominal pain-appendicitis) to identify certain high-risk conditions. Because triggers are unable to capture all types of events, the panel recommended use of additional strategies to identify errors for epidemiologic studies (e.g. cases referred to risk management or patient complaints).

Conclusion: Using consensus-based methods, we developed a list of triggers to identify potential missed opportunities to improve diagnosis in the ED. We plan to apply these triggers to electronic clinical data from five pediatric EDs to validate and quantify missed opportunities in diagnosis.

- 1. Unscheduled return visit with admission within 10 days (time period to be refined)
- 2. Return visits with new interventions that do not result in admission (e.g. diagnostic test that was abnormal, new medication). Potential Indicator: Visit chains of 3 or more visits in a pre-defined time period
- 3. All deaths within 24 hours of ED arrival or admission - exclusive of palliative care (time period needs refining)
- 4. Change of Service— a proxy for discrepancy in diagnosis may be change of service (admitted medical, changed to surgical) - within 48 hours
- 5. Care Escalation to ICU from the floor within 6/12/24 hours (validation for ED attribution would determine most optimal time)
- 6. High Risk diseases/conditions New diagnosis, symptom/disease dyads - e.g. child abuse

## Think Globally, Act Locally: Utilizing a Survey of Internal Medicine Provider Perceptions of Diagnostic Error to Inform **Future Quality Improvement Efforts**

J. C. Matulis III, S. Kok, A. Majka Mayo Clinic, Rochester, MN

Background: Advancements have been made in understanding individual and system level contributors to diagnostic error, however, little is known about how front-line clinicians perceive diagnostic error. Understanding the perceptions of those doing this work may be valuable in targeting future improvement opportunities.

Methods: A 16 question survey was administered across 2 Internal Medicine practices within a large, academic medical center. One practice is referral based, providing care for patients travelling to receive specialty evaluation while the other is a traditional community based Internal Medicine practice. The survey elicited perceptions of clinical conditions at greatest risk for diagnostic error as well as individual and system level contributors to error as defined by the National Academies of Medicine. Responses were categorized by provider demographic information and collated so that potential strategies for improvement could be identified.

Results: Of 124 clinicians surveyed, 65 (51%) responded. Thirty-Seven respondents (57%) practiced in the referral practice, while 28 (43%) practiced in the community based practice. Forty-Seven (47) percent had been in practice for less than 10 years and 44% of respondents were female. Conditions felt to be at greatest risk of diagnostic error included Cancer, any (46%), Pulmonary Embolism (43%), and Infection, any (37%). The most commonly cited factors surrounding clinician interactions included atypical patient presentations (83%), failure to consider other diagnoses (63%) and inadequate follow-up of test results (53%). The most commonly cited system level factors included "cognitive burden of the volume of data available in the electronic health record" (68%), "lack of time to think" (64%), and "systems that do not support collaboration" (40%). The vast majority of respondents (82%) were confident that diagnostic errors can be reduced.

Conclusion: Front-line clinician perspectives on conditions at risk of diagnostic error and systems level factors contributing to diagnostic error may differ from the published literature. Interventions which provide point of care clinical support, reduce the clinician's cognitive burden and closed loop communication of test results may hold promise in error reduction.

#### Patients' Use of an Online Symptom-Checker and Impact on Diagnosis-Related Outcomes

A. N. Meyer<sup>1</sup>, T. Giardina<sup>1</sup>, C. Spitzmueller<sup>2</sup>, H. Singh<sup>1</sup>

<sup>1</sup>Center for Innovations in Quality, Effectiveness and Safety, Michael E. DeBakey Veterans Affairs Medical Center and Baylor College of Medicine, Houston, TX

<sup>2</sup>University of Houston, Houston, TX

Background: Patients increasingly seek online symptom-checkers to obtain diagnoses. Little is known, however, about characteristics of patients who use these resources, their rationale for use, and whether patients find them accurate and useful. We evaluated how the Isabel Symptom Checker, a popular online symptom-checker, impacted patient experiences with their triage and diagnosis process. Methods: We invited the first 385 respondents from 4,000 registered US users of Isabel Symptom Checker to complete an online survey. Respondents received \$20 gift-card incentives; thus, available funds determined sample size. In addition to demographics, questions focused on why patients used the tool, self-reported health and financial outcomes related to its use, whether patients discussed symptom-checker use with their physicians, and if so, physicians' receptiveness to patients' use of the tool. Results: From the sample of 385 patients, 329 provided usable and complete responses. Mean respondent age was 48.0 years (SD=16.7); most were women (n=230;75.7%), white/Caucasian (n = 271;89.1%), with Bachelor's degrees or higher (n = 191;58.1%), and had <\$100,000 in household income (n = 216;65.6%), health care coverage (n = 296;97.4%), and chronic health conditions (n = 216;65.7%). More than half reported history of previous misdiagnoses (n = 181;55.0%). Patients used the symptom-checker to better understand what could cause their symptoms (n = 232;76.3%), to decide whether to seek care (n=101;33.2%), or what care setting to visit (e.g., primary or urgent care; n=63;20.7%). Many respondents used the symptom-checker before (n = 119;39.1%) or both before *and* after seeing a physician (n = 113;37.2%). Most users thought the symptom-checker gave them useful information for their health problems (n = 274;83.3%) with about half reporting positive health effects (n = 154;50.9%). While over half were neutral in terms of benefitting financially (n=172;56.8%); most found the symptom-checker useful: perceiving it to be satisfying (n=263;79.9%), easy to use (n = 182;55.3%), useful as a diagnostic tool (n = 253;76.9%), and providing them with insights leading them closer to correct diagnoses (n=231,70.2%). Additionally, most reported they would use it again (n=278,84.5%). Of those who visited physicians after using the tool (n = 213;70.1%), almost half discussed the findings with their physicians (n = 103;48.4%). Experiences were mixed, but patients did not always feel physicians were interested (n = 24;23.3%) or open to discussing the tool's results (n = 21;20.4%). **Conclusion**: While accessing a popular online symptom-checker for triage and diagnosis, patients reported receiving information that they found useful in their diagnostic process. However, evaluation of longer-term, objective health benefits, and evaluation of benefits for more diverse patient groups, is needed to better understand the broader impact of symptom-checkers on diagnosis and health outcomes.

#### How Available Is Availability? a Study to Identify Factors That Influence Reliance on Available Information

S. Monteiro<sup>1</sup>, J. Sherbino<sup>1</sup>, G. Norman<sup>1</sup>, J. Ilgen<sup>2</sup>, E. Hayden<sup>3</sup> <sup>1</sup>McMaster University, Hamilton, ON, Canada <sup>2</sup>University of Washington, Seattle, WA <sup>3</sup>Harvard Medical School, Boston, MA

Background: Cognitive bias is often cited as an explanation for diagnostic errors. Availability bias, defined as "The current case reminds you of a recent similar example or diagnostic category" is most well-known. Despite the ubiquity of availability bias in medical and popular literature, there is surprisingly little evidence to substantiate these claims. The present study sought to measure the influence of availability bias. **Methods:** We devised a sequential 2 phase experiment. Phase 2, the test phase, consisted of 8 written cases, each with two approximately equally likely diagnoses (Diagnosis A and B). The experimental intervention was in a "validation" phase preceding the test, where participants were asked to first verify a diagnosis which was either representative of i) Diagnosis A, and similar to a test case, b) Diagnosis A and dissimilar, c) Diagnosis B and similar. Each participant verified 2 cases from each condition, for a total of 6. Participants were not told of the manipulation. They then diagnosed all 8 test cases (the remaining 2 test cases had no prior). All cases were counterbalanced across conditions. We compared accuracy for cases from Condition a) and b) with the no-prior cases to evaluate the effect of prior exemplar; comparisons between c) and no-prior evaluated the effect of category priming. Because cases were designed so that both Diagnosis A and B were equally likely, overall accuracy was measured as the sum of proportion of cases in which either was selected. Subjects were emergency medicine staff (n = 40), residents (n = 39) and medical students (n = 32) from McMaster University, University of Washington, and Harvard Medical School. Results: Overall, staff had an accuracy (A + B) of 98%, residents 98% and students 85% (F = 35.6,p <.0001). For residents and staff there was no effect of condition (accuracy ranged from 97% to 100%); for students there was a clear effect of category priming, with accuracy of 84% for a), 87% for b) and 94% for d) but only 73% for the no prime condition (Interaction F = 3.54, p < .002)

Conclusion: Although prior research has shown substantial effects of similarity to prior cases, primarily in visual domains, the present study has shown such effects only for medical students. Possible explanations will be explored. Nevertheless, the study shows that with increasing expertise, availability may not be a source of error.

#### **Debiasing Versus Knowledge Retrieval Checklists to Reduce Diagnostic Error**

M. Sibbald<sup>1</sup>, J. Sherbino<sup>1</sup>, J. Ilgen<sup>2</sup>, L. Zwaan<sup>3</sup>, S. Blissett<sup>4</sup>, S. Monteiro<sup>1</sup>, G. Norman<sup>1</sup> <sup>1</sup>McMaster University, Hamilton, ON, Canada <sup>2</sup>University of Washington, Seattle, WA <sup>3</sup>Erasmus MC, Rotterdam, Netherlands <sup>4</sup>McGill University, Montreal

Background: There is an ongoing debate regarding the cause of diagnostic errors. One view is that errors result from unconscious application of cognitive heuristics; the alternative is that errors are a consequence of knowledge deficits.

Methods: Novice postgraduate year (PGY) 1 or 2 internal medicine residents and experienced cardiology residents (PGY 4 or 5) were randomly allocated to three conditions: a debiasing checklist, a content (knowledge) checklist, or control (no checklist). All groups received instructions on ECG interpretation and diagnostic errors. Those randomized to the debiasing checklist were taught about common biases whereas those randomized to the content checklist were taught about key components of the ECG. All participants were tested on a set of 20 ECGs, with 10 ECGs engineered to promote either search satisficing or confirmation bias. Participants in each intervention arm used a checklist to aid their interpretations.

Results: Forty novice residents and 21 experienced residents were enrolled. Participants made an average of 1.19 ± 1.34 errors per ECG. Neither intervention was significantly better than the control. As expected, more errors were identified among novice learners and when cases were engineered induce bias. The interaction term between the instructional condition and level of learner was of borderline significance. Novice learners had fewer errors when using either checklist, while errors rates were the same in all three conditions for expert learners.

Conclusion: Checklists attempting to help learners identify cognitive bias or mobilize domain specific knowledge did not systematically reduce diagnostic errors. Teachers should be cautious about integrating checklists into curricula.

#### Eyeballing: The Role of Experience in Rapid Visual Diagnosis of 'Sick'

A. LoGiudice<sup>1</sup>, M. Sibbald<sup>2</sup>, S. Watter<sup>3</sup>, G. Norman<sup>2</sup>, S. Monteiro<sup>2</sup> <sup>1</sup>McMaster University, ON, Canada <sup>2</sup>McMaster University, Hamilton, ON, Canada <sup>3</sup>McMaster University

Background: Medical reference textbooks do not list criteria to quickly classify someone as 'sick' or 'not sick'. Yet rapid visual assessment of illness severity—often called 'eyeballing'—is considered a vital skill by experts in acute care disciplines. Surprisingly little is known about the acquisition of this skill and how to assess it. Here we report ongoing research aimed at understanding the development of eyeballing skill

Methods: In a prior study, we validated a laboratory task where each case consisted of 15 s videos of a real patient presenting to the emergency room and a still image of their cardiac monitor. With this minimal information, and an average viewing time of only 11 s, emergency physicians could rapidly estimate illness severity with reasonable accuracy. Next we aim to (a) test residents at different levels of training to examine the trajectory of eyeballing proficiency as a function of experience; and (b) add an eye-tracking component to test the hypothesis that more proficient eyeballers quickly attend to the 'subjective' visual information of the patient, rather than the more 'objective' information derived from the cardiac monitor.

Results: Previous work has demonstrated high agreement between physicians on patient disposition. Inter-rater reliability judging 'sick' was computed as an intraclass correlation coefficient (ICC) of 0.54. Agreement among physicians in predicting disposition was 68% with ICC of 0.44, and accuracy at predicting disposition was 55%. We are currently collecting eye tracking data to help understand what information facilitates these decisions in novices and experts.

Conclusion: These data will inform dual-process theories of decision making that posit a System 1 process which rapidly compares a presenting patient to many prior patient encounters.

#### Serious Misdiagnosis-Related Harms Due to Missed Infections in the United States-Incidence, Settings, and Causes

N. Nassery<sup>1</sup>, A. C. Schaffer<sup>2</sup>, C. W. Yu-Moe<sup>3</sup>, A. S. Tehrani<sup>4</sup>, G. D. Clemens<sup>5</sup>, Z. Wang<sup>6</sup>, M. Fanai<sup>7</sup>, D. Siegal<sup>8</sup>, D. E. Newman-Toker<sup>7</sup>

<sup>1</sup>Johns Hopkins University, Baltimore, MD

<sup>2</sup>Brigham and Women's Hospital, Boston, MA

3CRICO, Boston, MA

<sup>4</sup>Massachusetts Eye & Ear Infirmary/Harvard University, Boston, MA

<sup>5</sup>The Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

<sup>6</sup>Johns Hopkins University School of Public Health, Baltimore, MD

<sup>7</sup>Center for Diagnostic Excellence, Armstrong Institute for Patient Safety and Quality, Johns Hopkins University School of Medicine, Baltimore, MD

8CRICO Strategies, Boston, MA

Background: Diagnostic errors cause substantial preventable harm, but population-based estimates of aggregate morbidity and mortality are lacking. Missed infections account for a large fraction of the serious morbidity and mortality attributable to diagnostic errors in clinical, autopsy, and malpractice studies. We sought to estimate the US national malpractice-based proportion and population-based incidence of missed infections, and define the most common settings and causes for these errors.

Methods: We performed a cross-sectional study of missed/delayed infection diagnoses. We analyzed coded malpractice claims in a national comparative benchmarking system representing ~30% of all US claims (CRICO Strategies 2006-2015 [n = 55,377]), and identified cases attributable to diagnostic error. Diagnostic error claims were grouped as infections using AHRQ Clinical Classifications Software to aggregate International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes, and the top high-risk-of-harm infections were identified. Annual US disease incidence for each infection code was measured using the Healthcare Cost and Utilization Project's National Inpatient Sample (NIS 2012-2014 [n = 21.5M sampled hospital discharges]). Disease-specific rates of diagnostic error and generic harm rates for the top five infections were estimated from published studies; plausible ranges were constructed with expert input when the data were derived from lower-quality studies. Total serious misdiagnosis-related harms (death or permanent disability) from infections were extrapolated to the population-level using the attributable fraction of infections in the malpractice data. Ranges were calculated as 95% confidence intervals for the individual infections and probabilistic plausible ranges (PPR) for aggregate estimates using Monte Carlo simulation. Settings and causes were derived from the malpractice claims.

Results: 11,592 diagnostic error claims were identified (median age 49, IQR 36-60; 51.7% female); 7,379 were classified as high-severity claims. Infections accounted for 13.4% of high-severity malpractice claims (n=992/7.379), 43.3% of estimated incident diagnostic errors (1,017,918/2,351,356; PPR 602,682-1,575,384), and 45.3% of incident misdiagnosis-related harms (250,137/552,532; PPR 156,342-417,113). Settings were ambulatory clinics (29.6%), emergency departments (28.2%), and inpatient units (41.2%). Top contributing factors were clinical judgment (84%), communication (33%), and clinical systems (21%). Top infections (error rate; errors/year) included: sepsis (20.8%; 279,558), pneumonia (11.6%; 170,100), meningitis/encephalitis (25.6%; 12,036), endocarditis (25.5%; 8,758), and spinal abscess (55.5%; 7,520). The top five accounted for 47.0% of infection-related errors (477,972/1,017,918) and 60.5% of harms (151,293/250,137).

Conclusion: Best available evidence suggests that ~250,000 Americans die or are permanently disabled from missed infections each year. Five infections account for about 60% of these serious misdiagnosis-related harms. Initial efforts to improve diagnosis should focus on these conditions.

### Serious Misdiagnosis-Related Harms Due to Missed Cancers in the United States—Incidence, Settings, and Causes

N. Nassery<sup>1</sup>, A. C. Schaffer<sup>2</sup>, C. W. Yu-Moe<sup>3</sup>, A. S. Tehrani<sup>4</sup>, G. D. Clemens<sup>5</sup>, Z. Wang<sup>6</sup>, M. Fanai<sup>7</sup>, D. Siegal<sup>8</sup>, D. E. Newman-Toker<sup>7</sup>

<sup>1</sup>Johns Hopkins University, Baltimore, MD

<sup>2</sup>Brigham and Women's Hospital, Boston, MA

3CRICO, Boston, MA

4Massachusetts Eve & Ear Infirmary/Harvard University, Boston, MA

<sup>5</sup>The Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

<sup>6</sup>Johns Hopkins University School of Public Health, Baltimore, MD

<sup>7</sup>Center for Diagnostic Excellence, Armstrong Institute for Patient Safety and Quality, Johns Hopkins University School of Medicine, Baltimore, MD

<sup>8</sup>CRICO Strategies, Boston, MA

Background: Diagnostic errors cause substantial preventable harm, but population-based estimates of aggregate morbidity and mortality are lacking. Missed cancers account for a large fraction of the serious morbidity and mortality attributable to diagnostic errors in clinical, autopsy, and malpractice studies. We sought to estimate the US national malpractice-based proportion and population-based incidence of missed cancers, and define the most common settings and causes for these errors.

Methods: We performed a cross-sectional study of missed/delayed diagnoses. We analyzed coded malpractice claims in a national comparative benchmarking system representing ~30% of all US claims (CRICO Strategies 2006-2015 [n=55,377]), and identified cases attributable to diagnostic error. Diagnostic error claims were grouped as cancers using AHRQ Clinical Classifications Software to aggregate International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes, and the top high-risk-of harm cancers were identified. Annual US cancer incidence was measured using data from the North American Association of Central Cancer Registries (NAACCR 2014, US-based registries). Disease-specific rates of diagnostic error and generic harm rates for the top five cancers were estimated from published studies; plausible ranges were constructed with expert input when the data were derived from lower-quality studies. Total serious misdiagnosis-related harms (death or permanent disability) from cancers were extrapolated to the population-level using the attributable fraction of cancers in the malpractice data. Ranges were calculated as 95% confidence intervals for the individual cancers and probabilistic plausible ranges (PPR) for aggregate estimates using Monte Carlo simulation. Settings and causes were derived from malpractice claims.

Results: 11,592 diagnostic error claims were identified (median age 49, IQR 36-60; 51.7% female); 7,379 were classified as high-severity claims. Cancers accounted for 37.9% of high-severity malpractice cases (n=2,793/7,379), 8.1% of estimated incident diagnostic errors (190,990/2,351,356; PPR 141,883-467,719), and 10.9% of incident misdiagnosis-related harms (60,454/552,532; PPR 44,610-149,583). Settings were ambulatory clinics (87.2%), emergency departments (3.8%), and inpatient units (8.4%). Top contributing factors were clinical judgment (80%), communication (32%), and clinical systems (28%). Top malignancies (error rate; errors/year) included: lung cancer (22.5%; 50.466), colorectal cancer (19.9%; 27,179), breast cancer (8.9%; 20,866), melanoma (10.2%; 7,800), and prostate cancer (2.4%; 5,551). The top five malignancies accounted for 58.6% of cancer-related errors (111,862/190,990) and 58.6% of harms (35,408/60,454).

Conclusion: Best available evidence suggests that ~60,000 Americans die or are permanently disabled from missed cancers annually. Five cancers account for nearly 60% of these serious misdiagnosis-related harms. Initial efforts to improve cancer diagnosis should focus on these malignancies.

#### Serious Misdiagnosis-Related Harms Due to Missed Vascular Events in the United States-Incidence, Settings, and Causes

A. S. Tehrani<sup>1</sup>, N. Nassery<sup>2</sup>, A. C. Schaffer<sup>3</sup>, C. W. Yu-Moe<sup>4</sup>, G. D. Clemens<sup>5</sup>, Z. Wang<sup>6</sup>, M. Fanai<sup>7</sup>, D. Siegal<sup>8</sup>, D. E. Newman-Toker<sup>7</sup>

<sup>1</sup>The Massachusetts Eye & Ear Infirmary, Harvard Medical School, Boston, MA

<sup>2</sup>Johns Hopkins University, Baltimore, MD

<sup>3</sup>Brigham and Women's Hospital, Boston, MA

<sup>4</sup>CRICO, Boston, MA

<sup>5</sup>The Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

<sup>6</sup>Johns Hopkins University School of Public Health, Baltimore, MD

<sup>7</sup>Center for Diagnostic Excellence, Armstrong Institute for Patient Safety and Quality, Johns Hopkins University School of Medicine, Baltimore, MD

<sup>8</sup>CRICO Strategies, Boston, MA

Background: Diagnostic errors cause substantial preventable harm, but population-based estimates of aggregate morbidity and mortality are lacking. Missed vascular events account for a large fraction of the serious morbidity and mortality attributable to diagnostic errors in clinical, autopsy, and malpractice studies. We sought to estimate the US national malpractice-based proportion and population-based incidence of missed vascular events, and define the most common settings and causes for these errors.

Methods: Cross-sectional study of missed/delayed vascular diagnoses. We analyzed coded malpractice claims in a national comparative benchmarking system representing ~30% of all US claims (CRICO Strategies 2006-2015 [n = 55,377]), and identified cases attributable to diagnostic error. Diagnostic error claims were grouped as vascular events using AHRQ Clinical Classifications Software to aggregate International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes, and the top high-risk-of-harm vascular events were identified. Annual US incidence for each vascular disease was measured using the Healthcare Cost and Utilization Project's National Inpatient Sample (NIS 2012-2014 [n=21.5M sampled hospital discharges]). Disease-specific diagnostic error rates and generic harm rates for the top five vascular diseases were estimated from published studies; plausible ranges were constructed with expert input when data were from lower-quality studies. Total serious misdiagnosis-related harms (death or permanent disability) from vascular events were extrapolated to the population-level using the attributable fraction of vascular events in malpractice data. Ranges were calculated as 95% confidence intervals for individual diseases and probabilistic plausible ranges (PPR) for aggregate estimates, using Monte Carlo simulation. Settings/causes were derived from malpractice claims.

Results: 11,592 diagnostic error claims were identified (median age 49, IQR 36-60; 51.7% female); 7,379 were high-severity claims. Vascular events accounted for 22.8% of high-severity malpractice claims (n=1,684/7,379), 22.7% of incident diagnostic errors (533,817/2,351,356; PPR 298,363-667,314), and 17.9% of incident misdiagnosis-related harms (98,922/552,532; PPR 59,885-130,175). Settings were ambulatory clinics (34.3%), emergency departments (31.1%), and inpatient units (34.1%). Top contributing factors were clinical judgment (86%), communication (33%), and documentation (22%). Top vascular conditions (error rate; errors/year) included: stroke (8.7%; 82,832), venous thromboembolism (19.9%; 63,527), arterial thromboembolism (28.2%; 48,886), myocardial infarction (2.2%; 27,458), and aortic aneurysm/dissection (9.0%; 26,719). The top five accounted for 46.7% of vascular-event errors (249,422/533,817) and 79.8% of vascular-event harms (78,950/98,922).

Conclusion: Best available evidence suggests that ~100,000 Americans die or are permanently disabled from missed vascular events each year. Five vascular diseases account for ~80% of these serious misdiagnosis-related harms. Initial efforts to improve diagnosis should focus on these conditions.

#### Total US Societal Costs of Harms from Diagnostic Error Estimated from Malpractice and Population-Based Data

D. E. Newman-Toker<sup>1</sup>, M. Keita<sup>2</sup>, N. Nassery<sup>3</sup>, A. C. Schaffer<sup>4</sup>, C. W. Yu-Moe<sup>5</sup>, A. S. Tehrani<sup>6</sup>, G. D. Clemens<sup>7</sup>, Z. Wang<sup>8</sup>, M. Fanai<sup>1</sup>, D. Siegal<sup>9</sup>, W. Padula<sup>10</sup>

<sup>1</sup>Center for Diagnostic Excellence, Armstrong Institute for Patient Safety and Quality, Johns Hopkins University School of Medicine, Baltimore, MD

<sup>2</sup>The Johns Hopkins University School of Medicine, Baltimore, MD

<sup>3</sup>Johns Hopkins University, Baltimore, MD

<sup>4</sup>Brigham and Women's Hospital, Boston, MA

5CRICO, Boston, MA

6Massachusetts Eye & Ear Infirmary/Harvard University, Boston, MA

<sup>7</sup>The Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

<sup>8</sup>Johns Hopkins University School of Public Health, Baltimore, MD

9CRICO Strategies, Boston, MA

<sup>10</sup>Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

Background: Misdiagnosis-related harms incur major societal costs, but aggregate estimates are lacking. Costs from missed and delayed diagnosis of dangerous disorders such as vascular events, infections, and cancers (the "Big Three") accrue from both "direct" costs from excess resource use (e.g., more intensive care from treating illness later, after progression to higher severity) and "indirect" costs from lost work (e.g., reduced productivity from disability). Both types of cost can be represented by malpractice claim payouts and net monetary benefit (NMB) to society as a whole. We combined malpractice data with epidemiologic data from nationally-representative samples to estimate costs and NMB. Methods: Cross-sectional study of missed/delayed diagnoses. Big Three diseases identified from coded malpractice claims in a national comparative benchmarking system representing ~30% of all US claims (CRICO Strategies 2006-2015 [n = 55,377]). Annual US disease incidence measured from Healthcare Cost and Utilization Project's National Inpatient Sample (2012-2014) and North American Association of Central Cancer Registries (2014). Disease-specific rates of diagnostic error and generic harm rates estimated from published studies; plausible ranges constructed with expert input when data derived from lower-quality studies. Overall serious misdiagnosis-related harms (death/permanent disability) extrapolated to population-level using attributable fraction of harms from Big Three diseases in malpractice data. Ranges calculated as 95% confidence intervals for individual diseases and probabilistic plausible ranges (PPR) for aggregate estimates, using Monte Carlo simulation. Societal costs calculated using two approaches: (1) population-based harms (n) x mean malpractice payout per harm (\$); and (2) NMB as a function of Quality-adjusted Life Years (QALYs) x willingness-to-pay (WTP) threshold of \$50,000-100,000 per QALY. We considered a range of possible parameters.

Results: We estimated 552,532 (PR 400,974-870,698) serious harms from all diseases. The mean per claim payout was \$248,976, so cost estimate #1 was \$138B (\$100B-217B). Under minimum assumptions (QALYs lost up to point of retirement [age 65], but not after; mean 0.5 QALYs lost due to death; mean 0.2 QALYs lost due to disability; WTP = \$50,000/QALY), NMB #2 was \$135B. Under maximal assumptions (to death [age 78.8]; 0.8 QALYs death; 0.5 QALYs disability; WTP = \$100,000), NMB #2 was \$994B.

**Conclusion:** Our analysis suggests the societal costs of missed dangerous diseases exceeds \$100B per year, and could be as high as ~\$1T (~6% of US gross domestic product) if downstream economic consequences are considered. Improving diagnosis in healthcare is a public health imperative, and societal resources commensurate the economic impact of the problem should be devoted to addressing it.

# Rapid Detection of Mycobacterium Tuberculosis Complex Species from Formalin-Fixed Tissue Specimens By Pyrosequencing-Based PCR Assay

D. B. Rabeneck<sup>1</sup>, M. DeLeon-Carnes<sup>1</sup>, V. W. Rowlett<sup>2</sup>, S. Reagan-Steiner<sup>2</sup>, R. B. Martines<sup>3</sup>, J. Ritter<sup>2</sup>, G. Hale<sup>2</sup>, J. Gary<sup>2</sup>, T. Huynh<sup>2</sup>, W. J. Shieh<sup>2</sup>, S. R. Zaki<sup>2</sup>, I. Bhatnagar<sup>2</sup>

<sup>1</sup>Centers for Disease Control and Prevention, ATLANTA, GA <sup>2</sup>CENTERS FOR DISEASE CONTROL AND PREVENTION, ATLANTA, GA <sup>3</sup>INFECTIOUS DISEASES PATHOLOGY BRANCH, ATLANTA, GA

**Background:** *Mycobacterium tuberculosis* is an infectious agent of great public health concern because of high morbidity and mortality rates among high risk populations and the emergence of drug resistance. The timeliness and accuracy of diagnosis is critical for prevention of disease transmission and selection of appropriate antimicrobial treatment. Culture-based identification of *Mycobacterium tuberculosis* complex species (MTBC) is tedious and time-consuming. Also, often, due to clinical suspicion of non-infectious conditions such as cancer or sarcoidosis, specimens are not processed for culture and formalin-fixed, paraffin-embedded (FFPE) tissues are the only specimens available. Acid-fast stains and mycobacterial immunohistochemistry can be performed on FFPE tissues, but these methods cannot differentiate between tuberculous and non-tuberculous *Mycobacterium* spp.

**Methods:** For rapid, sensitive and accurate identification of MTBC in FFPE tissues, we developed a pyrosequencing-based PCR (PSQ) assay targeting the IS6110 region. FFPE tissues from 105 cases submitted to CDC's Infectious Diseases Pathology Branch (IDPB) between January 2017—May 2018 with clinical and histopathological suspicion of MTBC were evaluated by the IS6110 PSQ assay. DNA was extracted from FFPE tissues and the assay was performed on the Qiagen PyroMarK Q48 pyrosequencing system.

**Results:** Of 105 cases, MTBC was identified in tissues of 73 (70%) cases by the IS6110 PSQ assay. The PSQ assay identified MTBC in all 66 cases in which MTBC was previously detected by a conventional *Mycobacterium* genus PCR targeting the 16S rRNA gene and Sanger sequencing. In addition, the PSQ assay also identified MTBC in seven out of the remaining 39 (18%) cases that were negative for MTBC by the conventional *Mycobacterium* PCRs. In three of these seven cases, conventional 16S rRNA and other PCRs previously detected *Corynebacterium* spp. or *M. gordonae*, suggesting that MTBC detection might be masked by other bacteria or environmental contaminants. Also, the PSQ assay identified MTBC from multiple tissue types (lung, lymph node, brain, liver, spleen, heart, and testis).

**Conclusion:** The IS6110 PSQ assay enhanced the sensitivity of MTBC detection in FFPE tissues, demonstrated specificity for MTBC, even in the presence of environmental contaminants, and is rapid and cost-effective compared to conventional PCR and sequencing. The PSQ assay can be a useful adjunct for MTBC diagnosis, particularly in cases where conventional specimens are unavailable for testing.

#### Clinical Reasoning: How Should We Teach It?

<u>V. Schaye</u>, K. Eliasz, M. Janjigian, D. Stern New York University School of Medicine, New York, NY

**Background:** Diagnostic errors have a significant impact on our health care system with cognitive errors contributing to the majority of cases. Educators have theorized that interventions grounded in dual process theory (DPT) may improve the clinical reasoning (CR) process of physicians but little empirical evidence of this theory exists.

**Methods:** This study was a quasi-experimental design in the New York University Internal Medicine Residency Program from June 2017-January 2018. We implemented two educational interventions in CR grounded in DPT during this 6-month period, leading to a natural experiment with three groups: no intervention (N=25), partial intervention (received part 1, N=23), and full intervention (received part 1 and part 2, N=23). The educational interventions covered the concepts of DPT, impact of diagnostic errors, and case-based discussions introducing techniques to develop fast and slow thinking. We used the diagnostic thinking inventory (DTI) at baseline (a 41 item self-assessment questionnaire to assess one's approach to CR). At 6 months, participants completed a follow-up DTI and a post-survey assessing their ability to apply concepts to cases as well as workplace experiences of CR teaching. Participants who completed pre- and post-surveys were included in the analysis. Case examples were scored by two independent reviewers blinded to group status (Table). Differences between groups were analyzed using a one-way analysis of variance.

**Results:** Those in the full intervention group reported more teaching of DPT on attending rounds and in conference than the partial intervention and no intervention groups (60% often or always taught compared to 25% and 8.3%, p = 0.002; 73.3% often or always taught compared to 58.3% and 33.3%, p = 0.033, respectively). Otherwise workplace experiences were not significantly different between the groups. There was a significant difference between groups in ability to apply concepts to cases with a medium to large effect size (Table). There was also a significant difference in giving case examples in problem representation format with a large effect size (Table). There was no significant

difference in change in DTI scores (mean change in score no intervention 7.0 (SD 16.3), partial intervention 8.8 (SD 9.8), full intervention 7.8 (SD 12.0), p = 0.946).

Conclusion: This study provides evidence that interns who participated in a CR curriculum grounded in DPT were effective in applying principles of CR in cases from their own clinical practice and supports the argument that we should be designing educational interventions in CR grounded in DPT. Subsequent studies would need to assess further impact on patient outcomes.

	No Intervention Mean Score (SD) n = 12	Partial Intervention Mean Score (SD) n = 12	Full Intervention Mean Score (SD) n = 15	p Value	Partial η²	Cohen's d (between no and full intervention)
Named Clinical	1.8	2.3	2.3	0.37	0.047	0.51
Reasoning Concept	(0.99)	(0.90)	(0.97)			
Applied Clinical	1.6	2.3	2.2	0.05	0.132	0.76
Reasoning Concept	(0.65)	(0.81)	(0.91)			
Gave Example in	1.2	1.5	2.1	0.004	0.241	1.2
Problem	(0.38)	(0.55)	(0.93)			
Representation						
Format						
Total Score	4.7	6.2	6.6	0.034	0.152	0.92
	(1.8)	(1.7)	(2.3)			

Items were scored as not done = 1, partially done = 2, well done = 3 with a total score range of 3-9 Linear weighted Cohen's Kappa=0.89

#### What Stumps Primary Care Clinicians? First Phase Analysis of Diagnostic Uncertainty Cases

L. Sommers1, V. Vaghani2, T. Giardina2, H. Singh2

<sup>1</sup>University of California, San Francisco, San Francisco, CA

<sup>2</sup>Center for Innovations in Quality, Effectiveness and Safety, Michael E. DeBakey Veterans Affairs Medical Center and Baylor College of Medicine, Houston, TX

Background: Diagnostic difficulty is frequent in primary care, yet we know little about situations that present uncertainty to primary care practitioners (PCPs). In this study, we explored how PCPs discuss diagnostic conundrums with colleagues, frame dilemmas, and seek and receive support. This knowledge could advance understanding of knowledge/skill needs of PCPs and systems-based problems contributing to diagnostic error.

Methods: We analyzed data from a cohort of 459 uncertainty cases compiled by the first author while PCPs presented cases to peers in "Practice Inquiry Colleague Groups" conducted at 17 San Francisco-Bay Area primary care practices between 10/2002 and 9/2015. Clinicians discussed patients for whom they encountered uncertainty of any type (e.g., diagnostic, management, ethical) for ideas and support. Meeting notes on each case included: 1) PCP presenter's uncertainty statement/question; 2) patient's medical history; and 3) colleague group's discussion summary. We assessed meeting notes for diagnostic uncertainty building on criteria developed in prior work on measuring uncertainty from medical record review. We collected the following data for further analysis: 1) presenters' uncertainties: symptoms, abnormal physical exam, and/or abnormal test/imaging findings; 2) diagnoses and conditions under consideration by presenter; and 3) patients' prior diagnoses and known conditions. Two authors analyzed data and used a consensus approach to reconcile differences.

Results: 0f 459 cases discussed, 258 (56%) involved diagnostic uncertainty. Most cases involved women (61%); median age decade 50-59 years; 16% included mention of ethnicity. Most (93%) had one or more physical complaints and/or abnormal findings contributing to uncertainty (See Table.) PCPs speculated about specific diagnoses and conditions that could contribute to uncertainty in 129 cases (50%); of these, 27% were mental, behavioral/neurodevelopmental conditions. They discussed patients' prior diseases and known conditions related to the uncertainty in 183 (71%) cases; of these, 17% were mental, behavior/ neurodevelopmental conditions. Adverse events were reported in 14% of cases including 6 deaths.

Conclusion: Practice inquiry-based techniques illustrate unique diagnostic uncertainties faced by PCPs. Certain situations, such as patients with mental, behavioral or neurodevelopmental conditions, might pose particular diagnostic challenges. This small-group process also facilitated open discussions of diagnostic error and patient harm. We now plan to analyze colleagues' responses to these cases.

**DE GRUYTER** 

PCP Presenters' Data Category	#/% Cases
Types of Diagnostic Uncertainties (n = 258 cases)	
≥1 non-pain, physical symptom	182 (71%)
<ul> <li>≥1 abnormal lab/imaging finding</li> </ul>	87 (34%)
≥1 pain symptom	72 (28%)
2. Possible Diseases/Conditions (ICD-11)	
(n = 171 diseases/conditions)	
<ul> <li>Mental, behavioral or neurodevelopmental disorders</li> </ul>	46 (27%)
<ul> <li>Diseases of the nervous system</li> </ul>	14 (8%)
Diseases of the circulatory system	14 (8%)
3. Past, Known Diseases/Conditions (ICD-11)	
(n = 402 diseases/conditions)	
<ul> <li>Mental, behavioral or neurodevelopmental disorders</li> </ul>	70 (17%)
<ul> <li>Endocrine, nutritional or metabolic diseases</li> </ul>	50 (12%)
Diseases of the circulatory system	42 (10%)

#### Examining Referral Patterns to Improve Diagnosis of Neuro-Ophthalmologic Conditions.

L. Stunkel<sup>1</sup>, B. Bruce<sup>2</sup>, D. D. Mackay<sup>3</sup>, N. J. Newman<sup>1</sup>, V. Biousse<sup>1</sup> <sup>1</sup>Emory University, Atlanta, GA <sup>2</sup>na, na, GA <sup>3</sup>Indiana University, Indianapolis, IN

Background: Neuro-ophthalmology specializes in the evaluation of patients with complex problems that may be urgently vision- or lifethreatening. Neuro-ophthalmologists often encounter diagnostic dilemnas or manage complex work-ups. However, access to neuro-ophthalmologists is currently limited by the relatively small number of providers, and patients may be affected by incorrect or delayed diagnosis of these conditions. The objective of this study was to analyze referrals patterns to neuro-ophthalmologists, and to characterize the rate of misdiagnosis and delayed diagnoses in patients who are ultimately referred to neuro-ophthalmology.

Methods: Retrospective chart review of 300 new patients seen over 45 randomly chosen days between 06/2011-06/2015 in one neuro-ophthalmology clinic affiliated with a tertiary university hospital [medical records of patients seen before 06/2015 were extensive paper charts, and therefore more complete than current electronic medical records]. Demographics, distance traveled, time between disease-onset and neuro-ophthalmology consultation, time between appointment request and neuro-ophthalmology consultation, number and specialty of providers seen before referral to neuro-ophthalmology, unnecessary tests prior to referral, referral diagnosis, final diagnosis, impact of the neuro-ophthalmology consultation on the patient's outcome were collected.

Results: Patients traveled a median of 36.5 miles [IQR:20-85; range:0-1059] for their neuro-ophthalmology appointment. Median time from first symptom to visit was 210 days (IQR:70-1100). Median time from request to visit was 34 days (IQR:7-86) with a bimodal distribution: one peak within one week of request (urgent requests addressed directly to physicians); another peak at about 3 months after request (through a call center). 247/300 (82.3%) were complex/very complex; referral was felt appropriate in 242/300 (80.6%). Median number of providers seen before neuro-ophthalmology was 2 (IQR:2-4; range:0-10). The accuracy of referral diagnosis was 39.6% (119/300). Mismanagement of the patient/delay in providing appropriate care was present in 85 cases (28.3%), with unnecessary tests in 56/300 (18.6%), unnecessary consultations in 65/300 (21.6%), and misinterpretation of imaging in 16 patients (5.3%). The impact of the neuro-ophthalmology consultation on patients' outcome included: 1) no impact (5/300, 1.6%); 2) provided reassurance/avoided further visits/tests (55/300, 18.3%); 3) provided a diagnosis and direction to treatment (178/300, 59.3%); 4) played a major role in preserving vision/prevented life-threatening complications/ avoided harmful treatment (62/300, 20.6%).

Conclusion: The majority of referrals were appropriate, but referral was often delayed and misdiagnosis prior to referral to neuro-ophthalmology was not rare. In this study, neuro-ophthalmologic evaluation often played a major role in preserving vision and preventing lifethreatening complications. Improving access to neuro-ophthalmologists has the potential to reduce diagnostic error.

#### **Evaluation of a Mobile Diagnostic Support Tool**

<u>P. Taber</u><sup>1</sup>, D. Borbolla<sup>2</sup>, F. Drews<sup>2</sup>, H. Kramer<sup>2</sup>, T. Taft<sup>2</sup>, P. Ghani<sup>2</sup>, C. Weir<sup>2</sup> <sup>1</sup>Veterans Healthcare System Salt Lake City, Salt Lake City, UT <sup>2</sup>University of Utah, Salt Lake City, UT

**Background:** Mobile diagnostic decision support applications for clinicians are thought to promise improved accuracy while fitting easily within modern clinical workflows. We evaluated effectiveness and usability of a mobile diagnostic clinical decision support (CDS) application. Application design criteria included: 1) putting the application on a mobile platform to improve usability at the point of care; 2) integrating diagnostic CDS with electronic health records to better match clinicians' workflow; and 3) reducing the need for data entry.

**Methods:** *Design.* We used a 4 (within) by 2 (between) design. Nurse practitioners (NPs) were randomly assigned to use the application or the internet for each vignette. Order of vignettes was randomized. We used two easy and two difficult validated clinical vignettes. Diagnostic difficulty was validated through expert consensus. *Participants.* The study included 18 NPs with one to seven years of experience. *Dependent variables* included: 1) time to complete diagnosis, 2) proportion correct diagnosis on differential, 3) perceptions of self-confidence, 4) NASA Task Load Index (TLX), and 5) the System Usability Survey (SUS). Qualitative feedback was elicited via semi-structured interviews.

Results: We used a repeated measure ANOVA, controlling for years of experience. Results for our first four dependent variables are presented in Table 1. <u>Time to complete</u> a diagnosis was significantly longer across all vignettes. <u>Accuracy in diagnosis</u> was not improved significantly by the application. The application improved accuracy more for harder cases. The TLX indicated that the application increased perceived time and effort for one easy diagnosis, and that it increased frustration for both difficult cases. The SUS showed the application to have low usability based on national norms, although <u>intentions to use</u> (mean = 4.11, SD = 3.61) and <u>perceived effectiveness</u> (mean = 3.39, SD = 1.15) were more moderate suggesting openness to using the mobile application if certain design features were improved. Qualitative responses highlighted four themes about the application: 1) it is hard to learn; 2) NPs had a hard time finding symptoms in it; 3) it helped them think; and 4) it might be useful for complex cases. **Conclusion:** The application negatively impacted time, effort and frustration in diagnosis. Our results suggest that mobility and continuous feedback from the application are not sufficient to guarantee a successful diagnostic tool. Additionally, mobile applications are not likely to be equally useful for all diagnostic problems. If usability issues can be solved, mobile applications like the one evaluated here are likely to be the most useful with difficult diagnostic cases.

Table 1. Results of mobile diagnostic decision support application evaluation for time to completion, perception of self-confidence, proportion correct diagnosis on differential and NASA TLX.

	Gro		Statistic
	App	Control	
Time to Completion	mean (sd)	mean (sd)	F (p-value)
Appendicitis (easy)	12.78 (4.32)	6.67 (3.04)	11.95 (0.000)
Colon cancer (easy)	12.78 (3.92)	6.2 (31.92)	21.03 (0.000)
Metastatic hepatic cancer (difficul		7.44 (3.75)	15.64 (0.001)
Pheochromocytoma (difficult)	12.00 (3.72)	6.57 (1.13)	10.75 (0.005)
Perception of Self-Confidence	6 70 (1 20)	6 67 (2 74)	NS
Appendicitis (easy)	6.78 (1.30) 7.38 (2.45)	6.67 (2.74) 7.7 (1.25)	NS NS
Colon cancer (easy) Metastatic hepatic cancer (difficul		6.71 (2.34)	NS NS
Pheochromocytoma (difficult)	6.82 (1.7)	6.71 (2.34)	NS NS
, , ,			
Proportion Correct Diagnosis	%	%	X <sup>2</sup> (p-value)
Appendicitis (easy)	44	89	4.00 (0.046)
Colon cancer (easy)	80	75	NS
Metastatic hepatic cancer (difficul		78	NS
Pheochromocytoma (difficult)	10	0	NS 5 (5 115/115)
NASA TLX Mental Demand	mean (sd)	mean (sd)	F (p-value)
Appendicitis (easy)		12.00 (6.083)	
Colon cancer (easy)	11.5 (6.824)		NS
Metastatic hepatic cancer (difficul			NS
Pheochromocytoma (difficult)		12.86 (6.517)	
NASA TLX Physical Demand	mean (sd) 2.22 (1.302)	mean (sd)	F (p-value)
Appendicitis (easy)			NS NS
Colon cancer (easy) Metastatic hepatic cancer (difficul	5.25 (5.922)		
			NS NS
Pheochromocytoma (difficult)  NASA TLX Temporal Demand	6.45 (6.758) mean (sd)	mean (sd)	F (p-value)
Appendicitis (easy)	10.44 (4.902)		3.67 (0.07)
Colon cancer (easy)	10.44 (4.902)		3.67 (0.07) NS
Metastatic hepatic cancer (difficul			NS
Pheochromocytoma (difficult)	9.64 (6.169)		NS NS
NASA TLX Performance	mean (sd)	mean (sd)	F (p-value)
Appendicitis (easy)	9.22 (4.236)	, ,	NS
Colon cancer (easy)	10.75 (4.027)		NS
Metastatic hepatic cancer (difficul			2.99 (0.10)
Pheochromocytoma (difficult)		10.71 (5.469)	
NASA TLX Effort	mean (sd)	mean (sd)	F (p-value)
	12.33 (3.841)		2.94 (0.10)
Colon cancer (easy)	10.63 (5.78)		NS
Metastatic hepatic cancer (difficul			6.28 (0.023)
Pheochromocytoma (difficult)		12.43 (3.552)	
NASA TLX Frustration	mean (sd)	mean (sd)	F (p-value)
Appendicitis (easy)		10.56 (4.978)	
Colon cancer (easy)	11.63 (4.955)		NS
Metastatic hepatic cancer (difficul			
Pheochromocytoma (difficult)		8.14 (5.08)	
· · · · · · · · · · · · · · · · · · ·		(0.00)	(0.010)

#### Potential Usefulness of Virtual Reality Simulation for Learning Clinical Reasoning in Japan

T. Watari<sup>1</sup>, Y. Tokuda<sup>2</sup>, M. Owada<sup>3</sup>, H. Kanda<sup>4</sup>, S. Okazaki<sup>5</sup>, K. Onigata<sup>6</sup>

<sup>1</sup>Shimane University Hospital, Shimane, Japan

<sup>2</sup>Okinawa Muribushi Project for Teaching Hospitals, Urasoe, Okinawa, Japan

<sup>3</sup>Toho University Oomori Hospital, Tokyo, Japan

4Shimane University, Izumo-shi, Japan

Shimane University Hospital, Izumo, Japan

<sup>6</sup>Shimane University Hospital, Shimane prefecture, Japan

Background: Clinical education of Japan has mostly been conducted with didactic lecture style in a large room without interactive sessions. Little education for symptomatology has been provided. In addition, there have been few teachers who specialize in clinical education, medical students have the limited chance to learn clinical reasoning. Since virtual reality simulation (VRS) has recently introduced as a novel learning strategy in medicine, VRS would also be useful for effectively learning clinical reasoning with minimum support by teachers. The purpose of this study was to examine the effects of VRS on outcome and satisfaction of learning clinical reasoning among medical students. Methods: Total of 210 fourth-grade medical students who had passed Objective Structured Clinical Examination were enrolled in a two-year period (2017 and 2018). A VRS software program (Body Interact, Portugal) was used for 2-hour regular classes in a longitudinal experimental design. The learning outcome was measured using a pre-post 20-item multiple-choice questionnaire (10 knowledge items, 10 clinical reasoning items), which had been selected from the Japanese National Medical Practitioners Qualifying Examination with real-time audience response system (®Turning point clicker, Tokyo, Japan). Repeated measurements were assessed for statistical analyses. McNemar's test was used for paired nominal data, while Wilcoxon signed-rank test was employed for non-parametric repeated measurements with Bonferroni correction. Results: A total of 169 students, who provided informed consent, completed the program in this study (19 absent, 13 drop-out, 9 online connection trouble). Compared with pre-test baseline scores, participating students showed significant increases in the average total post-test score (mean  $scores \pm standard$  deviation of post- vs. pre-test,  $12.97 \pm 2.65$  vs.  $10.1 \pm 2.80$ , p-value < 0.001), the score of knowledge part  $(5.17 \pm 1.77 \pm 1.7$ vs.  $4.79 \pm 1.54$ , p-value 0.028), and the score of clinical reasoning part (7.81  $\pm 1.59$  vs.  $5.28 \pm 1.97$ , p-value <0.001).

**Conclusion:** Clinical reasoning scores among medical students increased by a new virtual reality simulation program. VRS software programs could help medical students improve their clinical decision-making skills with minimum supervision of lecturers. The widespread use of VRS software programs in clinical education should be considered.

#### Malpractice Claims Related to Diagnostic Errors in Japan

T. Watari<sup>1</sup>, S. Mitsuhashi<sup>2</sup>, K. Kono<sup>2</sup>, N. Nagai<sup>2</sup>, Y. Tokuda<sup>3</sup>, S. Okazaki<sup>4</sup>, K. Onigata<sup>5</sup>, H. Kanda<sup>2</sup>

<sup>1</sup>Shimane University Hospital, Shimane, Japan

<sup>2</sup>Shimane University, Izumo-shi, Japan

<sup>3</sup>Okinawa Muribushi Project for Teaching Hospitals, Urasoe, Okinawa, Japan

<sup>4</sup>Shimane University Hospital, Izumo, Japan

Shimane University Hospital, Shimane prefecture, Japan

Background: Little is known about the incidence or significance of diagnostic error in Japan. We used the malpractice claims database to examine diagnostic error-related malpractice claims in Japan. We aimed to characterize the frequency, health outcomes, and economic consequences of diagnostic errors in Japan.

Methods: The largest malpractice claims database in Japan (Westlaw Japan K.K) was used to identify error types (diagnostic error, problem of procedure or operation, management problem), outcome severities, and payments in the malpractice claims closed between 1961 and 2016 (accessed on June 29th, 2017). The exclusion criteria were duplication of the case, intentional crimes, robbery, money troubles, and veterinary claims.

Results: Of 1345 medical malpractice claims analyzed, 772 cases (57.4%) were accepted. The median age of the patients was 36 years (interquartile range (IQR) 14-58), while the mode of the patients' age was 0 years (6%). The patients comprised 675 men (50.19%) and 601 women (44.68%). The medical outcomes were deaths (691, 51.38%), sequelae (453, 33.68%), complete recovery within 12 months (66, 4.91%), and complete recovery after a year (11, 0.82%). According to the classification of subspecialties in Japan, 290 cases (21.56%) were of internal medicine, followed by 253 cases of surgery (18.81%), 240 cases of obstetrics and gynecology (17.84%), 97 cases of orthopedic surgery (7.21%), 86 cases of pediatrics (6.39%), 76 cases of neurosurgery (5.65%), and 54 cases of dentistry (4.01%). The clinical settings of the medical claims were: clinic-based, 327 cases (24.16%); small-scale hospitals, 262 cases (19.48%); medium-scale hospitals, 457 cases (33.98%); and large-scale and university hospitals, 286 cases (21.26%). In total, 276 cases (20.52%) were in outpatient clinics and office rooms, 347 cases (25.80%) were in wards, 72 cases (5.35%) involved emergency outpatients; procedures rooms such as surgery, treatment rooms, and endoscopic rooms accounted for 616 cases (45.8%) at the site concerning the case and 30 at other sites (2.23%). In total, 319 of 1345 malpractice claims (23.72%) were diagnostic error-related cases, and 874 (64.98%) were system error-related claims. Compared with system errors, diagnostic errors were associated with a 1.33-times greater risk of death (95% confidence interval 1.20-1.48; P < 0.001). The median claims payments of diagnostic error-related and non-related claims were \$37471 (IQR \$0-211118) and \$9280 (IQR \$0-154107) (P<0.001), respectively.

Conclusion: Among the malpractice claims database in Japan, diagnostic errors were not rare causes of medical claims and were relatively common. Diagnostic error-related cases were associated with the risk of death and higher claims payments.

## **Poster Session 1: Clinical Vignettes**

Sunday, November 4, 2018 | 4:15 PM - 5:45 PM

#### Do Not Miss This Diagnosis: Discitis

S. M. Ahmed<sup>1</sup>, Z. Qamar<sup>1</sup>, S. Godil<sup>2</sup>, J. Akhtar<sup>3</sup> <sup>1</sup>Dow University of Health Sciences, Karachi, Pakistan <sup>2</sup>UPMC Mercy, Pittsburgh, PA <sup>3</sup>UPMC, Pittsburgh, PA

Learning objectives: Infectious spondylodiscitis is a diagnostic and therapeutic challenge because the clinical signs and symptoms caused by spinal infections are usually subtle. The correct diagnosis can be missed or delayed because of atypical presentation. Abdominal pain as a presenting symptom is a rare diagnostic clue for a spinal infection.

Case information: A 28 year-old woman with a past history of intravenous drug was transferred from an outside hospital on 6/11/2018 due to concerns for thoracic discitis. She reports that she has had intermittent right upper quadrant (RUQ) tenderness since Feb 2018. Pain started all of a sudden and is constant and sharp and non-radiating. Worsened by movement. Not affected by food. No associated n/v/d or fevers and chills. She had presented to a tertiary care center in May of 2018 with this complaint when CT abdomen/pelvis (A/P) was unremarkable, and was discharged home. Pain occurred several times since then. Now the RUQ pain was sharp/constant and worsened over last 2 days. No weakness, sensory loss, or bowel/bladder incontinence. She presented with these symptoms to the outside hospital. There the CT A/P was unremarkable but found thoracic discitis. MRI T-spine revealed T10-T11 discitis and and abnormal soft tissue concerning for multiloculated abscess formation. A review of the CT A/P done in May showed early changes of T10-T11 discitis, and was missed. She was treated with antibiotics and underwent surgery.

Discussion: Spinal infections requires prompt diagnosis and treatment to prevent serious neurological complications. When the common causes of RUQ pain have been excluded, further imaging is justifiable given the severe and lasting neurological effects that a spinal infection can cause. Physicians should maintain a high index of suspicion for this disease when treating patients with chronic abdominal pain in the emergency department.

#### Misdiagnosis of Excited Catatonia in an Adolescent with Autism

N. Bihani<sup>1</sup>, R. Klisz-Hulbert<sup>2,3</sup>, D. L. Levine<sup>4</sup>

<sup>1</sup>Wayne State School of Medicine, Detroit, MI

<sup>2</sup>Wayne State School of Medicine, Livonia, MI

<sup>3</sup>Detroit Medical Center

<sup>4</sup>Wayne State University, School of Medicine; Detroit Medical Center, Detroit, MI

#### Learning objectives:

- Describe Diagnostic Criteria for Excited Catatonia
- Define Treatment of Excited Catatonia
- Identify ways to avoid misdiagnosis of Excited Catatonia

Case information: A 13-year-old boy with established diagnoses of autism and intellectual disability presented with marked stereotypy, restlessness, impulsivity, frenzy, agitation, combativeness, and purposeless movements. He was given a small dose of lorazepam for sedation. He remained agitated and was labeled as having a "paradoxical reaction" to lorazepam. He was subsequently diagnosed with mania, and given quetiapine. Initially it appeared effective; he was discharged. Within one week, symptoms returned with marked aggression and agitation. He was re-admitted, titrated off quetiapine, and prescribed valproic acid. He calmed, but was withdrawn and irritable at discharge. At age 15, he developed hyperammonemia secondary to valproic acid. After discontinuation, many of his behavioral issues resolved. Recently, he experienced another episode of behavioral symptoms at age 22, manifesting with motor and verbal tics, excessive adherence to rules and routines, and increased impulsiveness. Psychiatry diagnosed excited catatonia and prescribed lorazepam. Within two weeks, Busch Francis Catatonia Rating Scale (BFCRS) decreased from 26 to 14. He did not have a mood disorder but rather episodes of excited catatonia due to his underlying diagnosis of autism.

Discussion: This case exemplifies premature closure error (PCE). Diagnosis was based on pattern recognition. Although PCE generally occurs when physicians assume the patient is having a common complication of a known illness, the opposite occurred in this case. Rather than considering catatonia as a complication of autism, providers diagnosed a second psychiatric (mood) disorder, possibly due to knowledge gaps (McGee, 2015). Catatonia occurs in 12-18% of patients with autism (Wing and Shah, 2000) and does not always involve stupor, mutism, and slowed motor activity—but can present with excitement. BFCRS is useful for quantifying symptoms. Diagnosis is confirmed with lorazepam challenge (Dhossche et al, 2006).

#### The Digital Rectal Examination Still Matters

A. Elder

Stanford University, Stanford,, CA Edinburgh Medical School, Edinburgh, United Kingdom

Learning objectives: Recognize the continuing importance of physical examination in the reduction of diagnostic error

Case information: An 89 year old man with dementia who lived in a nursing home was referred to a surgical OPD because of altered bowel habit. The patient was unable to give a detailed history but carers noted a tendency to infrequent bowel movements with intermittent diarrhoea. At times he appeared to have abdominal pain. Routine blood tests were normal. No physical examination was documented in the case record by the referring doctor, or in the surgical OP clinic. The reviewing surgeon felt that intra-abdominal malignancy was possible and arranged a CT scan of abdomen and pelvis. (Figure 1). This showed extreme rectal distension due to gross faecal impaction. The patient was treated with enemas and a stool softening agent, and all his symptoms resolved.

**Discussion:** This patient experienced significant symptoms and diagnostic delay simply because none of the doctors he saw undertook a digital rectal examination. Additional expense from investigation, and inconvenience and discomfort to the patient from investigation also occurred. The association between absent or poor physical examination and diagnostic error and delay is well recognised (Verghese et al Am J Med 2015). Much of the literature on diagnostic error however focusses on errors in processing of information that has been gathered by clinicians, or in systems issues that compromise the availability or collation of relevant information. However, it remains vital to ensure that clinicians gather information, particularly from the core bedside skills of history taking and physical examination. This patient, who could not give a clear history, serves as a simple, strong and clear reminder of the need for symptom targetted physical examination, if diagnostic error or delay are to be reduced.



#### Neurogenic Pulmonary Edema: A Diagnostic Challenge

A. Kanj, N. Abdallah, H. Tabaja, L. Kao, O. Chehab, S. El Zein, J. Dbaibou, D. L. Levine, J. Weinberger Wayne State University, School of Medicine; Detroit Medical Center, Detroit, MI

Learning objectives: Physicians need to be cognizant of neurogenic pulmonary edema (NPE), a rare form of non-cardiogenic pulmonary edema, to avoid misdiagnosis and adverse outcomes.

Case information: We present two patients with NPE following a cerebellar infarct (patient A) and a generalized tonic-clonic seizure (patient B). Both patients had normal echocardiograms and a low pro-BNP. Patient A was a 78-year-old woman who presented with shortness of breath associated with headaches and dizziness. Her SpO2:92% on room air and diffuse crackles were heard over the lungs. ECG showed no signs of acute coronary syndrome. Troponins were negative and Pro-BNP:20. Chest radiography showed pulmonary edema. The patient was admitted with an inaccurate diagnosis of new-onset heart failure and was started on diuretics. Her echocardiogram was normal. A brain CT done after aggravation of her headaches showed a large cerebellar infarct with an occlusion of the left vertebral artery, a mass effect and herniation of the cerebellar tonsils. Patient B was a 33-year-old man who was receiving his scheduled hyperbaric oxygen therapy for a skin burn when he suddenly developed tonic-clonic seizures followed by desaturation, requiring intubation. His PaO2:57mmHg on FiO2:100%. Chest radiography showed pulmonary edema. ECG showed no signs of acute coronary syndrome. Troponins were negative and Pro-BNP:127. His brain CT was unremarkable. The patient's pulmonary edema was initially attributed to heart failure but his echocardiogram was normal.

Discussion: The two patients presented above developed NPE following different insults to their central nervous system. Both patients had a good outcome. The first cases of NPE were described in patients with gunshot wounds to the head. Since then, this uncommon entity has been associated with strokes, seizures, traumatic brain injuries, and intracranial hemorrhages. NPE is diagnosed clinically and requires a high level of suspicion in the absence of more plausible alternative causes of pulmonary edema.

#### A Case of Missed Diagnosis of Acute Abdomen at Various Levels of Care

P. R. Kathi<sup>1</sup>, N. Thammineni<sup>2</sup>, S. Mangat<sup>3</sup>, K. Bazzy<sup>1</sup>

John D. Dingell VA Medical Center, Wayne State University, Detroit, MI

Wayne State University, Detroit, MI

Detroit Medical Center, Detroit, MI

**Learning objectives:** Recognize the effects of anchoring bias on physicians especially trainees leading to missed diagnosis and delayed patient care.

Case information: A 75 year old male with peptic ulcer disease, diverticulosis, and ESRD presented to ED with abdominal pain and vomiting. He reported left lower abdominal pain, constipation and decreased intake started 3 days before the presentation. Initial vitals: BP 78/49mmHg, HR 59/min. Initial labs showed lactic acidosis of 4.6. He received 400cc fluids with improvement in BP to 108/80 and decrease in lactic acid to 1.3. Chest X-ray was obtained in the ED which was read as normal by the radiologist. ED physician called the medicine team with admitting diagnosis of hypotension and dehydration. Primary team requested ED to obtain abdominal X ray, but it was not obtained. Primary team admitted the patient with constipation versus PUD as the working diagnosis and signed out to night float resident. He complained of worsening abdominal pain and it was found to be tense and distended on examination in the morning. A STAT CT abdomen was obtained showing large amount of free intraperitoneal air. On review of chest X ray taken at the time of admission, there was air under the diaphragm. The patient was then taken to the OR and an exploratory laparotomy was done showing perforated diverticuli in the sigmoid colon.

**Discussion:** With increasing dependence on radiology reports, physicians are not utilizing the opportunity to review the images and interpret findings in the clinical context. This in turn leads to delayed patient care due to anchoring bias if the initial report was wrong as described here. The diagnosis in this case was missed at various levels by radiologist, ED physician, primary and night float residents. This case provides a perfect example of anchoring bias limiting thought process which resulted in an unwanted patient outcome.



#### Delay in the Diagnosis of Vasculitis

<u>T. Kimura</u><sup>1</sup>, T. Kawahigashi<sup>2</sup>, M. Tokura<sup>1</sup>, M. Yoshizwa<sup>1</sup>, J. Branch<sup>1</sup>, S. Teshima<sup>1</sup>, I. Kitagawa<sup>1</sup> <sup>1</sup>Shonan Kamakura General Hospital, Kamakura, Japan <sup>2</sup>Shonan Kamakura General hospital, Kamakura, Japan

Learning objectives: Skin biopsy can be essential for diagnosis in cases of a systemic disease.

**Case information:** A 62-year-old woman presented to our hospital with a progressive rash. Two weeks prior to admission she noticed diffuse erythematous papules on her body and paresthesias of her lower limbs. Subsequently, she developed painful swelling of the small joints in her hands and feet. The paresthesias also worsened to encompass her finger tips and the soles of her feet. A local doctor prescribed an ointment suspecting an allergic reactions. On examination, variably sized erythematous papules with erosions associated with crusting were observed over her entire body excluding her face. Arthritis of her wrists, knees, and small joints of her hands and feet were observed. There

was no obvious mucosal involvement. Laboratory studies showed elevated leukocytes and C-reactive protein. Immunological studies were negative, including antinuclear antibodies and antineutrophil cytoplasmic antibodies. Serum IgA levels, urinalysis and blood culture results were unrevealing. A skin biopsy was performed, which showed leukocytoclastic vasculitis (LCV). Based on her pathological and laboratory findings, and without specific history to indicate a cause such as medications, a diagnosis of unclassified systemic small vessel vasculitis was made. Intravenous methylprednisolone 1000mg pulse therapy once daily for three days was performed. This was followed by oral prednisolone, which was tapered gradually. Corticosteroid therapy achieved a dramatical response. Three weeks later, her pain, arthritis and paresthesias had almost completely resolved.

**Discussion:** LCV is a rare type of small-vessel vasculitis. The diagnosis of LCV is often challenging due to its rarity and various presentations, as this case showed. In addition, premature closure lead to delay in the diagnosis in this case. A skin biopsy was a useful diagnostic tool in this case. When encountering cases that show a variety of systemic symptoms along with skin involvement, a skin biopsy can be essential for the diagnosis.



#### **Case of Chest Pain and Missed Coronary Artery Laceration**

D. Mayer

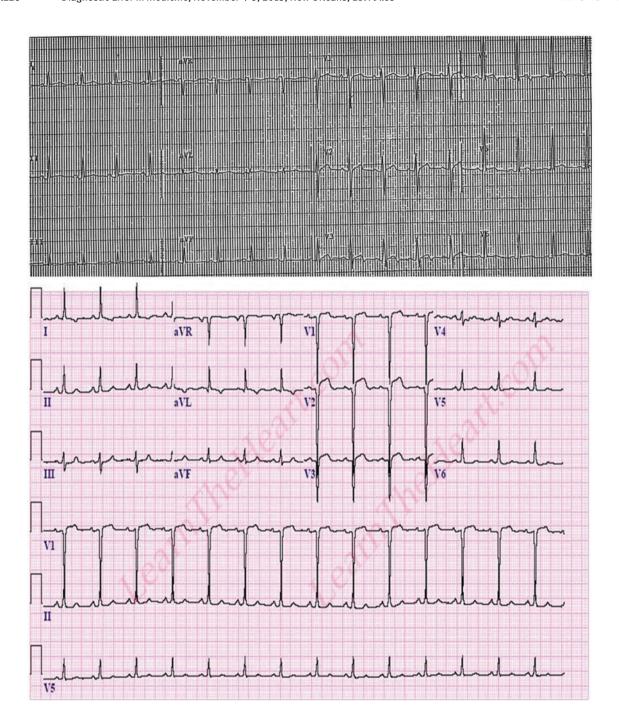
Retired from Albany Medical College, Niskayuna, NY

#### **Learning objectives:**

- 1. Recognize the difference between System 1 and System 2 errors
- 2. Identify the type of cognitive error made in the case.
- Identify how the learner can be taught to avoid this error.

Case information: A 34-year-old man was suicidal and stabbed himself in the left chest. He changed his mind, was taken to a local hospital, cleared surgically, and transferred to a psychiatric in-patient service. One week later he complained of severe sub-sternal chest pain. Except for a slight tachycardia of 108, his vital signs were normal. The medical team saw the patient with the cardiology fellow. An EKG showed diffuse ST elevation and PR depression. The physical examination was normal with a healing stab wound on his left anterior chest with minimal tenderness around the wound. A chest x-ray was done but not read. The team gave the patient t-PA for an acute myocardial infarction. Shortly after that he complained of increased pain and light-headedness. His blood pressure dropped then became unobtainable. He was unresponsive and bedside ultrasound showed a large pericardial effusion with poor ventricular contractility. A pericardiocentesis was attempted but he developed pulseless electrical activity, then asystole and could not be resuscitated. Post mortem revealed hemopericardium from a coronary artery laceration.

Discussion: Biases were premature closure of the differential diagnosis and incorrect prior probability. The consideration of coronary artery disease in a young man with chest trauma considered a disease with low prior probability more likely than diseases caused by trauma (pericarditis, cardiac tamponade or pneumothorax). Seeing the abnormal EKG led to diagnostic momentum in the direction of myocardial infarction rather than pericardial disease and effusion. Discrepancies in the initial diagnosis and re-evaluation to work through a differential diagnosis ruling out worst case scenarios. The chest x-ray was not read although it was later determined to be abnormal.



#### **Case of Missed Puerperal Sepsis**

#### D. Mayer

Retired from Albany Medical College, Niskayuna, NY

#### **Learning objectives:**

- 1. Recognize the difference between System 1 and System 2 errors
- 2. Identify the type of cognitive error made in the case.
- 3. Identify how the learner can be taught to avoid this error.

Case information: A 35-year-old black woman delivered her fourth child after an uneventful pregnancy, labor and delivery. She was observed until she urinated and defecated without problem. Two days later, she came to the emergency department with cramping and severe lower

abdominal pain, poor appetite, fever, nausea and vomiting. She had no dysuria or other pain. On physical examination her lower abdomen was tender with voluntary guarding. Pelvic examination showed vaginal blood and uterus "too tender to examine". White blood cell count was 2,300 with 15% neutrophils and 70% bands. Clean catch urine showed elevated specific gravity and too numerous to count white and red blood cells. Diagnosis was urinary tract infection and Bactrim was prescribed. Next day she felt worse, had a syncopal episode, returned to emergency with unobtainable blood pressure and had cardiopulmonary arrest. She could not be resuscitated. Post mortem showed severe endometritis from Group A Hemolytic Streptococcus with positive blood cultures.

Discussion: Biases were premature closure of the differential diagnosis. The urinalysis anchoring the diagnosis was improperly obtained using a clean catch despite post partum bleeding. The white blood cell count was not investigated. Diagnosis momentum, search satisficing and confirmation biases reflected the tendency to "stick" to the initial diagnosis. Initial finding of very tender uterus consistent with endometritiswas ignored. Stereotyping may have been present as her husband later said he felt the staff was trying to "get rid of them" and treated them as ignorant poor people. A broader differential required obtaining catheterized urine. Discrepancies in the initial diagnosis and re-evaluation to work through a differential diagnosis ruling out worst case scenarios.

#### Affective Bias Can Have a Big Effect: A Delayed Diagnosis of Malignant Catatonia

H. Pi1, M. Fleshner1, S. Tilstra2 <sup>1</sup>University of Pittsburgh Medical Center, Pittsburgh, PA <sup>2</sup>University of Pittsburgh School of Medicine, Pittsburgh, PA

#### **Learning objectives:**

- Recognize common triggers for acute catatonia
- Understand the role of framing bias, premature closure, and affective bias on diagnosis in patients with psychiatric illnesses

Case information: A 21-year-old with schizophrenia presented with diarrhea and vomiting. He was drowsy but arousable with an otherwise normal exam. Patient was treated with IV fluids and ondansetron for gastroenteritis and discharged home. Patient represented two days later with altered mental status and agitation. He was admitted for "dehydration/gastroenteritis" in the setting of schizophrenia. After admission, he became diaphoretic, tachycardic, and unresponsive to sternal rub. He had mild rigidity and contraction of bilateral hands. Laboratory studies revealed a CPK of 531. His mother reported giving patient scheduled ondansetron in the setting of increased risperdone dose prescribed by patient's psychiatrist. 1 mg IV lorazepam was administered out of concern for malignant catatonia with robust response in patient's mental status and vitals. Psychiatry later confirmed this diagnosis.

Discussion: Catatonia is an acute psychomotor depression with full physical capacity. Common triggers include dopamine blocking medications and dehydration/acute illness. Elevated CPK, autonomic instability, and fever increase concern for neuroleptic malignant syndrome or malignant catatonia. Framing bias, premature closure and affective bias all played roles in this delayed diagnosis. This case was framed by the ER and the admitting team prematurely closed on their diagnosis of "dehydration/viral illness." Patient's altered mental status was not further investigated and attributed to his psychiatric illness, despite the fact that he was high-functioning and oriented at the prior admission. Awareness of these biases, particularly affective bias, may help to broaden the differential in patients with mental illness and prevent anchoring on psychiatric illness as the root cause of altered mental status and agitation. Caution for use of dopaminergic drugs like ondansetron should be raised in this patient population to minimize risk of adverse events and drug interactions.

#### Pyuria, Diagnostic Momentum, and Intra-Abdominal Sepsis in the Elder

M. Pujara<sup>1</sup>, L. Bettner<sup>2</sup>, J. N. Lessing<sup>3</sup> <sup>1</sup>University of Colorado Hospital, Denver, CO <sup>2</sup>University of Colorado Hospital University of Colorado, Denver, CO

#### **Learning objectives:**

- Recognize that intra-abdominal infections in the elderly often present atypically and warrant a lower threshold for abdominal imaging.
- Use pyuria with equivocal urine culture results as an opportunity to reopen the diagnosis of an infectious source in sepsis.

Case information: An 86 year-old woman presented with leukocytosis and acute kidney injury found on pre-operative labs for elective aortic valve replacement. Aside from fatigue, she was asymptomatic without localizing signs of infection. Examination was notable for temperature 35.9°C, aortic stenosis murmur, non-infectious pressure ulcers, and an indwelling urinary catheter placed 1 month prior for urinary retention attributed to polymicrobial urinary tract infection (UTI) with documented normalization of her urinalysis after treatment. Admission urinalysis showed >75 Leukocytes, 11-30 Red Blood Cells, positive Leukocyte Esterase and negative Nitrites. She was empirically treated for suspected catheter-associated UTI. On hospital day (HOD)1 she developed hypotension unresponsive to fluids. Antibiotics were broadened upon ICU

transfer for vasopressor support. Urine cultures grew 10,000-100,000 CFU of multiple bacteria morphotypes. On HOD4 she developed new abdominal pain and ileus; CT Abdomen-Pelvis showed a 16.7cm x 19.0cm intra-abdominal abscess with bowel perforation. Given poor surgical candidacy, an IR drain was placed from which fluid cultures grew Pseudomonas Aeruginosa. On HOD7 the patient and family opted for comfort care.

Discussion: Intra-abdominal infections are an underappreciated etiology of sepsis in elder populations. This case highlights how an atypical presentation and diagnostic momentum due to pyuria resulted in a delayed diagnosis. Many factors contribute to the increased rates of mortality in older patients with intra-abdominal sepsis, among which delays are potentially avoidable with a lower threshold for abdominal imaging. Furthermore, this patient's urine culture results suggesting colonization rather than active infection could have prompted reevaluation for an alternate etiology of septic shock, which did not occur until the development of catastrophic abdominal symptoms.

#### Missed Aortic Root Dissection in a 44-Year-Old Male

T. Riveros<sup>1</sup>, T. Hedayati<sup>1</sup>, K. Cosby<sup>1,2</sup> <sup>1</sup>Cook County Health and Hospitals System, Chicago, IL <sup>2</sup>Rush Medical College, Chicago, IL

Learning objectives: Identify signs and symptoms associated with aortic root aneurysm. Recognize that atypical presentations combined with premature closure can have catastrophic consequences.

Case information: A 44-year-old man with hypertension presented to the ED with epigastric pain radiating to the anterior chest, nausea, vomiting, diarrhea, subjective fever and myalgias after eating at a restaurant. His mother also had diarrhea. The patient reported a syncopal episode while vomiting prior to arrival. Vital signs were significant for blood pressure of 104/53 mmHg and heart rate of 53 bpm. Labs revealed a leukocytosis of 12.6 k/uL, normal hemoglobin, non-elevated troponin, and creatinine of 1.9 mg/dL. EKG was sinus bradycardia with LVH. CXR showed a tortuous agrta and no cardiomegaly. After 2L of normal saline, blood pressure improved to 134/66 mmHg. The patient was admitted for syncope and acute kidney injury to a telemetry unit. A systolic ejection murmur at the LICS was noted by the admitting team. The following morning the patient had vomiting, epigastric pain now radiating to the back, and elevated blood pressure to 202/101 mmHg, Labs demonstrated increasing leukocytosis, improved creatinine, and stable hemoglobin. Several hours later the patient was found unresponsive. Resuscitation was unsuccessful. Autopsy results revealed an aortic root dissection with cardiac tamponade.

**Discussion:** Aortic dissection (AoD) is a condition with a wide range of presentations and high mortality. In a registry of more than 2000 patients with AoD, younger patients were less likely to be hypertensive on presentation and more likely to have proximal dissections affecting the sinuses of Valsalva which may result in bradycardia. An abnormal CXR is present in 90% of patients with AoD. In this case the ED focused on dehydration and acute kidney injury; the inpatient service focused on gastroenteritis. Over-simplifying patterns of symptoms and premature closure combined with an atypical presentation led to this diagnosis failure.

#### **Misdiagnosis of an Aortic Dissection**

T. Sakamoto, Y. Harada, S. Kakimoto, H. Yamamoto, T. Shimizu Dokkyo Medical University Hospital, Shimotsuga, Japan

Learning objectives: To recognize the importance of validating the final diagnosis with diagnostic criteria

Case information: A 67-year-old man presented to our emergency department with a five-day history of diarrhea, malaise, and arthralgia. He was a heavy drinker, and his past medical history was remarkable for right thalamic hemorrhage with left hemiparesis and hypertension. The patient was in mild distress. His vital signs were stable. Chest exam showed no murmur or crackles. Abdominal exam revealed mild tenderness over the periumbilical area, but otherwise unremarkable. Laboratory studies revealed serum amylase level of 179 U/L (upper limit of normal: 132 U/L). Abdominal computed tomography (CT) without contrast showed peripancreatic fat stranding and mild pericardial effusions. The patient was suspected with acute pancreatitis. However, given he did not fulfill the diagnostic criteria of acute pancreatitis, the gastroenterology consultants recommended additional abdominal contrast-enhanced CT to evaluate pancreatitis. The CT scan coincidentally revealed abdominal aortic dissection to lateral common iliac artery, which extended from the ascending aorta that elucidated by subsequent chest contrast-enhanced CT. The diagnosis was confirmed with Stanford A aortic dissection. He was initially treated conservatively due to co-existing chronic subdural hematoma followed by surgical treatment, which ended up with successful recovery.

Discussion: The present case was confounded by availability and confirmation biases. The patient's abdominal symptoms with heavy alcohol drinking could readily be attributed to acute pancreatitis. In addition, abdominal tenderness, mildly elevated serum amylase level, and peripancreatic fat stranding might have anchored to the incorrect diagnosis. At this point, premature closure overshadowed clinical information which were not consistent with acute pancreatitis such as pericardial effusion and absence of typical CT signs of pancreatitis. This case underscores the importance of validating the final diagnosis with clinical diagnostic criteria even when most of findings seem to be consistent with the initial diagnosis.

#### Missed Coarctation of the Aorta in a Neonate with Multiple Congenital Abnormalities

G. J. D. C. Shafer<sup>1</sup>, G. K. Suresh<sup>2</sup>

<sup>1</sup>Baylor College of Medicine & Texas Children's Hospital, Houston, TX

<sup>2</sup>Baylor College of Medicine & Texas Children's Hospital

**Learning objectives:** Describe how a 'red-flag' for a serious diagnosis can be missed by busy clinicians in the presence of information overload for neonates with complex conditions.

Case information: An infant was born at 37 weeks gestation by cesarean section. Prenatal evaluation showed multiple congenital anomalies including: bilateral cerebral ventriculomegaly with mass effect on the cerebellum, caudal regression spectrum, kyphosis and scoliosis. A fetal echocardiogram (echo) revealed a ventricular septal defect and biventricular hypertrophy. The cardiologist recommended an echo after delivery. On admission to the Neonatal Intensive Care Unit (NICU), blood pressure measurements in all four limbs revealed a difference in the systolic blood pressure between the upper and lower limbs of 30 mmHg, which was documented in the electronic medical record. This abnormality – a 'red flag' for a ortic coarctation – was not recognized by either the neonatology or cardiology clinicians. As the patient was otherwise hemodynamically stable after being intubated, a postnatal echo was deferred until the following day, while imaging of the head and kidneys was performed that night. An echo at 22 hours of life was suspicious for aortic coarctation. The patient was started on a prostaglandin E1 infusion, and a computed tomography angiography scan confirmed critical aortic coarctation.

Discussion: A large differential between upper and lower extremity systolic blood pressure measurements – a well-known indicator of aortic coarctation - was not recognized by the medical team, likely because it was not present on fetal echo, and the team prioritized diagnostic evaluations of non-cardiac organs on admission. Neonates with numerous congenital abnormalities present a diagnostic challenge in the perinatal period, when multiple defects may need to be evaluated emergently. This case illustrates the importance of not overlooking critical details during the often-hectic postnatal stabilization period.

### "What Is Essential Is Invisible." Lessons from Near-Miss Diagnosis in a Case of Severe Unilateral Abdominal Pain with Unknown Origin

T. Suzuki<sup>1</sup>, Y. Tokuda<sup>2</sup>

<sup>1</sup>Juntendo University Hospital, Tokyo, Japan

<sup>2</sup>Okinawa Muribushi Project for Teaching Hospitals, Urasoe, Okinawa, Japan

#### Learning objectives:

Identify an extra-abdominal cause of abdominal pain presented as referred pain.

#### Case information:

A 50-year-old man presented to the emergency department (ED) with a severe (9/10) abdominal pain in right mid-back and lower quadrant, which suddenly developed when he sat up from supine position. The pain became worse by sitting or standing positions. He did not have rash, vomiting or diarrhea. Vital signs were normal and abdominal examination showed no abnormal findings. There was no tenderness over the back. Blood and urine test were unremarkable. Contrast-enhanced abdominal CT did not show any findings explaining his symptom. At the time when ED physician was about to discharge, the physician referred the patient to a generalist diagnostician. The consultant physician asked about history of the similar pain, revealing that he had the same kind of pain ten days ago, developed when his dog hopped on the genital area, and it relieved within two days. Subsequent genital examination revealed tenderness on right testis and ultrasonography showed tear of tunica albuginea. A diagnosis of testicular torsion with referred pain on the abdomen was made. The patient was referred to urologist's observation and recovered without complication.

#### Discussion:

This case represents a near mistake of diagnosis or a salvaged diagnostic error. The following lessons can be learned: 1. Detailed historytaking can reveal a crucial diagnostic clue despite negative imaging studies regarding avoidance of diagnostic error; 2. physicians should consider extra-abdominal causes, including genital area (invisible if not examined thoroughly), in those of abdominal pain with unknown origin. The overconfidence on negative imaging results may lead to diagnostic error, like the current case. Patients in ED are at high risk for diagnostic errors, but these errors could be reduced by the cross-checking between emergency physicians and generalist physicians, thus lead to the reduction of adverse events in ED by avoiding near misses.

#### Ockham's Razor Is Still Sharp

S. Tobe, K. Nakano, K. Akazawa Shonan Fujisawa Tokushukai Hospital, Fujisawa, Japan

Learning objectives: Heredity angioedema (HAE) is a rare disorder caused by deficiency or dysfunction of the C1 inhibitor, and presents recurrent attacks in skin, bowel, and upper air way. The symptoms are swelling without hives or pruritus, gastrointestinal colic, nausea,

vomiting, diarrhea, and/or laryngeal edema. If the patient only presented gastrointestinal symptoms, we may diagnose it as gastroenteritis. We propose to list HAE in the differential diagnosis of recurrent gastroenteritis

Case information: The case is 36 years old man with chief complaint of abdominal pain. He has visited ER numerous times and have been admitted as gastroenteritis 5 times in the past 10 years. Suspected of Crohn disease or eosinophilic enteritis, he has undergone gastrofiberscopy and colonfiberscopy many times. Fiberscopy showed no findings and symptoms disappear spontaneously. 2 days before admission, he had upper abdominal pain and watery diarrhea. On the day of admission, he came to our emergency department with worsening abdominal pain. After the admission, we performed colonscopy and found edematous intestinal mucosa. Further history taking revealed that he has had transient edema of extremities. History of edema was seen in our past medical records many times. Genetic testing revealed that he has C1 deficiency which lead to the diagnosis of HAE.

Discussion: HAE presents symptoms that are not severe and disappear spontaneously. Considering the rarity of HAE, it is difficult to link those symptoms to the disease. But as William of Ockham suggests in Ockham's Razor, we should try to explain the symptoms in one simple answer. Had we noticed the link between the symptoms, the patient may have lived with more satisfaction during the lost 10 years.

#### What Do You Think When You See Hyperamylasemia?

Iizuka Hospital, Iizuka, Japan

Learning objectives: In the differentiation of hyperamylasemia, it is necessary to consider ectopic amylase-producing tumors as well as salivary gland diseases and pancreatic diseases.

Case information: A 78-year-old woman was admitted to our department by a nearby doctor due to hyperamylasemia. Physical examination demonstrated no parotid swelling or abdominal tenderness, and there were no noteworthy findings. Blood test showed that amylase was as high as 768 U/L, and in amylase fractionation, salivary gland type was 95% confined. Antinuclear antibody, anti-SS-A antibody, and IgG4 were both below the reference values. There was no salivary gland swelling by echo, and Sjogren's syndrome and IgG4-related diseases were absent. Although it was taken as from an outpatient follow-up policy, cough and chest pain occurred three months later. When chest X-ray was taken, there was decreased permeability of the upper right lung field. Chest CT showed tumor and pleural effusion with atelectasis in the right upper lobe, and adenocarcinoma cells were detected in pleural effusion cytology. Based on these findings, I finally diagnosed him with lung adenocarcinoma.

Discussion: The diagnosis was delayed since I did not know of the disease concept of ectopic amylase-producing tumor, and I thought it would be feasible if I confirmed that the salivary glands had no problem, except that the amylase value was increased. In ectopic amylaseproducing tumors, increases in amylase levels are due predominantly to a salivary gland type of amylase, with a particularly high incidence of lung cancer and ovarian cancer. Amylase-producing lung cancer accounts for 1-3% of total lung cancers, most of which are adenocarcinomas, and it also occurs in many cases with poor prognosis with carcinomatous pleurisy or distant metastasis at the time of diagnosis. In the case of asymptomatic salivary gland type amylase elevation, the possibility of ectopic amylase-producing tumors should be considered.

#### **Does This Antibiotics Truly Effective?**

T. Yabuki<sup>1</sup>, R. Ozawa<sup>2</sup>, Y. Tokuda<sup>3</sup>

<sup>1</sup>National Hospital Organization, Tochigi medical center, Utsunomiya, Japan

<sup>2</sup>National Hospital Organization, Tochigi medical center, Tochigi, Japan

<sup>3</sup>Okinawa Muribushi Project for Teaching Hospitals, Urasoe, Okinawa, Japan

Learning objectives: Recognize that Urinary Tract Infection(UTI) can be wastebasket diagnosis for elderly people living in nursing home. Don't prescribe Fluoroquinolone(FQ) easily in Tuberculosis(TB) endemic countries such as Japan.

Case information: An 81-year-old man living in nursing home presented to our hospital for evaluation of difficulty in oral intake. His past medical history was remarkable for dementia and cystostomy for neurologic bladder one year ago. After cystostomy he was diagnosed with "UTI" frequently, and he took levofloxacin as needed when he developed fever. His fever resolved promptly after taking antibiotics, but his appetite gradually declined. On admission, chest X-ray revealed nodules in the left upper lobe and bilateral peripheral infiltrates. Microscopic exam with staining for acid fast bacilli was performed on three consecutive sputum smear specimens in the morning after admission, and two of three exams were positive. Mycobacterium tuberculosis infection was confirmed by additional DNA probe testing. The patient was transferred to another medical institution specialized for treatment of TB.

**Discussion:** Levofloxacin, one of FQ, is one of alternative drugs of choice for treatment of TB in patients who cannot take first-choice drugs due to drug resistance or adverse effect. However, if used simply as a treatment for common bacterial infections in patients with TB,FQ exposure is associated with negative sputum smear results with greater risk for diagnostic delay and even death.FQ has been the most frequently used drug for patents with UTI.But TB is still in endemic in Japan, and careless prescription of FQ should be avoided. In this case, fever caused by TB was misdiagnosed as UTI and partial and incomplete treatment with levofloxacin was employed several times. As a result, the accurate diagnosis was delayed, and the risk of secondary infection also increased. In elderly patients with fever, UTI tend to be a wastebasket diagnosis. It is important to pay attention to the possibility of premature closure of diagnosing TB, a treatable infectious disease potentially spreading to others.

#### "Common Diseases"

M. Yokose, Y. Harada, T. Shimizu Dokkyo Medical University Hospital, Shimotsuga, Japan

#### **Learning objectives:**

- Seemingly apparent medical conditions can hide some important diagnostic clues, thereby generating anchoring bias.
- Rare diseases can be common in specific patient population.

Case information: A 70-year-old woman was referred with a two month history of fever which was attributed to cellulitis of the left lower limb. Her fever had not improved despite several antibiotics. She also admitted mild headache and nausea. She underwent renal transplantation 8 years prior, and was taking tacrolimus, mycophenolate mofetil, and methylprednisolone. Vital signs were notable for high fever (38.1). Physical examination showed swelling and redness of the left lower leg. There was no meningeal irritation. Blood test, urinalysis, and imaging studies were normal. Blood culture showed no bacterial growth. Her fever, headache, and nausea were considered due to cellulitis alone. Her cellulitis resolved with a seven-day course of cefazolin, however, her symptoms sustained. She was reevaluated. Urine culture on day 9 became positive for Cryptococcus neoformans, and cryptococcal antigen was detected in both of serum and cerebrospinal fluid. She was diagnosed with cryptococcal meningitis, and treated with antifungal agents successfully.

Discussion: This case underscores a good example of anchoring bias: The "mildness" of headache and nausea served as anchoring bias and herein diagnostic delay occurred. When a patient with seemingly apparent common medical condition, i.e. cellulitis as in this case, physicians can be inclined to be anchored to the diagnosis and be neglecting other accompanying mild symptoms rather than regarding them being important. In fact, headache and nausea are core manifestations of meningitis. On the other hand, cryptococcal meningitis is common disease in patients with renal transplantation (20%). In this sense, "mild" headache and nausea in this case should have been regarded as common symptoms of cryptococcal meningitis. Clinicians should always be aware that the prevalence of diseases can vary depending on patient population.

## **Poster Session 2: Scientific Abstracts**

Monday, November 5, 2018 | 7:00 AM - 8:00 AM

#### **Electronic Health Record Data Versus Patient-Reported Data**

K. T. Gleason<sup>1</sup>, D. E. Newman-Toker<sup>2</sup>, H. Lehmann<sup>3</sup>

<sup>1</sup>Johns Hopkins University, Baltimore, MD

<sup>2</sup>Center for Diagnostic Excellence, Armstrong Institute for Patient Safety and Quality, Johns Hopkins University School of Medicine, Baltimore, MD

<sup>3</sup>Johns Hopkins University School of Medicine, Baltimore, MD

Background: Electronic medical records (EMRs) are increasingly used for quality improvement purposes, and have the potential to meaningfully improve diagnosis. Methods such as electronic trigger tools to measure diagnostic process failures and SPADE (symptom-disease pair analysis of diagnostic error) to identify misdiagnosis-related harms rely on accurate EMR data. However, there remain concerns that EMRs are may be plagued by data quality issues, including both inaccurate and missing data. We sought to measure the accuracy of key EMR data in an existing cohort study.

Methods: The PaTH AF Longitudinal Cohort Study (N=95,302) recruited participants with an AF diagnosis by ICD 9-10 codes or electrocardiographic reading and age ≥18 years across 4 academic medical centers. EMR data and biannual patient-reported data were collected in tandem for 953. EMR documentation was compared to patient-report of anticoagulant use and demographics (age, sex, race, language). Level of agreement was determined through Cohen's Kappa coefficients. The association between EMR and patient-reported agreement with additional covariate factors (age, race, sex, age, education level, healthcare utilization, and AF diagnosis documented by ICD 9/10 codes vs. electrocardiographic reading) was explored using multiple logistic regressions.

Results: Our study population (n = 953) were 64% male (n = 623), 96% white (n = 890), and 72 (±10) years old on average. A small percentage of individuals (2.9%, n = 28/967) who reported they had AF diagnosis had no EMR documentation of an AF diagnosis other than an electrocardiographic reading, and half of these individuals had a documented anticoagulant prescription indicating that the providers in the health system were likely aware of the AF diagnosis. There was a moderate to good agreement (Kappa 0.5 to 0.8) between EMR documentation and patient report (Figure). In adjusted models, more comorbidities (OR 1.25, 95% CI, 1.06, 1.47), higher education level (OR 1.46, 95% CI, 1.26, 1.70), and ICD 9/ICD 10 coded AF diagnosis (OR 4.60, 95% CI: 2.12, 9.97) were associated with a higher likelihood of agreement.

Conclusion: We found that up to ten percent of patient-reported data differed from EMR data, and roughly 3% of patients who reported AF had no EMR diagnosis of AF. As EMRs are increasingly used to improve diagnosis, maximizing the accuracy of EMR data will likely require direct patient engagement to ensure data are accurate and up to date.

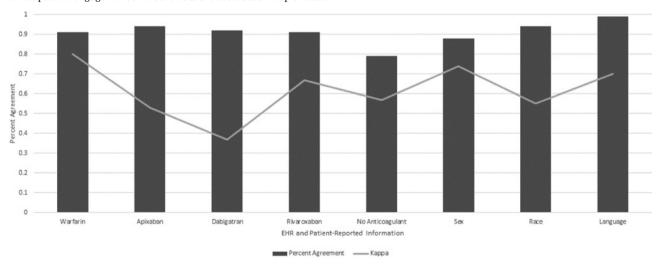


Figure 1. The Percent Agreement and Cohen's Kappa Between Patient-Reported Information and EHR-**Documented Information** 

## **Poster Session 2: Education Abstracts**

Monday, November 5, 2018 | 7:00 AM - 8:00 AM

#### Morning Report for All: The Use of Podcasts for the Dissemination of Clinical Reasoning Tools

R. C. Augustin<sup>1</sup>, S. Tilstra<sup>2</sup> <sup>1</sup>University of Pittsburgh Medical Center, Pittsburgh, PA <sup>2</sup>University of Pittsburgh School of Medicine, Pittsburgh, PA

Purpose/Problem: The use of social media and podcasts have drastically revolutionized medical education, resulting in a positive impact on learner engagement, collaboration, and professional development. However, the capability of these programs to disseminate high-quality and engaging resources for medical decision making has not been documented.

Description of program, assessment, or study: The podcast, Morning Report, is an audio podcast of a clinical reasoning session much akin to standard "morning report." Two expert physicians discuss a complex case with sequential disclosure of clinical information. Focus is on the clinical reasoning process, concepts and clinical pearls. 15 episodes were written, recorded, vetted for content validity, produced and disseminated using the iTunes podcast application in 2017-2018 by faculty, residents and students at Case Western Reserve School of Medicine. Goals were to: (a) create an open-access clinical reasoning resource for medical trainees, (b) role-model medical decision making, c) highlight cognitive biases and metacognitive principles, and (d) show that podcast is a feasible way to disseminate clinical reasoning concepts.

Outcomes: By June of 2018, over 1000 episodes had been downloaded. As expected, a majority of downloads occurred in Ohio; unexpectedly, listeners were documented worldwide (Fig. 1). The majority of episodes (67%) were downloaded via mobile applications (primarily the iTunes podcast app). The average number of downloads per episode was 67; the most popular episode was downloaded 114 times and the least popular episode was downloaded 43 times. Currently, the app holds a 5-star rating in the iTunes catalog, but with the caveat of only having 7 reviewers to date. Plans to collect qualitative data on the content, delivery, and usability of Morning Report are underway.

Discussion: Morning Report is an efficient, portable way to disseminate clinical reasoning concepts with the capability of reaching thousands of learners in a short amount of time, while allowing learners to critically assess, pause, and analyze the case along with the expert discussant. The ability of the digital platform to track number of downloads, popularity of topics, and user comments is an advantage over typical educational delivery systems to guide content and delivery in real-time.

Significance of findings: Disseminating clinical reasoning tools via podcasts is not only effective, but can also be done in a relatively inexpensive and efficient manner. More research in how disseminating clinical reasoning concepts via podcasts can affect decision making is needed.

- Cheston C, Flickinger T, and Chisolm M. Social media use in medical education: a systematic review. Acad Med. 2013 Jun;88(6):893-901. 1.
- morningreport-podcast.com

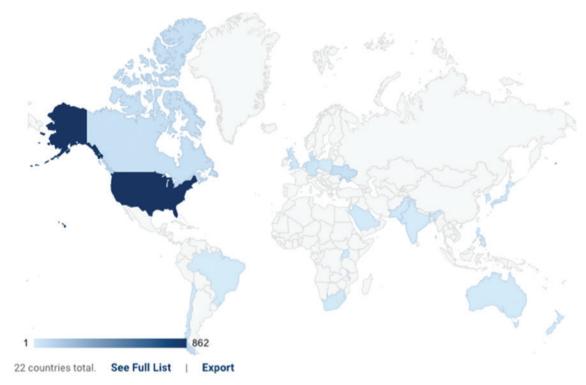


Fig. 1. Compilation of episode downloads by country of origin

#### The Effect of a Multi-Level Curriculum on Use of Clinical Reasoning Terminology in Residents' Nationally Presented Clinical Abstracts

E. Bonifacino<sup>1</sup>, W. Follansbee<sup>2</sup>, M. McNeil<sup>2</sup>, D. DiNardo<sup>2</sup> <sup>1</sup>University of Pittsburgh Medical Center, Pittsburgh, PA <sup>2</sup>University of Pittsburgh School of Medicine, Pittsburgh, PA

Purpose/Problem: Improving training in medical decision-making has been an increasing priority for the National Academy of Medicine and the ACGME. Despite this, optimal strategies for explicit instruction in clinical reasoning, as well as assessment methods for efficacy of existing curricula, remain unclear. Our objective was to evaluate the impact of a multi-modal multi-level clinical reasoning curriculum on resident interest in and comfort with clinical reasoning concepts as reflected by inclusion of clinical reasoning terminology in national conference submissions.

**Description of program, assessment, or study:** From 2014 to 2018, we implemented curricular efforts targeted at medical students, residents, and faculty in our division. First, we created faculty development sessions that detailed the impact of diagnostic error, reviewed clinical reasoning terminology, and provided an overview of cognitive psychology principles as they relate to decision—making. We similarly exposed our residents and students to these concepts through the use of interactive online modules with reinforcement of concepts during interactive case-based workshops tailored to each learner level. Additionally, we used widely disseminated posters and pocket cards to keep these concepts at the forefront in the real clinical setting. Finally, during programmatic interventions such as our "clinical reasoning case conference," we created and maintained venues for role modeling of medical decision-making including the approach to diagnostic uncertainty as well as for open discussion of diagnostic error and cognitive biases.

Outcomes: We evaluated the impact of our curriculum on the use of clinical reasoning-specific terms in independently submitted clinical vignettes to the Society for General Internal Medicine conference by our internal medicine residents. We compared the number of published clinical vignettes that included these terms prior to our curricular efforts (2014) to the incidence four years after initiation (2018). In 2014, 1 out of 15 abstracts included a clinical reasoning term (7%). In 2018, 13 out of 33 abstracts (39%) included a discussion of clinical reasoning, a 6-fold increase in quantity. The most common term was "framing bias," referenced in 5/13 (38%) of the abstracts in 2018.

Discussion: We have demonstrated that a multi-pronged curricular effort can be associated with increased inclusion of clinical reasoning terms in our residents' published clinical abstracts. Prior studies have postulated that use of clinical-reasoning-specific terminology can serve as cognitive scaffolding that can inform self-directed learning and skill-specific education.

Significance of findings: Our findings suggest that a clinical reasoning curricular effort can impact trainee interest in and comfort with clinical reasoning concepts, and thus provide a scaffolding for application to clinical scenarios.

#### Assessing Clinical Reasoning in Medical Student Soap Notes: Validating Three Instruments Using Messick's Criteria

Y. Covin<sup>1</sup>, N. Wick<sup>2</sup>, K. Gavinski<sup>2</sup>, J. Wagner<sup>2</sup>, P. Longo<sup>2</sup> <sup>1</sup>UT Health San Antonio School of Medicine, San Antonio, TX <sup>2</sup>UT Southwestern Medical Center, Dallas, TX

Purpose/Problem: Several assessment instruments intend to measure clinical reasoning ability communicated in Subjective-Objective-Assessment-Plan (SOAP) notes, yet we lack evidence corroborating their validity to a standard. We set out to compare the validity evidence using Messick's Criteria, specifically relationships to other variables, of three clinical reasoning instruments: Clinical Reasoning Task (CRT) checklist, Patient Note Scoring Rubric (PNS), and Summary Statement Assessment Rubric (SSAR). Global instrument scores were compared to a gold-standard clinical reasoning test, the Clinical Data Interpretation test (CDI).

**Description of program, assessment, or study:** During the last month of an 18-month pre-clinical curriculum, 230 students completed the CDI test, and then created an individual SOAP note for a clinical case. Four students from each CDI quartile were randomly selected to create a second SOAP note and to participate in a Think Aloud interview to verbalize their reasoning process for a second clinical case. Due to scheduling, three students were unable to participate in the Think Aloud interview. Analysis used 230 SOAP notes derived from the first clinical case, 16 written SOAP notes from the second clinical case, and 13 Think Aloud interviews. Each instrument's global score was correlated to CDI with Pearson correlation. Thirteen student demographic domains (e.g., gender, race/ethnicity, college major, pre-medical clinical experience) were collected and compared to instrument scores with ANOVA.

Outcomes: Statistically significant agreement achieved in all qualitative coding analysis [Intra-class Correlation Coefficient range: 0.77 – 0.98]. Only SSAR demonstrated a significant correlation with CDI (r = 0.201; p = 0.03). A large, significant correlation between the PNS and CRT global scores was seen (r = 0.71; p = 0.002). As a mediating variable, the thirteen demographic domains did not interact with the instrument scores.

Discussion: Differing strengths of association between clinical reasoning instruments suggests varying degrees of overlap in the clinical reasoning frameworks underpinning the assessment domains. Strong correlation between CRT and PNS scoring may suggest early areas of clinical reasoning capability development, which may not be yet captured in CDI or SSAR, which are scales weighted toward clinical knowledge synthesis and hypothesis testing.

Significance of findings: Our findings represent the first comparison of medical student SOAP note assessment with a clinical reasoning assessment standard. Educators interested in clinical reasoning assessment should be encouraged that the four instruments used demonstrated objectivity across all measured demographics. Future directions of this work should include SOAP note assessment at all training levels to determine the clinical reasoning instruments most appropriate for each level of clinical reasoning capability development.

## In Systematic Analysis of a Clinical Diagnosis, Using Checklist after Brainstorming Almost Doubles Actionable Diagnostic Hypotheses in Comparison to Using Checklist to Frame Approach before Brainstorming.

P. Foster

Greater Baltimore Medical Center, Towson, MD

Purpose/Problem: In response to excessive variation in resident performance with an educational tool designed to promote structured team assessment of diagnosis, we hypothesized that creating a version that visually emphasized free brainstorming would enable residents to enhance Type 1 diagnostic reasoning, and allow more systematic incorporation of clinical data. Moreover, we hypothesized that using the checklist to frame the brainstorming would inhibit performance. Description of program, assessment, or study: Randomized cross-over study comparing two versions of our TACT (Team Analysis of Clinical Thinking) forms. One form presents a pathophysiology based checklist on the left border with the appearance of row headers and instructions to brainstorm within each category (Precheck form), and a second where three columns of blank space are followed by the same checklist on the right border (Postcheck form). With the second, residents were instructed to brainstorm freely during the presentation and then consider the checklist. The interns received a brief introduction to the history of the approach and how to use the tool. They were then randomized to a sequence of the forms and presented with two cases chosen by their senior residents as unusual presentation of common disorders. The interns then graded hypotheses by how much they felt action needed to be taken.

Outcomes: 10 interns participated. Case 1 was an 81-year-old with history of vascular risk factors presenting with neck pain and an unresponsive pupil. Case 2 was an 18-year-old presenting with a neck mass, palpitations, tremor, and abdominal pain.

Differential Diagnoses	Case 1			Case 2	
	Precheck	Postcheck	Precheck	Postcheck	
Group total	17	20	14	36	
Actionable	14	15	10	23	
Individual average	5	8.5	6.5	11.3	
Precheck average	5.6				
Postcheck average	10.2				
P-value	0.013				

Discussion: The TACT sheets were developed to help residents practice modes of type 1 and type 2 thinking, and learn the value of systematic teamwork. Traditionally, interns explore few hypotheses and include irrelevant or extremely rare ones. This study shows that small design elements can have a large impact on tool performance. In our case, instructions and design that clearly sequenced type 1 and type 2 thinking appeared to almost double the number of actionable diagnostic hypotheses.

Significance of findings: For educators designing tools to reduce diagnostic error and increase learner sophistication, careful testing of form factors and function may improve performance of the tool and provide insight into human decision making.

#### Can North Korean Doctors Settle Down South Korea after the Union of the Korean Peninsula?

S. H. Kang

Kangwon National University School of Medicine, CHUNCHEON, Korea, Republic of (South)

Purpose/Problem: The health care system of North Korea producing medical doctors has collapsed. Hence, there is reason to question the clinical competence of their doctors. But few studies about the clinical competence of their doctors have been conducted. We examined results of North Korean refugee doctors in clinical skill tests provided by a medical school in South Korea.

Description of program, assessment, or study: In two consecutive years, one medical school in South Korea administered a clinical skill test, composed of the clinical performance examination (CPX) and the objective structured clinical examination (OSCE), to seven North Korean refugee doctors and their fourth-year medical students. The annual test results of each group were compared using the Wilcoxon rank-sum test. Outcomes: The North Korean doctors scored lower on the history-taking domain of the CPX, and on the domain of practice and proficiency of the OSCE than South Korean medical students. The North Korean doctors displayed higher average score on physical examination domain in the CPX.

Discussion: The differences of the customs and expressing symptoms caused North Korean doctors' lower score of history-taking domain in spite of using the same language. Although their lower score of the practice and proficiency domain in OSCE seem to be caused by their inexperience due to the insufficient training system, their practiced physical examination skill imply the feasibilities of successful settlement in South Korean medical system.

Significance of findings: With the proper support North Korean refugee doctors could successfully settle down in South Korea after the union of the Korean Peninsula.

#### **Applying Tools of Great Diagnosis to Healthcare Leadership**

A. Miller

201 Dromara Rd, Guilford, CT

Purpose/Problem: Physicians seeking leadership positions understand they need expertise in new areas. Aspects of great diagnosis are essential to effective leadership and are not to be overlooked.

Description of program, assessment, or study: A half-day workshop was offered as a mandatory session in a 16-month Physician MBA Program. The objectives were to identify foundations of strategic thinking, to emphasize skills, interaction, and practice, and to catalyze out-of-the-box experiences for the cohort. The program design included: 1) A 1-hour didactic identifying key fundamentals of great diagnosis and their application to leadership practice; 2) A three-hour experience in a nearby art museum gallery in which participants practiced skills of listening, framing, and perspective-shifting in groups with art and gave feedback; and 3) A 1-hour discussion with applications to cases, forthcoming curricula, and personal leadership learning goals.

Outcomes: While the intention of the program was to focus on three skills (listening, framing, and perspective-shifting), the group prioritized rigorous focus on listening. Conversationally, participants identified communication blocks across difference as a major challenge in healthcare and described scenarios from their professional experiences in which their own voices were not heard. The MBA program organizers reported that most students enjoyed and found valuable the session, while some struggled with the new experience.

Discussion: There is an overlap between fundamental skills in great diagnosis and great leadership that stands to be emphasized in the training of emerging physicians.

**Significance of findings:** Further study is needed.

#### S.A.F.E.: A Scaffold Curriculum to Teach Safe, Appropriate, Timely and Value Based Imaging to Medical Students

J. A. Neutze, A. S. Burdette, J. M. Brian, P. L. Brian, P. J. Chiarolanzio, A. L. Chetlen Penn State Health, Hershey, PA

Purpose/Problem: To create an organized, easily remembered scaffold- approach curriculum to teach safe, appropriate, timely and valuebased imaging to medical students.

Description of program, assessment, or study: Medical students indicate their motivation for taking radiology electives is to learn how to interpret imaging studies. Some are becoming radiologists. Others are learning "basic" radiology to use during the course of their practice, regardless of specialty. Classic radiology electives have been created to meet this expectation, with students spending time at view boxes, learning radiologic anatomy and looking at unknown radiology studies. An important element of radiology education not previously stressed to students has evolved with increased public awareness of need to reduce unnecessary imaging exposure and to contain costs incurred as a result of unnecessary imaging. Addressing this need will provide students with knowledge and skills to be able to choose the safest and most appropriate imaging examinations now and in future practice. The Penn State Health Radiology Medical Student Education directors have refined the goals of the Diagnostic Radiology electives to not only fulfill the students' educational expectations, but also to provide them with new non-interpretative skills. We created a new program: SAFE: Safety -Appropriateness-(interpreting) Films- Expedite and Execute (table 1). We stress the rationale for imaging choices should be applied in that order; that is, without safety and appropriateness practiced FIRST, even the best imaging interpretation may not result in the indicated and desired value-based health care health care.

Outcomes: Post-elective survey data from medical students enrolled in Diagnostic Radiology electives for past two years indicate they understand and plan to use the concepts of the SAFE program.

Discussion: SAFE radiology is an easily learned, effective method to introduce learners to concepts of appropriate, timely and valuebased imaging. Since 2016 program implementation and national introduction by poster presentation at 2017 HVPAA conference, we have broadened applications of the SAFE radiology concept throughout the 4 year medical school curriculum in courses such as anatomy, clinical clerkships and other electives. We are also considering the creation of an "accomplishment designation," e.g., SAFE achievement award, similar to the Six Sigma program to recognize levels of mastery and to expand the programs to provide exposure to clinical residents and faculty.

Significance of findings: Most learners rotating though radiology will not become radiologists. Rotating learners now know that safety and appropriateness should procedure image procurement and interpretation and that they have resources to implement and to teach colleagues about SAFE concept.

## SAFE

A scaffold curriculum to teaching appropriate imaging

- S: Safety- Discuss uses of ionizing radiation, MRI, ultrasound, and radiology contrast safety. Compare and contrast radiology safety applicable to the pediatric, pregnant and elderly patient and the role of Health Physics personnel.
- A: Appropriateness-Use resources such as their radiologists, ACR Appropriateness Criteria ™, Image Gently ™, Image Wisely ™ and Choosing Wisely ™ to order appropriate imaging studies while managing resources and maximizing safety information obtained. Observe the role that radiology studies and radiologist play in the overall care and management of the patient.
- **F: Films-** Develop a systemic approach to evaluating chest and abdominal radiographs. List core concepts in advanced imaging such as Hounsfield units (CT), T1 and T2 (MRI) and fluid (ultrasound). Describe patient preparation for radiology studies. Observe how patient studies are performed.
- E: Expedite and Execute Expedite patient management by recognizing common emergent findings that a student or clinician might see before a radiologist interprets the study. Execute the knowledge you have learned about safety, appropriateness and image and report interpretation to provide safe and effective patient-centered care.

#### A Medical Student Elective to Improve Diagnosis in Health Care: Developing Solutions to Reduce Patient Harm

J. A. Neutze, J. Beatty-Chadha, T. J. Mosher Penn State Health, Hershey, PA

Purpose/Problem: Medical schools have begun teaching patient safety in response to the Institute of Medicine (IOM) report estimating up to 80,000 US patient annual deaths from medical errors. Most courses tend to focus on teaching and correcting system errors because they are visible and may be easier to address. Diagnostic errors, accounting for 10-20% of medical errors, may be given less attention for a variety of reasons including less frequent reporting and greater difficulty to fix. Department of Radiology faculty champions identified a curriculum need to expose students to concepts of diagnostic error and solution development. Using the IOM report Improving Diagnosis in Health Care (2015) as a framework, an elective focusing on creating solutions to reduce diagnostic error has been developed.

Description of program, assessment, or study: Third-year Penn State College of Medicine students have the opportunity to participate in two-week elective experiences between clinical rotations. This pass/fail elective includes an introduction to the concept of diagnostic error and solutions, readings, facilitated discussion, and creation and presentation of final projects; checklist tool documents assessment and evaluation (Table 1). During this elective, students work with radiology mentors as well as in small groups to design projects and present results and recommendations to address their chosen IOM goal(s). Students identify current barriers and resources needed to implement change and what metrics might be used to measure success. Presentation format is based on type of project and data collected. Students choose a project from a list of options or develop a new project.

Outcomes: To date, two sessions, 3-4 students/session, have been completed; next session is offered in September 2018. Separate groups have developed projects on workplace distractions and feedback as they relate to diagnostic error. Students have connected to existing literature, identified local barriers/opportunities and developed concrete solutions in a very short timeframe.

Discussion: Diagnostic errors and solutions are vague, new concepts to medical students and even facilitator faculty. Our first group demonstrated new and expanding knowledge; second group has built on prior session work expanding initial projects and developing new ones. Longer electives might be offered in the future, allowing for expanded projects and to involve more faculty and students.

Significance of findings: Students and facilitators have started developing banks of concrete solutions to address IOM goals. There are also possible positive implications for each medical student's learning and future practice as they gain awareness of causes of diagnostic error and identify steps to reduce patient harm in daily work.

## Improving Diagnosis in Health Care: Developing Solutions to Reduce Patient Harm **Requirements Checklist**

Requirement	Completion Information
Introductory Meeting or Module	Date completed:
Required Readings	Institute of Medicine (2015). Improving diagnosis in health care: Report in Brief. Download here: http://bit.ly/IOMBrief Date completed:  Institute of Medicine (2015). Improving diagnosis in health care pp. 1-18 and chapter(s) applicable to independent project. Download here (PDF available): http://bit.ly/IOM-Book Date completed:
Day 2 In-Person Session (2 hours)	Date attended:
Mentor Meetings	Name of Mentor:
(at least two in-person or online/phone meetings)	Date Meeting: Duration of Meeting: Describe what you discussed and how you will use the information to develop your project.
Independent Project  (approximately 40 hours spent on reading, development, and related activities)	Record the following for individual project preparation completed over the course of two weeks.  Date:  Duration:  What session(s) did you attend and/or independent work did you complete?
Final Project Presentation (portion of 2 hour session)	Date attended:
Final Project	Date Submitted: Date Feedback Received: How will you use the feedback on your paper to improve future practice?

#### Analysis of Radiology Reporting: Where Can Errors be Eliminated?

G. A. Wright, S. Piedra Abusharar, J. A. Holbert, J. Beatty-Chadha, T. J. Mosher, J. A. Neutze Penn State Health, Hershey, PA

Purpose/Problem: To evaluate the radiologic process from the initial order of a study to the final communication of results to the ordering physician, identify areas of inefficiency and problematic steps, and to subsequently suggest changes which will improve patient outcomes.

Description of program, assessment, or study: In order to determine areas of inefficiency and potential for error in the radiologic process at our center, a survey was distributed to Penn State Health attending (56) and resident (24) radiologists. Participants were asked to provide information about what they saw as the greatest barriers in efficiency in study interpretation and result communication, identify what part of the process was the biggest risk for patient harm, and what changes they would make to improve the process. The survey responses were then compiled and analyzed for common themes in both areas for potential error and ideas for improvement.

Outcomes: A total of 18 survey responses were received (11 attending, 7 resident). Of these, 17(94%) respondents identified interruptions and incomplete history as the greatest barrier to the interpretation of a study. A total of 16(89%) respondents identified not having the ordering physician call back or having the incorrect contact information as the greatest barrier in efficiency in communicating study results. As a result, the majority of respondents cited poor communication or lack of follow up as the biggest risk for patient harm. Respondents suggested minimizing distractions, improving communication to ordering physicians, and quality improvement via double reads of studies. A total of 14(78%) respondents did not think that standardized reporting would help improve the issues they identified in the survey.

Discussion: Although there are many ways that errors can be introduced into a system, it is important to identify and eliminate potential errors in order to improve the lives of patients. By performing this study, inefficiencies specific to this center were identified, and potential solutions were developed.

Significance of findings: One area of inefficiency may be another center's area of strength and vice versa. Surveyed stakeholders identify that they are aware of problems and may be able to offer solutions, such as enlist colleagues to help and using information technology and artificial intelligence. This study provides a framework for future work to focus on specific areas of potential error and to implement solutions to improve them.

## Radiology Reporting Improvement Survey

The purpose of this project is to evaluate the radiologic process from the initial order of a stu-	
the final communication of results to the ordering physician to identify areas of inefficiency as problematic steps in the process that may lead to poor patient outcomes.	
The Radiologic Process: Study ordered $\Rightarrow$ Study available for interpretation $\Rightarrow$ Interpretation completed $\Rightarrow$ Report delivered to provider	
* Required	
Are you a resident or attending physician? *	
Resident	
Attending	
What do you see as the greatest barrier in efficiency in the interpretation of a study? *	
Your answer	
What do you see as the greatest barrier in efficiency in communicating results ? *	
Your answer	
What step in the radiology process do you see as the biggest	t
What step in the radiology process do you see as the biggest risk for patient harm? *  Your answer	t
risk for patient harm? *	t
risk for patient harm? * Your answer	t
risk for patient harm? *  Your answer  What part of the process do you feel works well? *	
risk for patient harm? *  Your answer  What part of the process do you feel works well? *  Your answer  If there was one thing about the process that you could chan	
risk for patient harm? *  Your answer  What part of the process do you feel works well? *  Your answer  If there was one thing about the process that you could chan what would it be? *	
risk for patient harm? *  Your answer  What part of the process do you feel works well? *  Your answer  If there was one thing about the process that you could chan what would it be? *  Your answer  Do you think that standardized reporting could address the	
risk for patient harm? *  Your answer  What part of the process do you feel works well? *  Your answer  If there was one thing about the process that you could chan what would it be? *  Your answer  Do you think that standardized reporting could address the issues mentioned above? *	

#### Improving Feedback Sharing Culture in Radiology

S. H. Lee<sup>1</sup>, J. A. Neutze<sup>2</sup>, T. J. Mosher<sup>2</sup>, J. Beatty-Chadha<sup>2</sup> <sup>1</sup>Penn State College of Medicine, Hershey, PA <sup>2</sup>Penn State Health, Hershey, PA

Purpose/Problem: The purpose of this study is to investigate how to improve the current diagnostic error feedback sharing practice in radiology.

Description of program, assessment, or study: An electronic survey was used to assess the frequency and quality of diagnostic error feedback, barriers to feedback sharing, and opinions and attitudes of both resident and attending radiologists regarding feedback sharing at Penn State Health Department of Radiology, Literature search was used to help explain the barriers and propose potential solutions.

Outcomes: Only 31% of the radiologists surveyed stated that they were aware of diagnostic errors they made. This is despite the fact that the Department of Radiology has a policy that radiologists be informed by radiology colleague of a diagnostic error that has been made. Overwhelming, 92% agreed that receiving feedback was important to their professional growth. In contrast, only 38% were satisfied with feedback received, largely due to the scarcity of feedback. Greater than 75% reported that they received less than one feedback per month from either other radiologists or ordering physicians. Several radiologists commented on the importance of positive culture on feedback sharing, pointing out the need to create non-threatening, welcoming environment for feedback sharing.

Discussion: The authors suspect that negative experiences arising from feedback sharing may be attributed to the lack of training in giving and receiving feedback resulting in both parties leaving unsatisfied of the outcome of a feedback session partly due to misunderstanding of each other. Proper training in sharing feedback may improve the satisfaction of both parties. It is proposed that a focus group be created to allow participating radiologists to be trained in the above skills, track their feedback sharing rate, correlate to their diagnostic performance, as well as their job satisfaction rate.

Significance of findings: By helping radiologists to embrace feedback sharing, we may be able to increase the overall performance of radiologists by encouraging lifelong learning and thus reducing diagnostic error in radiology.

Feedback and Reducing Diagnostic Errors	I am satisfied with feedback I receive.	I receive feedback from clinicians (surgery, internal medicine, etc.)	
		On a daily basis	
The goal of this project is to investigate ways to improve feedback for reducing diagnostic errors in the Department of Radiology at Penn State Hershey Medical Center.	○ Strongly agree	○ Weekly	
	○ Agree	○ Murthy	
My current level of training is	Neither agree or disagree	O Rendy	
O POV2		○ Never	
O P013	Disagree		
O POY4	○ Strongly Disagree	I receive feedback from other healthcare workers (technicians, nurses, etc.)	
O POYS		Only a daily basis	
○ Felow	If I am not satisfied, this is because	Weekly	
Attending	Long answer text	Munthly	
	As a resident, I receive feedback from senior resident/attending	Ranky	
I am aware of diagnostic errors I make.	On a deliybesis	○ Never	
○ Strongly agree	Weekly	I prefer to receive feedback in the following way (check all that apply)	
○ Agree		n person	
Neither agree or disagree	Morelly	E-mail or other electronic communication	
O Disagrae	Revely	in small group	
Strongly disagree	Never	in conference	
	○ N/A	Other	
Receiving feedback is important to my professional growth.			
○ Strongly agree	I receive feedback from colleagues	Describe how you seek out feedback or use feedback to address diagnostic errors.	
○ Agree	On a daily basis	Long answer text	
Nother agree or disagree	Weekly		
O Disagree	Munthly	Please provide suggestions for how to improve the amount and quality of the	
○ Strongly disagree	Randy	feedback that you receive.	
	○ Never	Long answer feet	
		Any other comments?	
		Long answer text	

## Use of Learning Sciences-Derived Research Frameworks and Methodologies to Analyze the Cognitive Factors **Enabling the Diagnosis of a Complex Patient Presentation: A Case Report**

F. J. Papa<sup>1</sup>, S. Dar<sup>2</sup>, J. McKenzie<sup>2</sup>, R. Bais<sup>2</sup> <sup>1</sup>UNIVERSITY OF NORTH TEXAS, Fort Worth, TX <sup>2</sup>University of North Texas Health Science Center, Ft Worth, TX

Purpose/Problem: Medicine has yet to produce or endorse, a codified framework for investigating the cognitive factors underlying the development of diagnostic competencies. However, the Learning Sciences offer several theoretical frameworks and investigative methodologies which could lead to a deeper understanding of the cognitive factors enabling differential diagnosis (DDX).

Description of program, assessment, or study: This study involved the use of: 1) Pivot and Cluster (PC), a framework for investigating how clinicians formulate an evolving understanding of a patient's problem and thereby, define the various diseases/disorders to be considered during the workup (i.e., define the clinician's cognitive 'workspace'), and 2) Dual Processing Theory (DPT), a framework for understanding how two distinct forms of reasoning and knowledge determine which of the several competing diseases/disorders (represented within the workspace) is the most likely etiology for the problem at hand. (DPT posits that System 1 reasons via largely unconscious pattern recognition processes and uses knowledge in the forms of case exemplars and disease prototypes while System 2 reasons consciously via analytical processes and knowledge in the form of IF/THEN rules). PC and DPT frameworks were subsequently used to produce a DDX-oriented investigative methodology (a Protocol Analysis - PA) to qualitatively and quantitatively deconstruct and analyze, the evolving problem space, differentials, reasoning processes and knowledge utilized to perform DDX. An AV recording of how an expert clinician diagnosed a complex case of Acetaminophen toxicity (recorded at the 2017 SIDM conference) served as source material for the PA.

Outcomes: Protocol Analysis enabled the identification of five evolving problem spaces containing respectively, 7, 17, 8, 4 and 6 diseases/disorders; totaling 31 unique differentials. The clinician's listing of the differentials within each problem space represented a conscious System 2-like effort. However, there was little evidence to suggest that System 1 or System 2 reasoning processes or knowledge predominated during the clinician's workup. The author will utilize a multimedia tool called Camtasia to demonstrate when and how: 1) any given problem space evolved, and 2) System 1 and/or System 2 reasoning processes and forms of knowledge were utilized during the clinician's workup.

Discussion: PC and DPT enabled the construction of a protocol designed to deconstruct, analyze and explicate some of the cognitive factors enabling the diagnosis of a complex case of Acetaminophen toxicity.

Significance of findings: Learning Sciences frameworks involving PC and DPT can be used to produce a protocol useful as an instructional and assessment methodology during medical training.

#### Partnered (Patients Assigned to Research Teams with Nurses and ER providers to Enhance Diagnosis)

S. Peterson<sup>1</sup>, K. T. Gleason<sup>1</sup>, Z. Berger<sup>2</sup>, A. McClenney<sup>3</sup>, D. E. Newman-Toker<sup>4</sup>

<sup>1</sup>Johns Hopkins University, Baltimore, MD

<sup>2</sup>Johns Hopkins Univserity School of Medicine, Baltimore, MD

<sup>3</sup>Johns Hopkins, Baltimore, MD

<sup>4</sup>Center for Diagnostic Excellence, Armstrong Institute for Patient Safety and Quality, Johns Hopkins University School of Medicine, Baltimore, MD

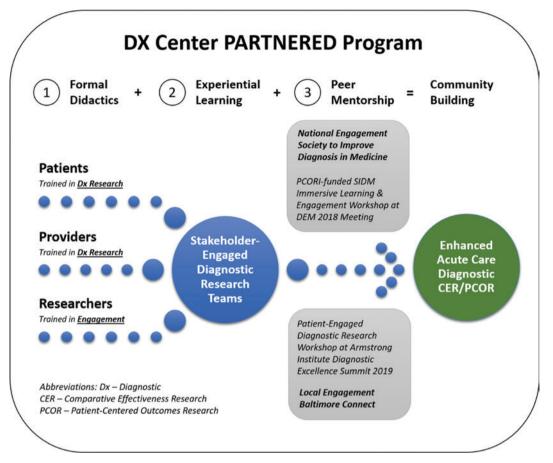
Purpose/Problem: Key stakeholders, specifically Patient partners, nurses and emergency department (ED) providers, have not been systematically engaged in patient-centered diagnostic research to improve ED diagnosis. Knowledge and skills in diagnostic comparative effectiveness research (D-CER) and diagnostic patient-centered outcomes research (D-PCOR) are lacking among key stakeholders. Building this capacity is essential to future research to improve patient-centered diagnostic outcomes. The Armstrong Institute Center for Diagnostic Excellence (DXC) seeks to create a growing cohort of patient partners, nurses, ED providers, and researchers who will be trained to participate in PCOR and CER focused on diagnosis.

**Description of program, assessment, or study:** PARTNERED (Patients Assigned to Research Teams with Nurses and ER providers to Enhance Diagnosis) is an 18-month initiative to train patients, nurses, and ED providers to engage in acute care diagnostic research methods. We aim to develop a sustainable program to ensure meaningful patient and stakeholder involvement in acute care diagnostic research. Patient partners, acute care nurses, and ED providers will receive training and be partnered with researcher teams. The DXC in partnership with SIDM and community organizations from Baltimore CONNECT and the Johns Hopkins Community Resource Advisory Council (CRAC) as well as the Johns Hopkins Patient and Family Advisory Council (PFAC) will collaborate to complete the project objectives.

Outcomes: The anticipated results are: 1) a sustainable training program for the development of patient nursing and provider leaders in D-PCOR and D-CER, 2) stronger partnerships between community patient partners and local diagnostic researchers, and 3) a dynamic cohort of patient partners, nurses, ED providers, and diagnostic researchers engaging in PCOR and CER.

Discussion: This PARTNERED program will develop a network of local patient partners and healthcare provider stakeholders with experience and knowledge in diagnostic research methods. Successful implementation of the program will help create a sustainable pool of stakeholders engaged in diagnostic research projects to improve patient-centered diagnostic outcomes.

**Significance of findings:** The development of a network of local patient partners and healthcare provider stakeholders with experience and knowledge in diagnostic research methods with a specific focus on acute care.



#### Thinking Fast and Slow: Training Faculty to Teach Clinical Reasoning

<u>V. Schaye</u>, M. Janjigian, K. Hauck, N. Shapiro, D. Becker, P. Lusk, S. Zabar, A. Dembitzer New York University School of Medicine, New York, NY

**Purpose/Problem:** Despite the call to action in 2016 to "Enhance health care professional education and training in the diagnostic process," most medical schools lack an explicit curriculum in clinical reasoning (CR) citing limited faculty expertise as a major barrier. We developed an easily implementable workshop to train faculty to teach CR using educational strategies grounded in dual process theory (DPT).

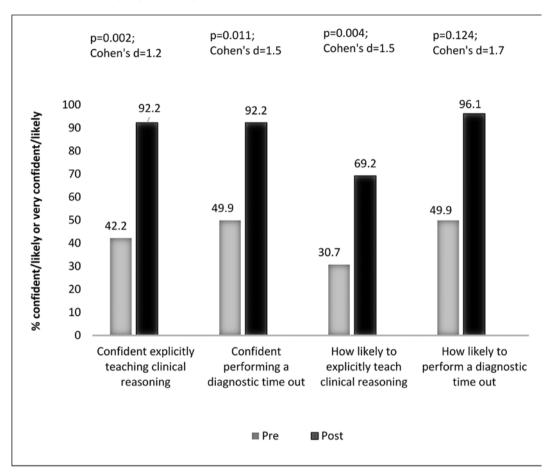
**Description of program, assessment, or study:** This workshop was part of a year-long faculty development program comprised of a preand post-Group Observed Teaching Exercise (GOSTE) and three workshops. Twenty-six medicine faculty from 6 different specialties participated. First, a baseline GOSTE simulated ward rounds with emphasis on learners' CR. Three standardized learners played the role of student, intern, and resident or fellow and each learner evaluated faculty performance on a scoring rubric on a 3-point Likert scale. Subsequently, we implemented a 3-hour workshop on teaching CR using a framework grounded in DPT. At the end of the workshop, participants completed retrospective pre-post (RPPs) assessments on change in confidence and in likelihood of using their newly acquired skills on a 4-point Likert scale. RPPs were analyzed using paired t-tests. Participants also completed commitment to change statements (CTCs) which were followed up on 2 months later.

**Outcomes:** In the baseline GOSTE, we found that 10% of faculty received a rating of well done for their use of problem representation and illness scripts, 60% well done for their strategies to broaden the differential diagnosis, 33% well done for reviewing strategies to prioritize the differential diagnosis, and 23% well done for discussing strategies to prevent cognitive errors such as diagnostic time outs. RPPs revealed that participants were more confident in their ability to teach CR and more likely to teach CR using strategies based in DPT with a large effect size in all domains (Table). In their CTCs, participants created specific plans for incorporating these CR strategies into their educational settings and on 2 month follow-up 81% of participants had partially implemented their CTCs.

Discussion: Similarly to what is found nationally, we discovered that our faculty lacked a framework for explicitly teaching CR.

Significance of findings: After participating in this easily implementable 3-hour workshop, faculty expressed greater confidence and increased likelihood to explicitly teach CR using a framework grounded in DPT with specific plans for incorporation into practice. Future studies would need to evaluate further the impact of training faculty to teach CR grounded in DPT on learners and diagnostic error rates.

Table. Retrospective pre-post survey results



#### Improving Diagnostic Thinking in Pediatric Residency

J. Seserinac, J. Lau The University of Arizona, Tucson, AZ

Purpose/Problem: The type 1 thinking, especially premature closing, anchoring, and diagnosis momentum, were predominant in residents' diagnostic process, while formulation of differential diagnoses, and ability to explain the diagnostic thinking process was lacking.

Description of program, assessment, or study: A new program, inspired by the NEJM Clinical Case Presentations, was developed in an effort to address the issues of diagnostic reasoning. It was integrated into the residents' teaching conferences. The chosen case was modified to obfuscate internet search and assigned to the third-year pediatric resident 2 weeks ahead of the scheduled session. The presentation comprised of case analysis, differential diagnosis development, discussion of diagnostic reasoning, ending with the final diagnosis and suggesting confirmatory test. The emphasis of the conference was on diagnostic decision-making process. All other pediatric residents received the case 1 week prior to the session with assignment to write their diagnoses on the attendance sheet at the beginning of the conference. They were encouraged to ask clarifying questions and challenge presenter's conclusions. In an effort to improve residents' involvement the conference format was adjusted. Modified sessions started with a quick review of the case's clinical aspects by the presenter. The audience, arranged into small groups (6-8), were given 10 minutes to discuss differential diagnoses. Each group announced their top three diagnoses with short discussion of their decision-making process, diagnostic test of choice and the final diagnosis. The presenter then proceeded with a diagnostic reasoning discussion, logic behind including and excluding certain diagnoses and the thinking process leading to the final diagnosis and diagnostic test selection.

Outcomes: 37 cases were presented. Two cases were excluded from the analysis due to presenter's misinterpretation of the instructions. The correct diagnosis was established in 33 cases.

Discussion: The pediatric residents not involved in the presentation guessed the correct diagnosis infrequently and were difficult to engage. Adjustment in session's format didn't much improve the diagnostic accuracy of diagnosis written on the attendance sheet, however, the overall involvement and participation in discussion was substantially improved.

Significance of findings: The goal of the curriculum was to improve diagnostic skills and advance differential diagnostic thinking of all pediatric residents during their training. The senior resident assigned to analyze the case showed substantial skills in diagnostic thinking, applying both types of clinical reasoning while solving diagnostic question. The audience engagement was more difficult to achieve, requiring creative approaches. Small group discussions followed by assigned, short, public, presentation substantially increased residents' involvement in diagnostic thinking process.

## **Poster Session 2: Practice Improvement Abstracts**

Monday, November 5, 2018 | 7:00 AM - 8:00 AM

#### Cognitive Huddles As Safety Nets to Minimize Risks of Nocturnal Diagnosis in the Emergency Department

C. Crock1,2, H. Singh3

<sup>1</sup>Royal Victorian Eye and Ear Hospital, East Melbourne, Australia

<sup>2</sup>Melbourne Medical School, Fitzroy, Australia

<sup>3</sup>Center for Innovations in Quality, Effectiveness and Safety, Michael E. DeBakey Veterans Affairs Medical Center and Baylor College of Medicine, Houston, TX

Statement of problem: Risks for diagnostic errors in emergency departments (ED) are likely worse at night due to fewer resources and impaired decision-making from fatigue. We developed an intervention based on 'cognitive huddles' to reduce risks related to nocturnal decision-making in the ED.

Description of the intervention or program: Setting was a specialty ED at an Eye and Ear, Nose & Throat hospital receiving 40,000 patientpresentations annually. To reduce risk of harm from cognitive errors on night-shift, we developed a quality improvement safety-net intervention that included a 15-minute cognitive huddle at morning hand-off. Components included 1) dedicated daily time-slot at 745 am for morning huddle of overnight and day team; 2) leadership support to pay certain huddle team members to come in 15 minutes early, 3) standardized procedures to discuss history, exam, tests results, pending studies and management plans for every patient seen overnight (including ones discharged), and 4) if errors were recognized, immediate implementation of amended decisions to prevent patient harm. Led by the overnight ED doctors, the huddle included all incoming junior doctors, senior consultants, nurses and scheduling clerks, and facilitated a protected non-punitive environment for anyone to question decisions of the overnight team. This motivated junior doctors to discuss uncertainty, openly acknowledging "I don't know" not just in huddles but at times to overnight patients with challenging decision-making. These patients were often told to expect call-backs after next-morning's huddle discussion. The morning teams thus acted as safety-nets and contributed their decision-making to overnight differential diagnoses, investigation and management plans. Huddle notes were recorded and analyzed. Findings to date: From 545 huddles conducted between Jan 1 2017-Jun 30 2018, 7740 case-notes were analyzed. Positive impact on the diagnostic process was found in 968 cases (12.5%) and included the following (n = cases): discussion of diagnostic uncertainty (600 cases), change in diagnosis (21) patient call-backs (95), change in management plans (92), additional consultations requested (252), additional tests ordered or pending test results followed-up (81). Qualitative feedback showed that junior night doctors felt supported and confident that challenging diagnoses will be reviewed in the morning and that non-clinicians felt more valuable as team members due to flattening of hierarchies. A culture of open discussion of errors and uncertainty was normalized.

Lessons learned: A cognitive huddle-based intervention led to recognition and mitigation of several risks related to the nocturnal diagnostic process in a specialty ED. Although additional outcome evaluation is pending, findings suggest cognitive huddles could potentially improve safety of diagnostic decision-making in high-risk settings.

#### **Error Reduction Strategy through Error Analysis**

C. Y. DiDonato

Self, Dudley, MA

Nurses have an important role in quality improvement and patient safety initiatives and are a significant part of error prevention and reduction.

Statement of problem: Developing methods to prevent errors from occurring may require recognition and comprehension of how and why errors occur, before an effective plan of action can be implemented.

**Description of the intervention or program:** The aim of this current integrative review was to identify, analyze and synthesize results from the literature that may be impeding rates and quality of medication error reporting. The purpose of this study is to investigate the factors that hinder reporting error and to find methods that could promote and improve those reasons that impede error reporting methods.

Findings to date: A database search produced 16 original research articles, 12 quantitative studies, and four outcomes research studies. Implications for nursing practice would include developing a curriculum or continuing education course dedicated to error analysis.

Lessons learned: The value of promoting and improving error reporting associated with medication errors can potentially give insight that could be employed to develop methods and strategies to reduce medication errors. Reducing medication errors reduces harm to patients and improves quality of care and patient safety. The goal would be to implement error curriculum to improve the awareness of error, the value of accurate and descriptive error reporting, and find possible solutions to improve error rates. Keywords: error, patient safety, evidenced-based, error prevention, error analysis, medication errors, and error reduction strategy.

#### Blood Cultures Practice Improvement: Improving the Volume of Blood Obtained for Routine Blood Cultures

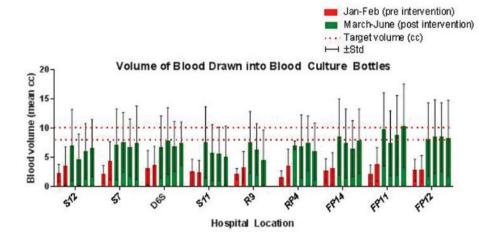
L. Glaser, M. Bulley, K. Eller, K. Hendricks, I. Nachamkin Perelman School of Medicine, University of Pennsylvania, Philadelphia, Phiadelphia, PA

**Statement of problem:** Blood cultures are one of the most common tests performed for the workup of suspected bloodstream infections in hospitalized patients. Several quality metrics serve to monitor pre-analytic processes that affect blood culture test results including blood culture contamination rates and volume of blood collected. For adult patients, collecting 8-10 cc of blood per bottle is the benchmark with a minimum of 2 sets (4 bottles) collected per episode. Insufficient volume reduces the sensitivity for detecting bloodstream infections and could have a negative impact on patient care.

**Description of the intervention or program:** In FY17, we began to monitor blood culture volumes on inpatient units on a monthly basis due to the availability of software enabling easy measurement for this metric. Following baseline data gathering, we instituted educational programs to improve the volume of blood collected.

**Findings to date:** Of 27 hospital units that draw more than 25 blood cultures/month, only 4 units approach or meet the benchmark volume. We discovered that although the contamination rates for phlebotomy team drawn blood cultures are extremely low (<1%), the volume of blood collected was below benchmark by the phlebotomy team. Following an educational program and continued surveillance and feedback to phlebotomists, we observed an immediate and sustained increase on blood volumes.

**Lessons learned:** Analytic tools for measuring pre-analytic variable for blood culture quality are essential for identifying previously unknown quality issues. Continued feedback to stakeholders and education are effective tools for improving laboratory test quality.



# Two-Physician Case Review to Identify and Characterize Diagnostic Error in Seven-Day Readmissions for General Medical Patients

M. A. Kantor<sup>1</sup>, K. E. Raffel<sup>1</sup>, P. Barish<sup>1</sup>, A. M. Esmaili<sup>1,2</sup>, H. Lim<sup>1,3</sup>, F. Xue<sup>1</sup>, S. Ranji<sup>1,3</sup>

<sup>1</sup>University of California, San Francisco, San Francisco, CA

<sup>2</sup>San Francisco VA Medical Center, San Francisco, CA

<sup>3</sup>Zuckerberg San Francisco General Hospital, San Francisco, CA

**Statement of problem:** Approximately 10% of inpatient deaths are related to diagnostic error (DE), and 10% of early readmissions may be related to missed diagnoses. Since DEs are difficult to measure and identify, their incidence may be underestimated and the underlying drivers of DE may therefore be difficult to eliminate.

**Description of the intervention or program:** We aimed to identify the burden of DE in 7-day readmissions and characterize the systems factors that increase the risk for DE in order to prospectively identify high-risk patients and implement improvement efforts. We created a two-physician review of the electronic medical record (EMR) for all patients discharged from an inpatient medical service and re-admitted to our institution within 7 days to any service (7-day readmissions), beginning 1/1/2018. Reviews were completed independently by two board-certified internists using a structured EMR review, including application of the Safer Dx Instrument to the readmission diagnosis. The entirety of the index admission was considered when determining whether a DE had occurred. Identified DEs were further characterized using the Diagnostic Error Evaluation and Research (DEER) taxonomy tool. Two-physician review teams met monthly to examine findings and reconcile discrepancies; all cases of suspected DE were evaluated with the larger group of reviewers (consensus review). If consensus review confirmed a DE, a standardized email was sent to the discharging physician(s) from the index admission to initiate discussion of the clinical decision-making and circumstances that may have contributed to the error.

Findings to date: Of 94 cases reviewed, 11 (12%) contained DE, representing a breadth of diagnoses. Ten (91%) had moderate impact (such as short-term morbidity, increased length of stay, or invasive procedure), and one had mild impact. Reviewers initially agreed on 84 reviews (89%) and 5 cases (45%) found on consensus review to contain DEs. The most common culprit areas within the diagnostic process were Laboratory/Radiology Tests and Assessment; the most common subcategories were failure/delay in ordering needed test(s) [5], erroneous clinician interpretation of test [4], and failure/delay to recognize/weigh urgency [4].

Lessons learned: DE is common among 7-day readmissions. Structured review of 7-day readmissions for DE is a novel mechanism to provide feedback to providers and may be an opportunity to improve diagnostic performance. Case review was feasible to perform, though low concordance on initial evaluation by physician-reviewers of cases found to contain DEs highlights the complexity of clinical care and difficulty in identifying DEs. Additional challenges include the time intensity of reviews, incomplete or ambiguous EMR documentation, and omission of the patient perspective.

# Integrating Genetic Testing into a Clinical Workflow for Long QT Syndrome: Challenges and Opportunities to Improve the Diagnostic Process

I. M. Lubin<sup>1</sup>, E. R. Lockhart<sup>2</sup>, J. E. Frank<sup>3</sup>, V. Y. See<sup>3</sup>, S. Vashist<sup>3</sup>, C. Greene<sup>3</sup> <sup>1</sup>Centers for Disease Control and Prevention, Atlanta, GA <sup>2</sup>Centers for Disease Control and Prevention <sup>3</sup>University of Maryland School of Medicine

Statement of problem: Analysis of a diagnostic workflow provides an opportunity to assess the applicability of professional guidance, setting-specific decisions, and other factors that influence the use and timing of genetic testing for informing clinical management decisions. The diagnosis of Long OT Syndrome (LOTS) was chosen as a model based on the need to appropriately integrate knowledge of the clinical presentation with physiologic and genetic tests in making a timely diagnosis and to reduce harms associated with diagnostic errors.

Description of the intervention or program: A diagnostic workflow was modeled using clinical and laboratory practices at the University of Maryland School of Medicine developed to convey how a teen with loss of consciousness during an athletic activity is generally diagnosed. Cross-referencing workflow elements to existing professional guidance provides insights to the availability and use of professional recommendations for genetic test selection and ordering, use of family history, and integrating knowledge of the test selected and its limitations to the diagnostic process.

Findings to date: Discrete steps in the diagnostic work flow are identified where genetic testing is used to inform aspects of the diagnosis and where errors may lead to delayed or missed diagnoses. Existing professional guidance recognizes the value of genetic testing but does not explicitly describe its integration with physiologic testing and family history in making an LQTS diagnosis. Likewise, professional recommendations are lacking with regards to test selection which is an important issue as the number and types of LQTS-related tests continue to increase. These tests vary by the number of genes analyzed, ranging from less than 10 to over 200. The interpretation of variant pathogenicity, especially for rare findings and those influenced by a family history must be expertly reviewed, often by more than one specialist, to minimize the potential for a diagnostic error.

Lessons learned: LQTS provides a model that demonstrates collaboration among clinicians, laboratory professionals, and genetic counselors can be helpful to clinical decision making and reduce potential harms, especially when professional guidance is lacking or not practical to develop. This is consistent with the use of diagnostic management teams as advocated within the 2015 Institute of Medicine report, Improving Diagnosis in Healthcare. Development and use of a diagnostic workflow diagram has value in understanding the opportunities and challenges for integrating genetic testing into the broader diagnostic algorithm to inform an appropriate diagnosis and reduce the potential for diagnostic errors.

### Structured Case Reviews for Organizational Learning about Diagnostic Vulnerabilities: Initial Experiences from Two **Medical Centers**

B. K. Mathews<sup>1</sup>, R. El-Kareh<sup>2</sup>

<sup>1</sup>HealthPartners and University of Minnesota Medical School, St. Paul, MN

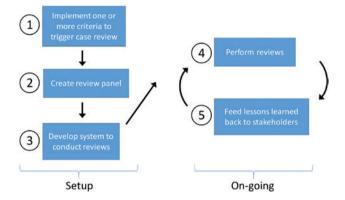
<sup>2</sup>University of California, San Diego, La Jolla, CA

Statement of problem: A key aspect of improving diagnostic safety at the organizational level is to identify and learn from diagnostic errors and near misses. However, few organizations have effective processes to capitalize on these learning opportunities. We describe two attempts to create such a process and describe lessons learned.

Description of the intervention or program: Triggered case review programs were created at two organizations, Site 1 (Regions Hospital, HealthPartners) and Site 2 (University of California, San Diego). At both sites, the goal was to establish an ongoing system to review cases and provide feedback in close to real-time. Both approaches used a similar 5-step process to create the review system (Fig). Step 1: Implement criteria to trigger case reviews-In Site 1, the criteria were developed using local expert opinion whereas Site 2 used criteria created from a modified Delphi process with national and international experts. Step 2: Establish review panel—Site 1 created a panel with hospital medicine physicians, nurse practitioners and physician assistants. Site 2 created a panel with hospital medicine, emergency medicine and surgery physicians. Step 3: Develop system to conduct reviews—Site 1 reviewed cases monthly and performed reviews using paper forms. Site 2 reviewed cases twice a month using an online system. Screening for diagnostic error varied by site and both sites categorized contributing factors using Reilly and colleagues' modified fishbone diagram for diagnostic errors. Step 4: Perform reviews—Site 1 reviewed 85 cases from June 2017 to May 2018. Site 2 reviewed 350 cases from Jan 2016 to June 2017. Step 5: Feed lessons learned back to providers and system—Site 1 provided individual feedback to clinicians and sent monthly "pearls" to the Department of Medicine, summarizing areas for improvement. Site 2 discussed cases and lessons with individuals/teams and system-level lessons were shared with the Patient Safety Committee and Risk Management.

**Findings to date:** The reviews at the two sites identified many opportunities for potential improvement in diagnostic safety [26/85 (42%) at Site 1 and 66/350 (19%) at Site 2]. These opportunities spanned a wide range of scenarios with some common themes of test result management, communication across teams in peri-procedural care and with consultants.

**Lessons learned:** Lessons include: (1) case review with timely feedback provides useful opportunities to learn and calibrate diagnostic decision-making; (2) sharing of cases supports a culture of open discussion of opportunities for improvement and learning; (3) reviews focused on diagnostic safety identify opportunities that complement those found with more traditional peer review.



# Analysis of Delayed Admissions to an Intensive Care Unit -How Can We Avoid Delayed Admissions to an Intensive Care Unit?-

T. Matsumoto<sup>1</sup>, T. Nakanishi<sup>1</sup>, T. Hiroe<sup>2</sup>, S. Watanuki<sup>3</sup>, S. Kosaka<sup>1</sup>

<sup>1</sup>Nerima Hikarigaoka Hospital, Tokyo, Japan

<sup>2</sup>Kyoto University Graduate School of Medicine, Kyoto, Japan

<sup>3</sup>Tokyo Metropolitan Tama Medical Center, Tokyo, Japan

**Statement of problem:** Patients with intensive care unit (ICU) transferred from general wards have higher mortality than those directly admitted from the emergency department (ED). More than 5% of patients admitted through the ED experienced an unplanned ICU transfer; the incidence and rates of transfers were highest within the first 24 hours of admission. In our hospital, we also have seen delayed admission to ICU have a negative effect not only on patients' outcome but also on medical staffs' exhaustion. Therefore, we aimed to evaluate the current causes of delayed admission to ICU and improve their causes like triage at ED, diagnostic errors, and systems.

**Description of the intervention or program:** There are some factors for delayed admission to ICU such as the longer consultation time at ED, and an ICU admission via general ward is also an important factor. In our hospital, we picked up the cases which were transferred to ICU from general wards within 24 hours after admission and evaluate the causes for delayed admission to ICU. Additionally, we choose characteristic two cases and analyzed the factors of delayed ICU transfer by fishbone diagrams.

**Findings to date:** From 1<sup>st</sup> April, 2016 to 31st May, 2018, 34 cases were transferred from general ward to ICU within 24 hours after admission from ED. The 30days mortality rate of delayed admission patients was 23.5% (8/34) and it's twice higher than total ICU admission patients mortality rate within 30 days (11.3%, 123/1085). Delayed admission cases were reviewed and causes are classified into four categories referencing preceding studies of bounces at ED<sup>2)</sup>. Four categories distributed as follows; patient-related 0% (0/34), illness-related 61.8% (21/34), physician-related 35.3% (12/34) and health-care system-related 2.9% (1/34). We determined the physician related three factors by using fishbone diagrams. They are the longer consultation time at ED, the lack of information before consultation and diagnostic error.

**Lessons learned:** An appropriate triage improves mortality. The analysis from 34 cases imply that we can improve the causes of physician-related and system-related delayed cases. Based on these results, we need to renew the system of ER such as human deployment to increase

the number of staff at busy time and educational environment to learn diagnostic process and estimate the severity. In addition to those, we chronologically evaluate the delayed admission to ICU and the patients mortality.

#### Get Rid of the Phrase "Speak up" and Improve Outcomes through the Power of Diverse Perspectives!

L. B. Mooney

Be Highly Reliable, Pomfret Center, CT

**Statement of problem:** Communication problems related to individuals reluctance to "speak up" continue to be a major cause of serious harm and poor outcomes. Attempts to create the psychological safety needed to facilitate "speaking up" have not been successful. Research documents that hierarchies are a problem, nurses perceive open communication to be unsafe and ineffective, role expectations impede speaking behavior and that managers can impact culture around employee voice.

Description of the intervention or program: "Speaking In" is a intentional, invitational leadership practice. "Speaking in" is based in the knowledge that sensemaking, "the making of sense" of the unknown or unexpected is a critical leadership skill key to achieving desired outcomes when working in environments that are rapidly changing, and contain surprises and adaptive challenges. Sensemaking, which is how we determine "what is this?" mostly precedes decision making, when the question becomes "what should be done?" The effectiveness of sensemaking is directly related to the inclusion of diverse perspectives. However, and in order to access diverse perspectives, leaders will have to provide an environment that invites them. The practice of "speaking in" is designed to support leaders at all levels in developing a climate of psychological safety by changing 3 things: 1. Mental model related to speaking: "Speaking In" reframes the purpose of speaking. Sharing a concern, question or idea is reframed a a "gift" contributing to clarity amongst ambiguity instead of a challenge to one's ability or decision making authority. 2. Focus: "Speaking In" redirects attention and voice to the situation at hand, and team verses to a person of hierarchical superiority. 3. Language: New phrase "Speaking In" is introduced to define and support practice. Scripted intentional, invitational questions are taught such as: "What might I not be aware of that you think I need to know?" "What concerns or ideas or questions do you have related to \_\_\_\_\_\_?" "What would you like to see happen?

Findings to date: This practice is in development, but is easily adopted and tried for usefulness by any individual. Pilot study is being developed by Renaldo Blocker PhD of Mayo Clinic

**Lessons learned:** We are currently seeking leaders to trial this practice in a variety of settings.

# Program to Improve Identification of Diagnostic Discrepancies in Preliminary Overnight Radiology Resident Reports on Trauma Patients

<u>T. J. Mosher</u>, M. Bruno, M. Moore, W. Murray Penn State Health, Hershey, PA

**Statement of problem:** Published discrepancy rates on critical findings between overnight radiology resident preliminary reports and final reports by attending radiologists vary from 1%<sup>(1)</sup> to 9%<sup>(2)</sup>. Variation is driven in part by methods used to identify and analyze the clinical impact of the discrepancy. To improve reliability of discrepancy reporting we are developing a technology driven process to evaluate 100% of radiology resident overnight preliminary reports and demonstrate feasibility in analysis of overnight trauma patients evaluated in the emergency department by on-call radiology residents.

**Description of the intervention or program:** Resident preliminary and final reports for all trauma patients seen between 10 pm and 7 am between 3/23/18 and 6/1/18 were exported from the radiology reporting system (Nuance Powerscribe 360) and analyzed electronically to identify any text modification. Differences between report versions were scored for clinical significance using criteria listed in Table 1. Any report scored D to F was subsequently evaluated to identify impact on patient care at 24 hours, 30 and 60 days. Weekly discrepancy reports will be provided to on-call residents to identify patterns of diagnostic errors and to target areas for educational intervention. We are currently developing natural language processing (NLP) and artificial intelligence (AI) tools to automate the analysis to allow the program to be scaled to handle greater report volumes.

**Findings to date:** During the period of analysis 77 of 178 (43%) trauma reports were modified on the final report. Report changes consisted of (A) grammatical edits 30%, (B) non-clinical changes 36%, (C) minor clinical significance 21%, and (D) low clinical significance 13%. There were no discrepancies of major clinical significance (E) or patient harm (F).

**Lessons learned:** Rigorous audits of 100% of preliminary radiology reports is possible using tools present in commercial radiology reporting systems but require human intervention. Development of automated NLP and AI tools will be necessary to scale this program to accommodate larger report volumes and provide timely feedback to trainees and providers.

Table	Table 1: Scoring of Clinical Significance				
Α	No substantive change (ex. editorial/spelling/grammatical change in report)				
В	Minor change of no clinical significance (ex. Addition of fracture nomenclature, trauma				
	grading scale to an otherwise accurate report, addition of incidental findings not directly				
	related to acute patient management)				
C	Minor change of doubtful clinical significance (ex. change in wording that may be interpreted				
	as a change in level of diagnostic confidence, addition of diagnosis in the impression that is				
	appropriately described in the body of the report, recommendation for additional non-				
	emergent studies that may alter diagnosis)				
D	Major change of low clinical significance (Finding related to patient's acute condition that				
	alters patient care but would not have changed overnight clinical management)				
E	Major change of high clinical significance (Clinically significant finding related to patient's				
	acute condition that would have changed overnight clinical management)				
F	Patient harm (Clinically significant finding related to patient's trauma that led to patient harm				
	because of delay in diagnosis or inappropriate treatment based on a wrong interpretation by				
	a radiology resident) Harm assessed at 24 hours, 30 days, 60 days				

# Perceptions of Oncology Patients and Their Providers Regarding Infection Prevention: Towards the Development of an Educational Tool.

G. Nadimpalli<sup>1,2</sup>, L. Maragakis<sup>3</sup> <sup>1</sup>Johns Hopkins Bloomberg School of Public Health, Baltimore, MD <sup>2</sup>University of Maryland, Baltimore, MD <sup>3</sup>Johns Hopkins Medicine, Baltimore, MD

Statement of problem: Cancer patients are vulnerable to infections following chemotherapy when their immune system cannot respond normally to pathogens. The mortality rate of cancer patients due to infections with febrile neutropenia is about 2%-21%. These patients, therefore, need to have sufficient information to take steps to prevent infections, to identify and report any signs of infection to their providers. The chronic nature of the disease makes patient empowerment through education vital especially during transition of care. Lowering the risk of infections is an important strategy to decrease poor outcomes, and effective educational programs may be a useful tool.

Description of the intervention or program: "Patient Empowerment Education Program" (PEEP), is a comprehensive easy to navigate online educational application aimed to overcome the barriers of education for both the providers and patients. The tool will ensure a system improvement in care of cancer patients and facilitate a shift in the paradigm of care from illness to optimum wellness. Salient features of the educational tool "PEEP" are -

- Application will be accessible for patients on mobile, tablet, desktop.
- 2. Assist patients education on Infection Prevention and monitor teach-back.
- 3. Aid patients in timely risk assessment, to identify red flag symptoms, and alert them to contact their provider.
- A reliable source of information for patients and can be used in settings beyond the hospital.

Findings to date: This tool is an outcome of an online survey and cross-sectional study of 61 providers and 21 oncology patients from December 2017- April 2018 at the Johns Hopkins Hospital. The results were- A majority (58.5%) of providers who daily educate reported that patient education is easy or very easy compared with 34.5% of providers who do not daily educate (p = 0.051). The barriers identified by the daily educators were lack of time (55.1%) and no defined format (37.9%). The daily educators frequently shared information on care of central line (93.1 %) in comparison to non-daily educators (62.9%)(p = 0.012). The patients who were daily educated by the provider reported 88.8% satisfaction in contrast to 33.3% satisfaction reported among patients who were not daily educated by their providers (p = 0.082).

Lessons learned: Patients are more satisfied when providers share daily education with them and the knowledge they gain can help prevent infection and improve outcomes. Providers who educate oncology patients daily share crucial information on infection prevention such as central line care with their patients. An easy to navigate online educational application is proposed to facilitate patient education and enhance the continuum of care and infection prevention in cancer patients.

Table1 Descriptive statistics of providers who self-reported they educate patients and comparison of providers who do and do not provide daily education to their oncology patients.

	Overall Educators (n=56)	Daily educators (n=29)	Non-daily educators (n=27)	p- value **
Percent of rate of ease of education as self- reported by providers.				
1-very hard	0	0	0	
2	5.4	0	10.3	0.051
3	47.2	31.0	65.3	
4	36.3	44.8	26.9	1
5- very easy	10.9	13.7	7.69	1
Percent of barriers, encountered by providers during patient education. **				
Improper tools for education	16.0	24.1	7.4	0.08
No defined format	37.5	37.9	37.0	0.945
No monitoring of education	25.0	31.0	18.5	0.28
Lack of time Chart Are	ea 53.5	55.1	51.8	0.80
Identify patient as a barrier	10.7	13.7	7.4	0.440
Percent of topics shared by providers to patients				
during education. **				
during education. ** Central line care	78.5	93.1	62.9	0.00
<u> </u>		93.1 93.1	62.9 96.3	_
Central line care	78.5			0.59
Central line care Hand Hygiene	78.5 94.6	93.1	96.3	0.00 0.59 0.03 0.00
Central line care Hand Hygiene Use of Personal Protective Equipment	78.5 94.6 62.5	93.1 75.8	96.3 48.1	0.59 <b>0.03</b>
Central line care Hand Hygiene Use of Personal Protective Equipment Isolation Precautions	78.5 94.6 62.5 78.5	93.1 75.8 96.5	96.3 48.1 59.2	0.590 0.03 0.000 0.000
Central line care Hand Hygiene Use of Personal Protective Equipment Isolation Precautions Routine surveillance	78.5 94.6 62.5 78.5 51.8	93.1 75.8 96.5 68.9	96.3 48.1 59.2 33.3	0.590 0.03 0.000 0.000
Central line care Hand Hygiene Use of Personal Protective Equipment Isolation Precautions Routine surveillance Spread of germs Percent of various tools as used by provider	78.5 94.6 62.5 78.5 51.8	93.1 75.8 96.5 68.9	96.3 48.1 59.2 33.3	0.590 0.03 0.000 0.000
Central line care Hand Hygiene Use of Personal Protective Equipment Isolation Precautions Routine surveillance Spread of germs Percent of various tools as used by provider while educating patients. **	78.5 94.6 62.5 78.5 51.8 66.0	93.1 75.8 96.5 68.9 75.8	96.3 48.1 59.2 33.3 55.5	0.590 0.03 0.000 0.100
Central line care Hand Hygiene Use of Personal Protective Equipment Isolation Precautions Routine surveillance Spread of germs Percent of various tools as used by provider while educating patients. ** Audio Video	78.5 94.6 62.5 78.5 51.8 66.0	93.1 75.8 96.5 68.9 75.8	96.3 48.1 59.2 33.3 55.5	0.59 0.03 0.00
Central line care Hand Hygiene Use of Personal Protective Equipment Isolation Precautions Routine surveillance Spread of germs Percent of various tools as used by provider while educating patients. ** Audio Video Verbally educate	78.5 94.6 62.5 78.5 51.8 66.0	93.1 75.8 96.5 68.9 75.8 0	96.3 48.1 59.2 33.3 55.5	0.590 0.03 0.000 0.100
Central line care Hand Hygiene Use of Personal Protective Equipment Isolation Precautions Routine surveillance Spread of germs Percent of various tools as used by provider while educating patients. ** Audio Video Verbally educate Paper	78.5 94.6 62.5 78.5 51.8 66.0 0 98.2 48.21	93.1 75.8 96.5 68.9 75.8 0 100.0 55.1	96.3 48.1 59.2 33.3 55.5 100 96.3 40.7	0.590 0.03 0.000 0.100 
Central line care Hand Hygiene Use of Personal Protective Equipment Isolation Precautions Routine surveillance Spread of germs Percent of various tools as used by provider while educating patients. ** Audio Video Verbally educate Paper Educational dummy model Percent of teach back technique used by the	78.5 94.6 62.5 78.5 51.8 66.0 0 98.2 48.21	93.1 75.8 96.5 68.9 75.8 0 100.0 55.1	96.3 48.1 59.2 33.3 55.5 100 96.3 40.7	0.590 0.03 0.000 0.100 
Central line care Hand Hygiene Use of Personal Protective Equipment Isolation Precautions Routine surveillance Spread of germs Percent of various tools as used by provider while educating patients. ** Audio Video Verbally educate Paper Educational dummy model Percent of teach back technique used by the provider while educating patients. **	78.5 94.6 62.5 78.5 51.8 66.0 0 98.2 48.21 17.8	93.1 75.8 96.5 68.9 75.8 0 100.0 55.1 27.5	96.3 48.1 59.2 33.3 55.5 100 96.3 40.7 7.4	0.59 0.03 0.00 0.10 - 0.29 0.28

<sup>\*\*</sup> indicates participant could choose ≥ 1 response

#### **Teleradiograph Interpretation to Reduce Missed Abnormalities of CT Scans**

K. Nakano, K. Akazawa

Shonan Fujisawa Tokushukai Hospital, Fujisawa, Japan

Statement of problem: Computed tomography (CT) scans contain tremendous amount of information. Because we order the scan to see certain areas, abnormality in other areas can be missed. Recently, cases where these misses lead to adverse event have been brought up in the newspaper. Because CT scan is now an essential tool to make a diagnosis, method to reduce this type of miss should be constructed.

Description of the intervention or program: Our hospital started teleradiology interpretations of CTs taken in our ED. All CTs taken in ED are interpreted by radiologists hired by a teleradiology interpretation company. Interpretations are reported in an hour at the shortest and in 2days the longest depending on the payment. To prevent unread reports, ED residents who ordered the scan is required to post the report on the medical record of the patient. ED staffs make sure that reports are posted and check the content. With report coming back several days later, our ED office worker prints out all reports that have any abnormality and the reports are checked by ED staffs. If abnormality that was not found at the ED is detected, the patient is called in for further testing.

Findings to date: Most discrepancy between ED resident and teleradiology interpretation were brain aneurysm, rib fracture, facial fracture, and thyroid abnormality. Many except fractures were not related to the visit and did not lead to an adverse event. While interpretation report

<sup>\*</sup> p-values are calculated using a Pearson Chi-square test; bold indicates statistically significant (< 0.05)

helped residents make the diagnosis in numerous occasions, some reports missed critical illness such as deep vein thrombosis, abdominal bleeding, appendicitis, fracture, and infarction which ED residents found.

Lessons learned: If the radiologist is absent, teleradiograph interpretation is a good way to reduce misses. Physicians must keep in mind that CTs contain information more than you wish for and telegraph interpretation is an effective method to find unexpected abnormalities. Physicians must also note that teleradiograph interpretation is not definite and that they can overlook findings too. Physician have the better hand in that the patient is in front of you and that you have access to laboratory tests and past images. Teleradiograph radiologists have the better hand in that they are specialized in examining in wide view. By knowing pros and cons of teleradiograph interpretation, we can reduce the missed abnormality to the minimum.

#### **How Distracting Are Distractions?**

K. M. Lessner<sup>1</sup>, R. Kalayanamitra<sup>1</sup>, R. Wu<sup>1</sup>, J. Beatty-Chadha<sup>2</sup>, J. A. Neutze<sup>2</sup>, T. J. Mosher<sup>2</sup> <sup>1</sup>Penn State College of Medicine, Hershey, PA <sup>2</sup>Penn State Health, Hershey, PA

Statement of problem: Every specialty relies on radiology to help diagnose and treat patients. There is a high demand to have images interpreted quickly and efficiently. People expect radiologists to increase quantity of images interpreted without decreasing quality of these interpretations. This can become difficult, especially when technologists and clerical staff need information or assistance and colleagues are calling, paging or stopping by for questions. The activation-based goal memory model applies here: radiologists resume prior tasks after distractions. They typically do not immediately pick right back up. Instead, they must first undergo the priming process again, known as resumption lag, which prolongs read-completion time (referenced). These distractions inevitably have negative impacts on radiologists' performances and can lead to more diagnostic errors. We sent a survey to Penn State Health attending and resident radiologists (table 1). Radiologists were asked how often they were distracted, how much distractions impeded workflow and any strategies to overcome distractions. The overwhelming majority reported they felt frequently distracted, usually due to phone calls and pages. They also reported it was difficult to return to original tasks after distractions and it often took them longer to finish an interpretation if they were distracted mid-interpretation. **Description of the intervention or program:** Asynchronous communications allows radiologists to address multiple interruptions. For example, Penn State Health already utilize an instant messaging program called *Primordial* that enables radiologists to wait and respond to messages between interpretations. Recently, Penn State Health has also adopted a new system called CareConnect that allows radiologists to view multiple protocol requests simultaneously and to respond to requests at time of their choosing. More integrative functions such as allowing radiologists to communicate non-urgent study results via electronic notification in CareConnect may allow for more seamless workflow and increased productivity. Findings to date: It is impractical to remove telephones and pagers from reading rooms as radiologists still need to protocol studies and communicate study results to clinicians. Ultimately, there were not many concrete suggestion suggested to prevent distractions or refocus afterwards. Most radiologists reported they restart interpretation processes to ensure nothing gets missed.

Lessons learned: Distractions serve as a significant barrier to radiologists and may result in longer interpretation time and may result in diagnostic errors. Given that interruptions produce a resumption lag or repriming period between the interruption and work resumption, implementing strategies that favor asynchronous over synchronous communication allows radiologists to address multiple interruptions, reducing total resumption lag times. New IT solutions will be key.

H	ow distracting are distractions?
now	elieve workplace distractions may lead to increased diagnostic error. Our goal is to look at often radiologist are distracted throughout their work day and how well they cope with these ctions.
* Rec	sured
Wh	at is your level of training? *
	PGY2
	PGY3
	PGY4
	PGY5
	Fellow
	Attending
	at images do you commonly read? *
	X-rays CT scans
_	Ultrasound
	MRI
_	all the above
	Other:
Ho	w often do you feel that you are distracted throughout the
0	Never
0	Rarely
0	Occasionally
0	Frequently
0	Always
Wh	at is the MAIN source of distraction? (select 1-2) *
	Phone calls/ pages
	Medical staff visits
	Confirming protocols
	Background conversations/ music

	t difficult for you to pick up where you left off in a reading er a distraction occurs? *
0	Never
0	Rarely
0	Occasionally
0	Frequently
0	Always
	you feel that it takes you longer to complete a read due to se distractions? *
0	Never
0	Rarely
0	Occasionally
0	Frequently
0	Option 5
Do	you have any stratgegies for avoiding these distractionsdfg
You	r answer
Do	you have any strategies for refocusing after a distraction?
	ranswer

# A Comparison of Baws (Brief Alcohol Withdrawal Scale) and Ciwa-Ar to Diagnose the Severity of Alcohol Withdrawal in Hospitalized Medical Patients

L. Schwartz<sup>1</sup>, A. Renson<sup>2</sup> <sup>1</sup>NYU Langone Hospital - Brooklyn, Brooklyn, NY <sup>2</sup>NYU Langone Health Center, New York, NY

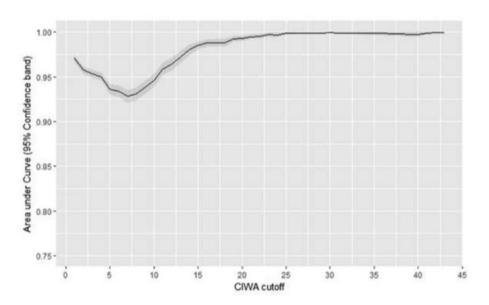
Statement of problem: Treating alcohol withdrawal in the inpatient medical setting requires timely identification of the severity of alcohol withdrawal so that appropriate treatment can be administered. Delayed or missed diagnosis can lead to increased morbidity and mortality, and it contributes to increased cost, length of stay and ICU admissions. The most commonly used measuring tool to quantify the severity of alcohol withdrawal is the CIWA-Ar scale. This scale is lengthy and prone to subjectivity, and it is also more difficult to administer in diverse patient populations where many may not speak English. Many have attempted to devise an alternative to the CIWA-Ar scale that is less subjective and quicker to administer. One such alternative measure is the BAWS scale developed by Darius Rastegar et. al at the Center for Chemical Dependence at the Johns Hopkins Bayview Medical Center. It is a 5-item scale that uses more objective measurements to determine the severity of alcohol withdrawal.

Description of the intervention or program: We created a scale to convert CIWA scores into BAWS scores as a quality improvement project and applied this conversion retrospectively to all patient encounters where a CIWA score was obtained to assess for agreement of the scores and compute AUC. The data was obtained from EPIC computerized records for all inpatient encounters on the medicine service in a 13 month period (6,478 scores in total).

Findings to date: Using BAWS to predict a range of CIWA, we get a high AUC of 0.92. When we compare a BAWS score of 3 to a CIWA of 8 (the thresholds for treatment with benzodiazepines), we get a sensitivity of 0.9 [95% CI 0.88,0.92] and a specificity of 0.8 [95% CI 0.79,0.81]. We also compared higher CIWA scores and found that for CIWA 12 compared to BAWS 6, we get a Kappa score of 0.71 [0.67-0.74] and for CIWA 15 and BAWS 7, we get kappa 0.76 [0.72-0.80].

**Lessons learned:** BAWS can be converted to CIWA scores retrospectively. We conclude that medical providers can use the BAWS scale to assess the severity of alcohol withdrawal and initiate medical treatment in diverse inpatient populations. The BAWS is easier and faster, especially when working with patients who are delirious or who speak different languages. Our findings can raise awareness and acceptance of this tool for assessing the severity of alcohol withdrawal and guiding appropriate treatment with medical inpatients.

# Using BAWS to predict a range of CIWA cutoffs, we get very high AUC – minimum about .92:



#### Institutional Learning from Diagnostic Safety Events - a Pilot Program

<u>B. Smith</u><sup>1</sup>, J. Wright<sup>2</sup>, M. Woodruff<sup>2</sup>, E. McKnight<sup>2</sup> 'Intermountain Healthcare, Salt Lake City, UT 'Intermountain Healthcare

**Statement of problem:** At Intermountain Healthcare, one-third of serious safety events are related to delay in diagnosis. One-half of avoidable patient deaths are related to delayed diagnosis. These diagnostic related events are an important source of learning for our providers, care teams and healthcare delivery system. It can be challenging to capture, analyze and broadly communicate lessons learned from these diagnostic safety events in an effective and systematic way. Here we present a pilot program that identifies and analyzes diagnostic safety events and communicates improvement opportunities across our integrated healthcare system.

Description of the intervention or program: Serious safety events resulting in significant patient harm due to delayed diagnoses are identified through our SafetyNet reporting tool and verified through a physician-led safety event review process. Two cases were selected from our pool of reported events for this pilot. The local safety manager and medical director were asked to perform a root cause analysis (RCA). Uniquely for these diagnostic related RCAs, the providers were asked to review a cognitive bias questionnaire prior to the interview. This questionnaire gave providers time for introspection related to their mental and emotional state during the patient encounter that may have contributed to the diagnostic error. After interviewing the providers and care team, a fishbone diagram was completed describing contributing factors felt to have led to the diagnostic error. Findings were discussed with the providers and care teams. Lessons learned were agreed upon and corrective action items were assigned as appropriate. Then, a de-identified clinical vignette describing the event was created and distributed system-wide as a talking point for all care teams to spread institutional learning and encourage safeguards against future events. Findings to date: The involved providers have responded favorably to the process described above, claiming they feel they will be more effective in preventing similar events in the future. Non-involved providers and care teams who reviewed the case vignettes reported doing so led to meaningful evaluation of their own work environments that could lead to similar negative outcomes. System-level leaders appreciated the vision on diagnostic safety events this process provided. By focusing on diagnostic safety events in this way, we are better able to identify gaps in our care teams and their environments and make changes to more effectively support diagnostic excellence.

**Lessons learned:** Serious diagnostic safety events can effectively be reported and analyzed using a coordinated evaluation approach. Institutional learning can be accomplished by distributing RCA findings as clinical vignettes for care team discussion.

#### **Provider Factors** Investigation/Evaluation Cognitive Biases Leading Factors Lack of knowledge Framing Appears well Anchoring Incomplete medical history high risk Satisficing Pt not 'red flagged' o DM ≈ CAD risk Pt self-selected to Urgent Care Failure to order appropriate (low risk setting) tests o EKG Failure to Initial assessment – low risk Failure to adjust risk perception ppropriate Dx - Costochondritis chest pain requiring ED eval. Leading Factors Investigation/Evaluation Troponin not available in Urgent Care Pt not red-flagged Pt self-selected to IC

# Diagnostic Error Fishbone Diagram

#### Delayed Diagnostic Process Related Claim Exposure Solved with Communication Tracking Numbers.

External Factors

S. Smith

CEO and Founder Complete.MD, Savannah, GA

Statement of problem: Coverys Medical Professional Mutual Insurance Company has identified failure to diagnose as the leading category of Medical Professional Liability (MPL) claim exposure. The majority of these claims emanate from non-hospital settings with non-integrated electronic medical record (EMR) software. In the diagnosis-related cases, process-related factors were noted in 12% of the claims with the most frequent subcategories being lost/misplaced specimens, unreliable transmission of test results, failure to manage the patient's follow-up care, and failure to inform patients of test results. No current software system exists enabling integrated patient/physician/consultant communication across EMR platforms and the care continuum with real-time performance metrics.

Description of the intervention or program: Addressing the Coverys process-related factors, a consortium of pathologist and dermatologist was held to develop a new communication system, Pathologytracker, base on "tracking numbers" created uniquely for each pathology specimen acquired in the exam room and then shared across the care continuum. The tracking numbers were linked to RFID (radio frequency identification) specimen labels enabling physical location tracking and were linked to the pathologist's report, patient communication, patient scheduling, and physician documentation of treatment completion. Additionally, pathologist linked codified "recommendation" codes with the pathology report to direct the next step in patient care. In real-time, all data were shared on a web-based platform with built-in time metrics for every step of the process. The entire care team - patient, physician, and pathologist had access the data with performance metrics. Findings to date: Pathologytracker software was studied from February 23, 2015 to February 23, 2018. There were 51,217 patient visits associated with a total of 108,631 pathology specimens. (average 2 biopsies per patient visit). Cancers including melanoma, basal cell skin cancer, squamous cell skin cancer were identified in 31,338. Only 261 (0.08% 261/31,338) cases remained not treated "Open". All cases had a disposition. Biopsies with non-cancer ICD-10 diagnosis but with pathologist recommendations for re-biopsy or re-excision were 4,833 (4.4% 4,833/108,631) Non-cancer cases remaining "Open" were 121 (2.5% 121/4833). All cases had a disposition.

Lessons learned: Unique tracking numbers generated at biopsy inception can be shared across the care continuum. Utilization of a pathology tracking number platform demonstrated its utility improving clinical processes. The platform eliminated lost/misplaced specimens, assured all results were transmitted, documented all follow-up care, and assured all patients were informed of their results. Expansion of tracking systems into radiology, clinical pathology, and other specialties will have a positive impact on process-related diagnosis claims through enabling communication across the care continuum with non-integrated EMR systems.

#### Learning about Missed Opportunities in Diagnosis from Institutional Risk-Management Data

D. K. Upadhyay<sup>1</sup>, D. Torretti<sup>1</sup>, J. Hunt<sup>1</sup>, S. Korukonda<sup>1</sup>, B. Hohmuth<sup>2</sup>, Y. B. Mehta<sup>2</sup>, P. J. Lamphear<sup>3</sup>, C. Schuerch<sup>2</sup>, A. Khanna<sup>4</sup>, H. Singh<sup>4</sup> <sup>1</sup>Geisinger Medical Center, Danville, PA

<sup>2</sup>Geisinger, Danville, PA

<sup>3</sup>Geisinger, Scranton, PA

\*Center for Innovations in Quality, Effectiveness and Safety, Michael E. DeBakey Veterans Affairs Medical Center and Baylor College of Medicine, Houston, TX

Statement of problem: Most institutions do not systematically analyze their risk-management data related to diagnostic errors. We garnered institutional support and leveraged partnership with patient safety researchers to create a program to analyze our institutions' risk-management data for opportunities to improve diagnosis.

Description of the intervention or program: We created an institutional program for data collection and analysis that included a diagnostic safety champion, an institutional 'Committee to Improve Diagnosis' and a team of 3 individuals (two physicians and one nurse). Initially, selected 'diagnosis-related' cases reported to risk-management office between February 2015 and June 2018 were discussed informally by the champion and committee. Subsequently, with help of a new collaboration with researchers, the-team-of-3 piloted the operational application and use of data collection instruments previously developed by the research team to analyze cases. Their goal was to create practical insights on reasons for diagnostic errors and identify learning opportunities. The team was trained by researchers and piloted the Safer Dx Instrument to identify diagnostic errors as well as another data-collection instrument to categorize which Safer Dx diagnostic process dimensions were involved. Dimensions included 1) patient-provider encounter (history, examination, ordering tests/referrals based on assessment); 2) performance/interpretation of diagnostic tests; 3) follow-up and tracking of diagnostic information over time; 4) subspecialty and referral-specific factors; and 5) patient-related factors. After independent reviews of each case, the team will use consensus methods to resolve differences.

Findings to date: A single reviewer has analyzed 138 cases referred by risk-management thus far, of which 126 were confirmed to involve diagnostic errors with application of the Safer Dx instrument. Initial analysis indicates most errors (96;70%) involved the patient-provider encounter category, of which 32 involved data gathering (12 history-taking and 21 physical exam) and 87 involved data-integration. This was followed by diagnostic test performance and interpretation (n=62), follow-up and tracking (n=8), and referrals (n=7). Detailed independent and consensus reviews are expected to be completed by September 2018 to better understand breakdowns in the diagnostic process. This data will feed into organizational learning with assistance from the champion and committee.

Lessons learned: We created an institutional program to gather and analyze risk-management data using evidence-based data collection instruments for diagnostic errors, previously applied only in research settings. Initial insights reveal several opportunities for improving diagnosis, specifically related to breakdowns in data gathering/interpretation within the patient-provider encounter. Other health care systems should similarly create institutional programs for assessment, and identify contributory factors to generate learning for improvement.

### Identification and Communication of Uncertain Diagnosis in Hospitalized Children

A. Ipsaro<sup>1</sup>, S. J. Patel<sup>1</sup>, D. Richardson<sup>1</sup>, S. T. Chan<sup>1</sup>, D. Warner<sup>1</sup>, E. Luksic<sup>1</sup>, K. Rohrmeier<sup>1</sup>, K. D. Bell<sup>2</sup>, A. Kammer<sup>1</sup>, P. A. Hagedorn<sup>1</sup> <sup>1</sup>Cincinnati Children's Hospital Medical Center, Cincinnati, OH <sup>2</sup>Geisel School of Medicine at Dartmouth College, Lebanon, NH

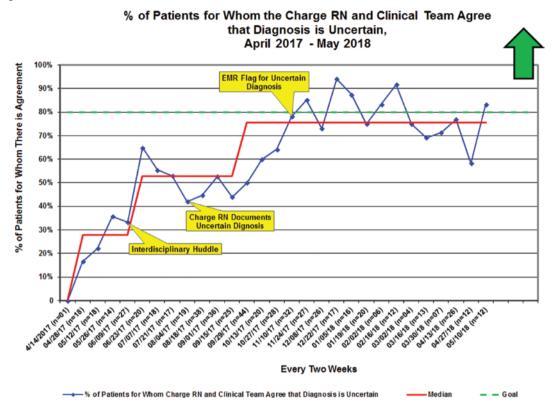
This project will be presented as a poster at the Pediatric Hospital Medicine conference July 19-22, 2018 in Atlanta, Georgia.

Statement of problem: Patients admitted to the hospital are frequently assigned a working diagnosis despite substantial diagnostic uncertainty. Increased provider transitions with shift work may make it difficult to recognize when accumulating evidence suggests a working diagnosis is incorrect. At our center, there was no process to facilitate discussion of diagnostic uncertainty between clinicians, nurses, patients, and families. Without such a process, patients may be at greater risk for diagnostic error.

Description of the intervention or program: This quality improvement project took place on an inpatient medical/surgical unit of a freestanding children's hospital. Our primary measure was percentage agreement between the charge nurse and clinician (physician, APN, or PA) regarding patients with uncertain diagnoses, calculated biweekly. The numerator was patients where both agreed the diagnosis was uncertain, and the denominator was all patients identified as having uncertain diagnosis by either the charge nurse or clinician. We collected data every shift, asking "Do you have any patients with an uncertain diagnosis?" We aimed to increase the percentage of pediatric inpatients concurrently identified as having an "uncertain diagnosis" by both the charge nurse and clinician from a baseline of 0% (with no process in place) to 80%. Utilizing Plan-Do-Study-Act-Cycles, we implemented interventions facilitating interdisciplinary communication, structured handoffs, a culture supportive of questioning, and patient identification in the electronic medical record (EMR). A run chart was used to track percent agreement by two-week increments (see included figure). Established rules of run chart interpretation were used to determine special cause.

Findings to date: The median percentage of patients for whom the charge nurse and clinical team agreed that the diagnosis was uncertain increased from a baseline of 0% to 76% over 13 months. The most impactful interventions were 1) scheduled interdisciplinary huddles for clinicians and charge nurses to discuss patients with uncertain diagnoses, 2) the addition of "uncertain diagnosis" checkboxes to the charge nurse documentation template, and 3) an EMR label for "uncertain diagnosis" that is visible on EMR patient lists and on electronic displays throughout the unit. Days on which a patient was identified as having an uncertain diagnosis accounted for an average of 11% of patient-days during each two-week period.

Lessons learned: Uncertain diagnosis had been infrequently communicated at handoffs between clinicians and nursing staff. Using the Model for Improvement, we increased the identification, communication, and agreement regarding uncertain diagnosis among clinicians and nursing staff.



#### **Epifinder: Helping to Improve Epilepsy Diagnosis**

R. Yao¹, S. Habib¹, H. K. Tarin¹, E. Salinas¹, E. M. Okazaki², J. I. Sirven², A. Z. Crepeau², K. H. Noe², J. F. Drazkowski², M. T. Hoerth², L. Csernak¹, C. Lima<sup>1</sup>, B. Guiang<sup>1</sup>, N. Mehta<sup>1</sup> <sup>1</sup>EpiFinder, Inc., Scottsdale, AZ <sup>2</sup>Mayo Clinic, Phoenix, AZ

Statement of problem: Estimates of epilepsy misdiagnosis are up to 40%, contributing to delay in management and treatment. This high rate is due to the wide array of neurological signs and symptoms, which leads to difficulty in distinguishing between an epilepsy seizure episode or a different condition. It is also due to the fact that epileptic persons are initially assessed by general practitioners instead of epileptologists. A reform in screening strategies is a necessary response to this issue in diagnostic error.

Description of the intervention or program: EpiFinder is a clinical decision support tool developed to strengthen information gathering by healthcare professionals for the purpose of providing more accurate screening for epilepsy. This tool is centered around patient care and works to eliminate errors during the assessment. The application utilizes a heuristic algorithm that takes into account key terms from a patient's medical history to produce a differential diagnosis of various seizure types as well as epilepsy syndromes. The ultimate goal of the application is to improve patient outcomes and improve the diagnostic and therapeutic gap.

Findings to date: EpiFinder's team collaborated with epileptologists from Mayo Clinic Arizona to conduct a clinical study. In the study, 54 patients were assessed using the EpiFinder tool upon admission to Mayo's Epilepsy Monitoring Unit (EMU). The final diagnosis, using the 'gold-standard' continuous video electroencephalogram (cVEEG), resulted in 26 epilepsy diagnoses and 27 non-epilepsy diagnoses. The cVEEG results were compared with EpiFinder's differential diagnosis for validation. EpiFinder's algorithm differentiated between epilepsy and nonepilepsy diagnosis with an accuracy of 86.8%, or 46 of 53 patients. EpiFinder was able to correctly screen for the presence of epilepsy in patients with a sensitivity of 86.4% and a specificity of 85.1%. The study was published in the medical journal Epilepsy & Behavior in May 2018.

Lessons learned: These findings demonstrate the utility of EpiFinder as a clinical decision support tool to aid healthcare professionals to potentially reduce the rate of misdiagnosis, allowing patients to receive proper and timely treatment and management. The significance of this study is that EpiFinder was able to provide a diagnosis before the patient underwent the diagnostic procedures performed by a neurologist (or epileptologist) as these procedures are costly and time consuming.

# **Poster Session 2: Clinical Vignettes**

Monday, November 5, 2018 | 7:00 AM - 8:00 AM

#### Pain in the Neck

A. Salim<sup>1</sup>, K. Atiq<sup>2</sup>, <u>J. Akhtar<sup>3</sup></u> <sup>1</sup>Dow University of Health Sciences, Karachi, Pakistan <sup>2</sup>United Medical and Dental College, Karachi, Pakistan <sup>3</sup>UPMC, Pittsburgh, PA

Learning objectives: Common symptoms usually point to common diagnosis. However, it is important to be aware of rare diseases which may present similarly to make a timely diagnosis. We report a case of acute calcific tendinitis of the longus colli muscle, a rare inflammatory process of the prevertebral muscles.

Case information: A 35-year-old woman presented to the emergency department with eight days of right sided neck stiffness. She developed sore throat a few days later with painful swallowing. This was followed by pain on the left side of the throat with low grade temperature. On examination, she had a heart rate of 110 beats per minute and a temperature of 38.4. Neck showed anterior neck tenderness and pharyngeal edema with erythema. There was painful range of motion of the cervical spine. Laboratory showed a mild leukocytosis with a WBC count of 11.2 per cubic millimeter. CT scan of neck with contrast done to rule out retro-pharyngeal infection showed acute calcification of the left longus colli muscle. Patient was discharged home on non-steroidals.

Discussion: The clinical picture includes acute neck pain, limited range of motion of the cervical spine with stiffness and odynophagia. However, this condition is often confused with life-threatening conditions such as infection (meningitis or retropharyngeal abscess), intracranial hemorrhage, trauma, herniation of cervical discs, or malignancy. Knowledge of the characteristic radiologic findings and the self-limiting course prevents the patient from needless medical and surgical interventions.

#### Superior Vena Cava Syndrome Misdiagnosed As Allergic Reaction

O. Chehab<sup>1,2</sup>, A. Kanj<sup>2</sup>, H. Tabaja<sup>2</sup>, N. Abdallah<sup>2</sup>, S. El Zein<sup>2</sup> <sup>1</sup>Imperial College London, London, United Kingdom <sup>2</sup>Wayne State University, School of Medicine; Detroit Medical Center, Detroit, MI

Learning objectives: Recognize superior vena cava (SVC) syndrome in a patient with facial swelling

Case information: A 31-year-old woman with sickle-cell disease and a history of moyamoya disease, presented to the emergency department after a syncopal episode and mild facial swelling following a routine exchange transfusion. The patient receives exchange transfusions every four weeks through a venous access port on the left side of her chest. She reported that on her previous transfusion episodes she developed facial edema which was relieved by benadryl. The patient was started on intravenous benadryl treating a presumed allergic reaction. Prednisone was then added, with no improvement. Workup for hereditary angioedema was unremarkable. The patient's condition worsened with progressive facial swelling and shortness of breath requiring intubation. A CT-thorax done to evaluate for SVC obstruction showed multiple venous collaterals. A duplex scan of upper extremities and neck veins showed chronic deep vein thromboses (DVTs) involving the right and left brachial veins and an acute DVT involving the right axillary vein. In addition, there was note of a chronic DVT in the left axillary vein near the venous port. The patient was started on unfractionated heparin managing SVC syndrome and was extubated 3 days later after a decrease in facial swelling.

Discussion: The vast majority of SVC syndromes are caused by mediastinal tumors. Other less common causes include vascular diseases, infections and thrombosis related to the presence of central venous catheters, such as venous ports. The obstruction of the SVC can result in collateral venous return mainly through the azygous and internal mammary venous systems. SVC syndrome often presents with facial and upper extremity swellings. This presentation can mimic angioedema or an acute allergic reaction. Our patient's history of possible allergic reaction following previous transfusions led to an anchoring bias, despite her mother mentioning that her reaction occurred once without transfusion.

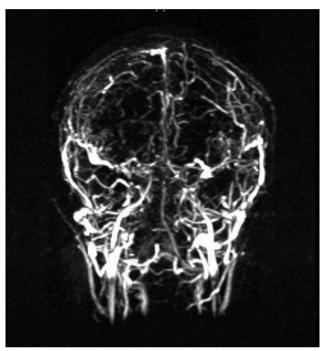
# Triage-Based Physician Anchoring: How Patient Rooming Decisions May Affect Patient Outcomes

Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL

**Learning objectives:** To discuss cognitive bias as it pertains to medical decision making and patient outcome.

Case information: A 15 year old male with a history of insulin dependent diabetes mellitus presented to a community hospital emergency department for evaluation of unusual behavior. At triage, the patient stated that he wanted to "kill myself and my mother". Initial BG 190 by finger stick. Pt was placed in a psychiatric evaluation room and waited 30 minute for a provider. Initial VS: HR 124, BP 151/96, remainder normal. Initial exam normal except for "odd statements". A psychiatric evaluation and basic laboratory studies ordered. 4 hours later patient's labs normal. He was noted to have continued altered mental status and elevated HR. Urine toxicology study sent and negative and psychiatry unable to perform evaluation. At change of shift a new MD ordered head CT which was abnormal. An MRI was delayed for an additional 3 hours due to need for sedation. The MRI result showed diffuse Cerebral Venous Sinus Thrombosis with herniation. Patient was intubated, started on anticoagulation protocol and airlifted to an appropriate pediatric intensive care unit for continued management.

Discussion: This case illustrates several cognitive biases at work. In this case, the patient's initial suicidal comment led to placement of this patient in a behavioral health room. It has been shown that placement of a patient in a non-medical room may not prioritize them appropriately. This phenomenon, known as "framing", caused a series of events which delayed his diagnosis and initiation of life-saving therapy for his condition. Additional cognitive errors such as the Availability heuristic, Premature closure and Anchoring contributed to the outcome of this case. It is important to be aware of how our own innate biases affect us and how initial labeling may have an impact on treatment, particularly in the emergency setting.



#### Negative d-Dimer? Yes. Positive Pulmonary Embolism? Absolutely!

Cook County Health and Hospital System, Chicago, IL

Learning objectives: Understanding the limitation of d-dimer testing and how it may impact patient outcome. Being able to recognize that pulmonary embolism may still be a working diagnosis in the setting of a negative d-dimer.

Case information: 48 year old Vietnamese speaking male with unremarkable past medical history who presented with 2 months of worsening dyspnea, bilateral chest pain, and dizziness. Patient with low health literacy and significant language barrier despite interpreter phone service. Nontoxic appearing at time of presentation. Negative workup in the department which include unremarkable EKG, normal chest-xray, neg initial trop, normal d-dimer, BNP and hemoglobin that was within normal limits. Patient admitted to hospital and later found to have bilateral pulmonary embolism as the source of his dyspnea. Patient thankfully had a good outcome requiring lifelong anticoagulation for unprovoked pulmonary embolism.

Discussion: Limitation of d-dimer includes type of assay used in Emergency Medicine departments. ELISA is gold standard with sensitivity nearing 97% but takes upwards of 2-4hours to result. This is clearly not ideal in our fast paced environment. Therefore most departments rely on less sensitive assays with sensitivities in the mid 80%. Other confounders include length of time of presentation. Several case reports documenting normal d-dimers in the presence of clinically relevant pulmonary embolism. There appears to be an inverse relationship with the time of presentation and d-dimer level. Lastly, d-dimer can be normal in patients who present with prior pulmonary embolism. This can be a challenge to assess if patients with "first" pulmonary embolism are not seen within a reasonable time frame of symptom onset.

#### The Fallacy of "TB Rule-out": Negative Afb Smear/Culture and Bayesian Reasoning Error

A. M. Esmaili1, K. E. Raffel2 <sup>1</sup>San Francisco VA Medical Center, San Francisco, CA <sup>2</sup>University of California, San Francisco, San Francisco, CA

Learning objectives: 1. To appreciate the multiple patient, laboratory, and institutional factors that can contribute to missed or delayed diagnosis of pulmonary tuberculosis (TB), 2. To understand the sensitivity and limitations of diagnostic testing for pulmonary TB,

Case information: A 83-year-old woman of Chinese origin with remote history of breast cancer and pulmonary fibrosis of unknown etiology presented to care with one month of fevers, night sweats, malaise, and cough. After three negative acid-fast bacillus (AFB) smears during hospital admission, she was discharged on oral antibiotics for community acquired PNA. Several weeks later, she was admitted to the hospital with persistent fevers, night sweats, and malaise. Working diagnoses included infectious, malignant, and inflammatory etiologies. ID was consulted several days into her hospital course and recommended repeat AFB smear with nucleic acid amplification test (NAAT). She was discharged again after three negative AFB smears and with her tuberculosis NAAT pending at time of discharge. She had planned follow-up after her discharge with her primary care provider. Two days after discharge, her primary physician was notified of the positive NAAT result. Discussion: Pulmonary TB is a must-not-miss diagnosis: the individual and public health implications are morbid if diagnosis is missed or delayed. Multiple levels of 'risk factors' -- provider cognitive biases, limited sensitivity of microbiological testing, and institutional inexperience -- can contribute to delayed or missed pulmonary TB diagnosis. In this case, delayed TB diagnosis was strongly attributable to provider anchoring on prior negative AFB smears (a test with limited sensitivity). Because the three utilized tests for pulmonary TB (AFB smears, AFB cultures, and NAAT) all have significant limitations in sensitivity, institutions have established unique testing algorithms. To avoid missed or delayed diagnosis of pulmonary TB, individual providers must seek to understand both the limitations of microbiological testing for pulmonary TB and their own institution-specific testing algorithms.

#### **Delayed Diagnosis of Meningoencephalitis**

K. A. Godil1, S. Godil2 <sup>1</sup>Lehigh Valley Health Network, allentown, PA <sup>2</sup>UPMC Mercy, Pittsburgh, PA

Learning objectives: Delayed diagnosis of HHV-6 meningoencephalitis due to Anchoring Bias.

Case information: 28-year-old female with no significant past medical history initially came to the Emergency Room with a headache. The headache was 10 out of 10 intensity and radiated to her neck. Ct head and bloodwork were normal and she was given Fioricet for a tension headache and discharged home. The next day she came back to the Emergency Room with her family for complaints of altered mental status and again was discharged home after normal CT head and blood work. She then came back to the emergency room, a few hours later, for the third time with altered mental status and respiratory compromise. CT head was negative for Subarachnoid hemorrhage, and then lab work showed an increasing white count. She then developed an episode of left face and leg numbness and was subsequently started on seizure prophylaxis. A lumbar puncture was done which was positive for HHV-6 on PCR as well as xanthochromia. A diagnosis of HHV meningoencephalitis was made. **Discussion:** Even though a case of HHV-6 meningoencephalitis is a rare occurrence in an immunocompetent individual, a high degree of clinical suspicion should be involved for meningitis when a young person comes in to the ED with classic signs for meningeal irritation, even if initial blood work and imaging were normal. This is a classic example of anchoring bias. The initial presentation may have seemed as if our young patient had a tension headache, however on her second visit to the Emergency Room, when she presented with Altered Mental Status, further differentials should have been sought.

#### The "Patient Zero" Challenge

V. Gonzalez<sup>1</sup>, J. Lu<sup>1</sup>, M. Mycyk<sup>1</sup>, S. Aks<sup>1</sup>, K. Cosby<sup>2</sup> <sup>1</sup>Cook County-Stroger Hospital, Chicago, IL <sup>2</sup>Cook County Health and Hospitals System, Chicago, IL

Learning objectives: Recognize that some medical diagnoses are novel and unlikely to be identified at the first visit. Identify unusual cases by being curious, asking probing questions, and maintaining general awareness of public health outbreaks. Unusual features of common medical conditions should alert the clinician to consider a broader differential.

Case information: A 44 year-old woman with no past medical history presented with 3 days of left flank pain and gross hematuria. Patient has 20-pack year smoking history and she denied drug use. Labs showed Hb 8.4 and a urinalysis "loaded" with RBCs. CT scan showed no hydronephrosis, and she was diagnosed with a passed urinary stone and discharged. The patient returned one day later for persistent bleeding from IV site. Coagulation studies revealed PT >120s, an unquantifiable INR, and worsening anemia. After repeated questioning, she admitted to synthetic cannabinoid use at her niece's birthday party. The patient admitted to ICU with severe coagulopathy and treated with prothrombin complex concentrates and vitamin K.

Discussion: This patient represents one of the first cases in the Chicagoland area of severe acquired coagulopathy following exposure to synthetic cannabinoids containing a long-acting vitamin K antagonist. Despite press releases, clinicians managing earlier cases were operating under "patient zero" conditions. The early outbreak time period is ripe for diagnostic error. In this case, a passed stone did not adequately explain the gross hematuria and anemia. These "patient zero" cases are challenging because the clinician must synthesize the presentation in the context of established medical knowledge, but also consider that the presentation may be a novel medical diagnosis. This requires meticulous history taking, creativity, and reconsideration of the diagnosis as the case evolves. Novel medical diagnoses are commonly seen in infectious disease and toxicology, but are not limited to these fields, and a high index of suspicion and critical thinking is essential.

#### A Diagnostic Delay Case of Nonconvulsive Status Epilepticus with Acute Behavioral Changes in Elderly

T. Harada

Showa Medical University Hospital, Tokyo, Japan

Learning objectives: Physician should consider nonconvulsive status epilepticus at any time in acute mental and / or behavioral changes in elderly

Case information: A 70-year-old man was referred emergency department for mildly hyponatremia and altered mental status. He lived alone and completely dependent on the insurance system. In addition, since his welfare official met two months ago, history during was absolutely unknown. He was disoriented and could not communicate and behaved like a child, showed automatisms, delusional symptom, talk to oneself, shout loudly with agitation. Physical examination showed no other neurological abnormality. Laboratory data, urinalysis, cerebrospinal fluid, and head images studies were unremarkable other than hyponatremia(128mEq/L) and mild hippocampal atrophy. Consultation for psychiatrist result was judged as BPSD like. We don't perform emergency electroencephalogram because of mildly consciousness, no history of central nervous disease and the result of psychiatrist consultation. However, on admission day8, electroencephalogram performed on standby and detected epilepsy. He was diagnosed with nonconvulsive status epilepticus(NCSE) and transported to a specialized facility **Discussion:** In this case there are mixed various diagnostic error factors, process of data integration, data gathering, especially confirmation bias and overconfidence effect. To recognize that NCSE is a disease easily lead to diagnostic error is important. The clinical symptom of NCSE is very various, mild to severe, positive symptom and negative symptom and cannot diagnosis unless perform electroencephalogram. Moreover, the sensitivity of the history of epilepsy is less than half, and even including the history of central nervous system disease is about 75%. It is difficult to diagnose NCSE from the various factors mentioned above, but NCSE is a treatable disease and early diagnosis and treatment are important. As in this case, NCSE shows acute behavioral change as the main symptom, and physician should consider NCSE at any time in acute mental and / or behavioral changes in elderly

# Listen to Your Patient, He Is Telling You the Diagnosis: Delayed Legionella Diagnosis in the Setting of a Known **Outbreak**

N. Hefzi<sup>1</sup>, K. J. McCollum<sup>2</sup>, M. J. Miller<sup>2</sup>, P. Shah<sup>3</sup>, O. Awad<sup>3</sup>, D. L. Levine<sup>1</sup>, P. Tannous<sup>3</sup>

Wayne State University, School of Medicine; Detroit Medical Center, Detroit, MI

<sup>2</sup>Wayne State University, Detroit, MI

<sup>3</sup>Wayne State University, School of Medicine, Detroit, MI

Learning objectives: Maintain a high index of suspicion for specific illnesses during known outbreaks of infections, obtain appropriate histories to ensure timely diagnosis, and initiate targeted treatment to include coverage for the likely pathogen.

Case information: A previously healthy 32-year-old man with a 7 pack-year smoking history presented with productive cough and shortness of breath. He was prescribed Augmentin. Several days later, his physician advised him to report to the ED when his x-ray revealed pneumonia. On presentation patient described progressively worsening shortness of breath and malaise. Creatinine was 7.7 mg/dL, BUN 95 mg/dL, CPK 3191 U/L, hemoglobin 9 g/dl, total bilirubin 3.2 mg/dL. Urine was positive for myoglobin but negative for casts. Patient was admitted to the ICU and started on Vancomycin and Cefepime. Later, a positive Legionella urinary antigen was reported. The patient was switched to azithromycin and treated aggressively for acute renal failure (ARF) with improvement. He was discharged with residual renal dysfunction.

Discussion: Legionella-associated pneumonia is a rare but fatal cause of ARF and rhabdomyolysis.1 This patient worked in a location where an outbreak of Legionella had recently been reported. It is particularly important to obtain appropriate histories, including work/exposure histories, in the setting of known outbreaks. If this patient's physician had been aware of the outbreak, Legionella may have been diagnosed earlier possibly preventing ARF and an ICU admission. Likewise, had the ED physicians obtained an appropriate history in the setting of a local outbreak, proper antibiotics would have been started instead of vancomycin and cefepime. Maintaining a high index of suspicion for Legionella during outbreaks and appropriate use of treatment guidelines for pneumonia, including DRIPS2 (Drug Resistance in Pneumonia Clinical Prediction Score), would have ensured this patient received antibiotic coverage for this potentially lethal pathogen. 1. Scand J Urol Nep hrol,vol.40,no.4,pp.345-346,2006. 2.Antimicrob Agents Chemother. 2016 Apr 22;60(5):2652-63.

#### Diagnostic Challenges When Taking Care of Tourists: A Case in Point

M. Imamura<sup>1</sup>, M. Saijo<sup>1</sup>, K. Moroi<sup>2</sup>, U. Nakagawa<sup>1</sup>, S. Constantine<sup>1</sup> <sup>1</sup>Sapporo Tokushukai Hospital, Sapporo, Japan <sup>2</sup>Niseko International Clinic, Abuta-gun, Japan

**Learning objectives:** To describe the challenges faced when making diagnosis in tourists visiting Hokkaido, Japan.

Case information: A 67-year-old Singaporean woman arrived to Hokkaido for sightseeing. She had been having cough for 1-month prior to arrival, but her symptoms worsened significantly on the plane when she developed dyspnea. She was found to be hypoxic and the chest radiograph demonstrated a left pulmonary infiltrate. She was diagnosed with pneumonia and received intravenous antibiotics in a local travel clinic for 2-3-days. The patient did not want to go to a hospital, however, given the lack of improvement, she was transferred to our institution. Due to anchoring bias and financial concerns, a CT was not ordered at admission. Two days later, we noticed bilateral lower extremity edema and later found to be a lower leg deep venous thrombosis. An enhanced chest CT demonstrated pulmonary embolism, pericardial effusion, irregular masses in her left lung, breast, adrenal glands and liver suggesting metastatic cancer. We initiated anticoagulation therapy and talked to the family about our findings and recommended to travel back to Singapore as soon as possible before other complications ensued. They agreed and after a prolonged process, she traveled back home through a private patient transfer company.

Discussion: Tourism in Japan is increasing as well as the chances to be in contact with travelers suffering all kinds of illnesses. In 2016 alone, Japan had more than 24 million visitors and this number will increase given the Tokyo 2020 Olympics. Obstacles for accurate and prompt diagnosis include the language barrier, the lack of access to previous records and many times, financial restrictions for testing and treatment from patients, families and hospitals. Having established institutional protocols can result in a streamlined process to handle these patients which would be beneficial for both patients and doctors.

#### Vertbral Osteomyelitis Misdiagnosed As Urinary Tract Infection.

A. Ishiguro<sup>1</sup>, S. Miyawaga<sup>1</sup>, S. Nagasaki<sup>1</sup>, I. Kitagawa<sup>2</sup> <sup>1</sup>Shonan Kamakura General Hospital, kamakura, Japan <sup>2</sup>Shonan Kamakura General Hospital, Kamakura, Japan

**Learning objectives:** Vertbral Osteomyelitis presenting with costovertebral angle tenderness.

Case information: A 85 year-old man who is a nursing home resident, with past medical history of lumbar spinal stenosis came to our hospital with chief complaint of fever and general malaise. On admission day, he even could not eat and had incontinence, so his family called an ambulance. On arrival, vital signs were BP140/69mmHg,HR83bpm,BT38.1 ,RR20/m, Physical examination showed costovertebral angle tenderness. The patient had elevated WBC of 12800,CRP of 7.65, urinalysis showed bacteriuria. The patient was admitted to internal medicine as urinary tract infection. After admission, another physician repeated physical examination. The patient had sacral decubitus and tenderness around L4/L5 and CVA tenderness. In addition, gram stain of urine showed gram positive cluster. On suspicion of vertebral osteomyelitis, MRI was taken. MRI showed increased signals between L4 and L5, which were consistent with vertebral osteomyelitis and discitis. The patient was treated with cefazolin 6g per day. On echocardiogram, he had no sign of infectious endocarditis. After 6 weeks of intravenous antibiotic treatment, the patient was discharged.

Discussion: Misdaignosing oseteomyelitis as urinary tract infection at our emergency department was caused by anchoring bias and overconfidence of physical examinations and typical medical history of urinary tract infection. What we can learn from this case is that we should listen to what exactly patients tell us and repeat basic physical examinations.

### **True Diagnosis of Transient Disturbances of Consciousness**

T. Kawahigashi<sup>1</sup>, M. Tokura<sup>2</sup>, I. Kitagawa<sup>1</sup> <sup>1</sup>Shonan Kamakura General Hospital, Kamakura, Japan <sup>2</sup>Shonan Kamakura General hospital, Kamakura, Japan

Learning objectives: Intracardiac thrombosis is a rare phenomenon, but any delay in treatment could be lethal.

Case information: An 88-year-old Japanese woman presented to the emergency department with a disturbance of consciousness. Her caregiver noticed that she did not respond to the call one hour before her visit. Her symptoms had disappeared when she arrived to the hospital. However, when she was about to go home, she presented with new symptoms including left facial paresis, arm drift, and abnormal speech. Although we evaluated her by performing laboratory examinations, as well as computed tomography and magnetic resonance imaging of the head, these modalities did not reveal any specific finding and her symptoms completely disappeared within a few minutes. She was subsequently discharged. On further examination, an echocardiogram showed a large free-floating mass (69 × 11 mm) in the right atrium, which was protruding into the right ventricle (figure). Pulmonary hypertension was also detected (mean pulmonary artery pressure = 45 mmHg); she died 5 hours after arriving to the hospital. A full body autopsy revealed that she died from a large thrombus in the right atrium and bilateral pulmonary emboli, which made her cardiac output lower and led to both the disturbance of consciousness and stroke-like symptoms.

Discussion: Right atrium thrombosis is uncommon, but possibly underdiagnosed. Pulmonary embolism is often accompanied by right heart thrombus, which is associated with a high (>40%) mortality rate. Echocardiography is a key examination modality for this setting. In this case, we missed a life-threatening condition due to the rarity of this condition, nonspecific symptoms and cognitive bias, premature closure. Physicians need to have a high level of awareness of this fatal condition and pursue the appropriate differential diagnosis when facing complex presentations.



#### An Axillary Schwannoma Mimicking a Malignant Lymphoma – a Case of Anchoring to a Diagnosis?

I. Kitagawa, J. Branch

Shonan Kamakura General Hospital, Kamakura, Japan

Learning objectives: The application of the dual process model and avoiding cognitive bias will enable physicians to diagnose a Schwannoma as a differential diagnosis of an axillary nodule, mimicking lymphadenopathy.

Case information: A 39-year-old man with no significant medical history presented with a two-week history of a non-tender nodule in his left axilla. It was also noted that he had a left inguinal nodule that was unchanged in size for the last 3 years. He denied any recent sexual history, contact with animals or exposure to unwell persons. On examination of the axillae, a soft, non-tender nodule was palpated of approximately 3 cm in diameter. Ultrasonography revealed a well-defined, 2 cm nodule suggesting a lymph node as a most likely cause. Despite the absence of other symptoms, the physician suspected a malignant lymphoma (ML) and undertook a biopsy of the nodule. The biopsy result revealed a Schwannoma, not a ML.

Discussion: Cognitive bias should be avoided in every diagnostic case. In cases which are diagnostically-challenging, the cognitive bias is one of the key factors of diagnostic errors. This case exhibits an example of misdiagnosis, revealing the process of a typical cognitive error anchoring. The physician in this case was anchored to an intuitive diagnosis of lymphadenopathy. However, the true diagnosis was not what was expected. In such a case, debiasing measures play a key role to fill the gap between correct and incorrect diagnosis. Here it is emphasized the usefulness of a cognitive debiasing strategy, which can be universally applied to avoid unexpected misdiagnosis, improving diagnostic accuracy and quality of patient care. We believe that the dual process model and cognitive bias plays a pivot role for debiasing cognitive error and helping physicians to make efficient and effective diagnoses.

# Anchor Aweigh! A Diagnostic Error Calling for Standardized Chest-Ultrasonography Training for Internal Medicine Residents

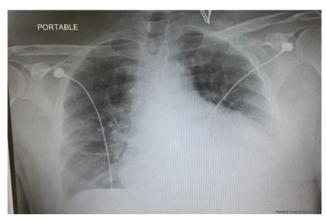
L. E. Krowl<sup>1</sup>, A. S. Dhamoon<sup>2</sup>

<sup>1</sup>State University of New York (SUNY) Upstate Medical University, Syracuse, NY <sup>2</sup>State University of New York (SUNY) Upstate Medical University, Syracuse, NY

#### Learning objectives: see prior page

Case information: Elderly male presented to emergency department with increased shortness of breath and anuria. Vital signs were normal though exam showed bibasilar crackles, jugular distention and lower extremity edema. Labs showed creatinine 8.0 mg/dL (baseline 1.1), blood urea nitrogen (BUN) 110 mg/dL, Lactic Acid 10.9 mmol/L. Chest radiograph showed cardiomegaly with bilateral pleural effusions. Patient was started on furosemide infusion for cardiogenic shock secondary to heart failure. Nephrology and cardiology were consulted, and the medicine resident was called to admit the patient. Nephrology sees the elevated creatinine, BUN with nil urine output and recommends fluids. Cardiology sees the chest radiograph and recommends continuing furosemide. Given the contradiction, the medicine resident used bedside ultrasound to assess patient's volume status. Inferior vena cava was plethoric; there was large pericardial effusion and right ventricular collapse. Cardiology then performed pericardiocentesis because of impending collapse due to cardiac tamponade.

**Discussion:** With the advent of electronic medical records physicians are spending more time assessing data from electronic records and less time assessing patients at bedside. Ultrasonography amplifies physical exam beyond the limits of a stethoscope; and mastering basics of ultrasonography during resident training, gets physicians back to the bedside. Residents are becoming more familiar with using ultrasound for procedures mandated by Accreditation Council for Graduate Medical Education for graduation, yet using ultrasound to amplify the physical exam is practice limited by resources. While cardiac tamponade is a less frequent cause of shock, it must be ruled out prior to diuresis. A clinician competent with basic chest ultrasonography may be able to make a life-saving diagnosis of cardiac tamponade. Ultrasonography training is self-taught in many residency programs, which poses questions of competence. Reporting cases such as this may lead to a greater call-to-action for standardized chest-ultrasonography training across all medicine residency programs.



#### "out of Sight, out of Mind": Delayed Diagnosis of a Necrotizing Soft Tissue Infection

D. McCurry<sup>1</sup>, D. DiNardo<sup>2</sup>

<sup>1</sup>University of Pittsburgh Medical Center, Pittsburgh, PA

<sup>2</sup>University of Pittsburgh School of Medicine, Pittsburgh, PA

#### Learning objectives:

- 1. Define posterior probability error and its similarity to the gambler's fallacy.
- Evaluate the accuracy of the laboratory risk indicator for necrotizing fasciitis (LRINEC) score and review attempts to improve upon its diagnostic value.
- 3. Define two variations of the availability bias and how strategies to counter each variation differ.

Case information: We report a case of diagnostic delay of a necrotizing soft tissue infection (NSTI). During a routine outpatient oncology visit, providers incorrectly diagnosed a recurrent peri-rectal abscess without imaging or surgical evaluation, despite significant exam and laboratory abnormalities, leading to diagnostic delay and inappropriate triage of illness severity. Hours later, the patient was admitted to an inpatient oncology service where diagnostic reevaluation, including calculation of a LRINEC score, raised suspicion for NSTI, resulting in urgent imaging, NSTI diagnosis, and surgical intervention. During multiple recent emergency department visits prior to this presentation, the patient was repeatedly diagnosed with a peri-rectal abscess.

**Discussion:** We hypothesize that the patient's history of emergency department visits led to a type 1 clinical reasoning failure, specifically posterior probability error leading to premature closure. In NSTI, where early recognition is challenging but crucial to survival and recovery, this case illustrates how adjunctive scoring systems, such as the LRINEC, can be used to counter biases in type 1 reasoning and prompt type 2 thinking leading to urgent reevaluation. However, given that NSTI is not commonly encountered in the outpatient oncology setting, we argue that the non-availability formulation of the availability bias, the "out of sight, out of mind" bias, likely contributed to diagnostic delay, as well. While more reflective and methodical thinking has been shown to mitigate the classic formulation of the availability bias, the non-availability bias can persist despite the use of type 2 reasoning, making this case a diagnostic challenge on multiple levels.

#### Careful Language at Handoffs: Noticing How Labels Impact Thinking

A. Miller

201 Dromara Rd, Guilford, CT

**Learning objectives:** 1. Appreciate the value of making careful observations and describing findings with accuracy 2. Describe how problem representation influences seeing

Case information: A 1914 work of art by Henri-Gaudier Brzeska (attached) serves as a case to teach key lessons about the strengths and limits of human cognition in the challenge of accurate language use in problem representation and at handoffs. Descriptions of this image in CME settings reveal many paradoxes can be seen: enemies battling, lovers at rest, brutal disfigurement, a sexual encounter, non-sexual play, former lovers parting ways, same-sex interaction, heteronormative interaction, an ancient relic of bodies preserved in volcanic ash, a modern-times nod to the past. Upon first glance, many remark "they are wrestling." The title is indeed "The Wrestlers," which is revealed only after careful inspection of the picture.

Discussion: An emerging body of studies evaluating patient records as they transmit uncertainty and bias indicates that words have the potential to influence the subsequent person's understanding, curiosity level, and interest in a case. It is indeed a challenge to place words to what we see with both accuracy and room for development. An art piece rife with ambiguity and paradox gives an opportunity to appreciate the importance, complexity, and satisfaction of this challenge: "seeing" multiple meanings unfold and then receiving the title only afterwards prompts remarking on propensity to close (would you have looked twice at the picture had you known the title?) and the value of forcing strategies (does the title definitively mean that it is wrestlers presented here? What else could it mean?). Separating fact from inference is an advanced clinical skill; cognitive science and Gestalt theory suggest this takes deliberate action to do this. Complexity and paradox in art further reminds us each patient is a unique human being.



#### Fool Me Twice, Shame on Me: Prescription Errors Causing Adverse Events

M. J. Miller<sup>1</sup>, K. J. McCollum<sup>1</sup>, N. Hefzi<sup>2</sup>, K. Beydoun<sup>1</sup>, C. Denholm<sup>1</sup>, D. L. Levine<sup>2</sup> <sup>1</sup>Wayne State University, Detroit, MI

<sup>2</sup>Wayne State University, School of Medicine; Detroit Medical Center, Detroit, MI

Learning objectives: Recognize importance of documenting adverse drug events(ADE) and reporting ADE across settings (outpatient, hospital, etc.) and in all systems including community pharmacies. Encourage healthcare providers to advocate for electronic systems that automatically communicate with each other.

Case information: A 66-year-old African American man presented to the ED for severe angioedema of the lips and face. Review of medical records revealed the patient had a previous episode of angioedema one year prior. In this instance, the primary care physician (PCP) prescribed a combination pill which contained hydrochlorothiazide and lisinopril. Fortunately, with appropriate treatment, the patient's edema resolved without complications. The patient was informed of the error and measures were taken to prevent future events including calling his pharmacy to ensure documentation, contacting the PCP to ensure documentation in the outpatient electronic record, ordering an allergy-alert bracelet and discarding any medications containing ACE-Is.

**Discussion:** Up to 45.11% of adverse medication reactions requiring hospitalization are considered preventable, accounting for more than \$100 billion annually in healthcare expenditures. Key to this case was failure of the clinician and the EMR to detect the patient's allergen, lisinopril, when it was prescribed as a combination drug. ACE-I induced angioedema can be life-threatening requiring intubation and even cricothyroidectomy,3 It can be difficult to readily identify the components of combination drugs when ordered as trade names. Although there was a documented lisinopril allergy, Zestoretic (lisinopril/hydrochlorothiazide) was not recognized. Most healthcare facility EMRs alert physicians about drug allergies and interactions. Improving existing safeguards to better recognize combination drugs while reducing unnecessary alerts4 is critical to maintaining patient safety. 1.Pharmacoeconomics.(1999).15(5):445-458. 2.Eur J Clin Pharmacol.(2016).73(4):358-398. 3.Immunol Allergy Clin North Am. 2006;26(4):725-737. 4.JAAPA.(2016)28(7):41-2.

#### First-Hand Clinical Information Is More Important Than Referral Letters.

T. Miyagami<sup>1</sup>, T. Abe<sup>2</sup>

<sup>1</sup>Juntendo University Faculty of Medicine, Tokyo, Japan

<sup>2</sup>Juntendo University Urayasu Hospital, Urayasu, Japan

**Learning objectives:** Pyuria in a young man warrants careful workup for underlying structural cause. Overreliance on a referral letters can lead to minimization of patient history and complaints.

Case information: A 21 year old man with no significant medical history was referred to our hospital by an outside provider for three weeks of left back pain without fever. He had been hospitalized at another facility for the previous 3 weeks for severe back pain during which time he developed fever and pyuria. He was given 2 weeks of antibiotics. Though his pain continued, he became afebrile. On laboratory testing, serum albumin was noted to be 1.8mg/dl and CRP remained elavted, despite clinical improvement. He was discharged from the previous hospital and referred to our outpatient department for further investigation. He continued to complain of cloudy, foul-smelling urine at this time. Though his urine remained pyuric, we overlooked this; in the setting of low albumin and persistenly high CRP, plus a newly identified history of chronic diarrhea going back ten years, we pursued colonoscopy which identified active Crohn's disease. Three days later, however, he returned to our emergency department with recurrent fever and back pain. As urine revealed fecal contamination, vesicointestinal fistula was suspected. Ileal cystocele was subsequently identified on abdominal CT with oral contrast. Antibiotics was started and standby surgery was planned.

**Discussion:** Reexamination of the CT findings from the previous hospital identified possible vesicointestinal fistula given that there was air in the bladder. The patient's back pain was also likely due to Crohn's disease, as it resolved with 5ASA administration during his hospitalization. The referral letter we received from the previous hospital focused only on CRP and serum albumin, with neither information about pyuria nor a thorough clinical history. This reinforces the need to listen to patient's complaints without an overreliance on accompanying referral letters.

#### Take a Second Look: Diagnostic Reasoning and the Consequences of Anchoring

S. Grewal<sup>1</sup>, S. Bano<sup>1</sup>, C. Sittambalam<sup>1</sup>, <u>N. Muganlinskaya<sup>2</sup></u>

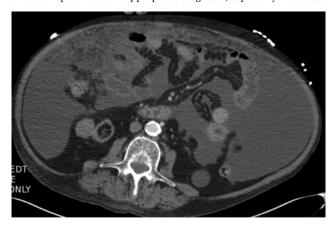
<sup>1</sup>MedStar Health Internal Medicine, Baltimore, Baltimore, MD

<sup>2</sup>MedStar Health Internal Medicine, Baltimore, MD

**Learning objectives:** Acceptance of a diagnosis before sufficient verification has occurred and failure to consider plausible alternatives are common causes of diagnostic error.

Case information: A 77-year-old man with history of failed renal transplant presented with weakness, nausea, and vomiting for 1 week. Of note, he also received weekly paracentesis for ascites that had been worked up extensively in the past and attributed to hypo-albuminemia and malnutrition. He was found to be hypotensive and tachycardic, with leukocytosis (26.3) and elevated ESR (88) and CRP (239). He revealed a significant weight loss (40 lbs) over the past year with physical exam notable for a cachexia with diminished breath sounds, irregularly irregular heart rhythm, trace dependent edema, moderately distended abdomen, and left heel stage 1-2 ulcers. He was empirically treated for presumed sepsis with vancomycin and piperacillin-tazobactam along with fluids. His abdominal exam prompted a CT of the abdomen which revealed omental infiltration. Cultures remained negative and further workup for infection returned negative. Ascitic fluid cultures remained negative and cytology was negative for malignant cells. Antibiotics were discontinued after 24 hours. Despite negative cytology, the case was reviewed again with radiology as it was felt that the infiltration seen could represent 'omental caking' due to malignancy. An omental biopsy was done with immuno-histochemical staining consistent with malignant peritoneal mesothelioma (MPM).

**Discussion:** MPM is a rare (400 cases/year), but rapidly fatal condition (<1 yr) and usually presents with nonspecific symptoms. Ascites is present in 60-80% of newly diagnosed patients. Careful attention to history, physical exam findings, and review of radiographical findings prompted further workup for malignancy despite the negative findings. This case highlights the importance of reflective reasoning and avoiding anchoring to pursue more definitive workup to make the appropriate diagnosis, especially in the setting of recurrent symptoms.



#### A Normal Finding Is Not Always Normal: A Teachable Moment for Diagnostic Error.

T. Nishinobu, Y. Harada, T. Shimizu Dokkyo Medical University Hospital, Shimotsuga, Japan

Learning objectives: To recognize the lower sensitivity of brain magnetic resonance imaging for early cerebellar infarction.

Case information: A 55-year-old man visited the emergency department with a history of vertigo, nausea, and vomiting for four hours before presentation. His past medical history included cerebral infarction, diabetes mellitus, and chronic kidney disease on hemodialysis. Physical examination revealed vertical nystagmus but otherwise normal. Electrocardiogram, head computed tomography (CT) without contrast, and brain magnetic resonance imaging (MRI) findings were interpreted as normal by a radiologist. While his vertigo improved, the patient was admitted due to severe nausea. The patient was discharged with complete recovery on the next day; however, he presented to the emergency department several hours later as his vertigo and vomiting recurred. Repeated head CT detected a low-density area in the right cerebellar hemisphere. Another diffusion-weighted MRI was performed, showing a high intensity lesion on the identical area. The MRI also revealed another lesion in the right cerebellar vermis, which was overlooked on the previous MRI. Diagnosis of acute right cerebellar infarction was made, and the patient was admitted again.

Discussion: A diagnostic delay occurred in this case. Diagnostic error in this case seemed to be avoidable even with the misinterpretation of the first MRI findings. Based on the presence of vertical nystagmus, which was a characteristic finding for central vertigo, and the patient's high-risk profile for stroke such as diabetes, chronic kidney disease, and prior stroke, the probability of new stroke should have been estimated to be very high. Therefore, even if the MRI findings was documented as "normal," stroke still should be on the differential list. This case underscores the importance to know the limitation of each diagnostic test. Clinicians should precisely utilize each clinical tool for decisionmaking, thereby reducing diagnostic error rate in case false-negative results occur.

#### The True Diagnosis of Nonspecific Symptoms Mimicking Aging Process

R. Ono1, I. Kitagawa2 <sup>1</sup>Shonan Kamakura General Hospital, Kanagawa, Japan <sup>2</sup>Shonan Kamakura General Hospital, Kamakura, Japan

Learning objectives: It is important to recognize anterior pituitary deficiency and malignancy sometimes come with only nonspecific symptoms.

Case information: A 67-year-old man with the past medical history of hypertension and diabetes mellitus presented with a history of generalized weakness, lethargy, weight loss and loss of appetite. The patient had been a current smoker with a half a pack a day for 47 years. One and a half months before, he had noticed lethargy and anorexia with 3kg weight loss during only one month and visited our hospital, but he was diagnosed with diabetes mellitus and aging process. On admission day, he could not move because of the generalized weakness and lethargy. Laboratory data showed hypoglycemia, low thyroid stimulating hormone and free thyroxine. We suspected hypopituitalism and performed head imaging modalities revealing multiple tumors, in one of which was on suprasellar region. A chest computed tomography showed a tumor shadow and the bronchoscopic biopsy pathologically showed small cell lung carcinoma. Hormone profile demonstrated hypothalamic pan-hypopituitarism. We diagnosed hypothalamic hypopituitarism secondary to metastases from the primary lung carcinoma and initiated radiation, chemotherapy and hormone replacement, but he died 10 months later.

**Discussion:** Metastasis to the pituitary gland is a rare condition revealed in about 5% of patients with cancers, although approximate two thirds of them are latent. Recognition of hypopituitarism can be challenging, especially in the elderly in whom symptoms such as lethargy and visual decline may be mistaken for the aging process. In patients with established metastatic disease, symptoms may be wrongly attributed to malignancy or to the side effects of therapy. Systemic complications of malignancy, including nonspecific symptoms (weakness, lethargy and weight loss) and central nervous system involvement, may also mask anterior pituitary deficiency. Physicians need to recognize anterior pituitary deficiency and malignancy sometimes come with only nonspecific symptoms.

#### Missing the Forest for the Trees

Z. Qamar<sup>1</sup>, J. Akhtar<sup>2</sup> <sup>1</sup>Dow University of Health Sciences, Karachi, Pakistan <sup>2</sup>UPMC, Pittsburgh, PA

Learning objectives: Multisystem diseases require a comprehensive evaluation to establish a correct diagnosis. The cohesively abnormal 'forest' may be obscured by confounding 'trees'. It is important physicians pay attention to details.

Case information: A 22-year-old man with no significant past medical history presented with a chief complaint of rash of his legs bilaterally and leg pain. He denied fevers or chills with this. He had a recent upper respiratory infection, which included a cough and some nasal congestion. He stated he has had approximately 2 months of nonproductive cough. He denied any night sweats. Examination of the lower extremities showed scattered purpura which are very tender and raised. He also has multiple petechiae, and a moderate amount of swelling of the ankles bilaterally. Dermatology was consulted and thought it was likely viral or autoimmune in nature. Skin biopsy shower superficial perivascular neutrophil-rich dermatitis suggestive of an early phase of leucocytoclastic (aka hypersensitivity) vasculitis. The initial complete blood count showed leukocytosis of 22.1 with monocytic predominance. This was initially ignored as the treating physician was focussed on the rash, and satisfied once the diagnosis of vasculitis was established. Further evaluation with a bone marrow biopsy confirmed AML with monocytic differentiation.

Discussion: Failure to elicit all relevant information (unpacking) in establishing a differential diagnosis might result in significant possibilities being missed. If patients are allowed to limit their history giving, or physicians otherwise limit their medical decision process, diagnosis might be missed or delayed.

#### Missing a Beat: The Delayed Diagnosis of the Need for Palliative Care in a Patient with End-Stage Heart Failure

C. Schifeling<sup>1</sup>, J. N. Lessing<sup>2</sup> <sup>1</sup>University of Colorado School of Medicine, Aurora, CO <sup>2</sup>University of Colorado, Denver, CO

Learning objectives: Recognize, in timely fashion, the need for and benefit from palliative care in patients with non-malignant illnesses. Case information: A 75 year old woman with atrial fibrillation and ischemic congestive heart failure for two years, four hospitalizations in the past three months, and recent left ventricular ejection fraction of 10-15% was transferred to an academic medical center many hours away from her home for placement of a cardiac resynchronization defibrillator for decompensated heart failure refractory to intravenous diuretics and inotropes. Fifteen days, a right heart catheterization, two thoracenteses, two blood transfusions, four nights of agitated delirium and an infection with C. difficile later, palliative care met with the family who opted to pursue home hospice. She died at home surrounded by friends and family five days after discharge.

**Discussion:** While the clinically meaningful quality of life benefits of palliative care extend to at least three months (1), the median time from palliative care consultation to death is only 23 days (2). This discrepancy is greater for patients with non-cancer terminal illnesses, in part because of poorer prognostication measures in this population (3). As such, there is an imperative to better describe the diagnosis of the need for palliative care, particularly in patients with non-malignant disease. 1. Kavalieratos D, Corbelli J, Zhang D, et al. Association Between Palliative Care and Patient and Caregiver Outcomes: a Systematic Review and Meta-analysis. JAMA. 2016 Nov;316:2104-114. 2. Kozlov E, Carpenter BD, Thorsten M, et al. Timing of Palliative Care Consultations and Recommendations: Understanding the Variability. Am J Hosp Palliat Care. 2015;32:772-5. 3. Coventry PA, Grande GE, Richards DA, et al. Prediction of Appropriate Timing of Palliative Care for Older Adults with Non-Malignant Life-Threatening Disease: a Systematic Review. Age Aging. 2005:218-27.

## Persistent Fever in a Medically Complex Child

J. Seserinac, J. Lau The University of Arizona, Tucson, AZ

Learning objectives: Identify difficulties in diagnosing etiology of fever in a medically complex child

Case information: A 16-year-old medically complex girl presented with fever and abdominal discomfort. Her past medical history was significant for spastic quadriplegic cerebral palsy, developmental delay, nonverbal, VP shunt placement, spinal fusion, constipation, urinary tract infections. She was non-toxic and uncomfortable with presumed abdominal pain. Initial CBC showed elevated WBC, urinalysis minimal pyuria, abdominal x-ray and VP shunt series were essentially normal. The urine culture specimen was obtained prior to starting IV Ceftriaxone, however was incorrectly ordered in the computer and the specimen was rejected and discarded by the laboratory. The patient was admitted to the pediatric hospitalist service for presumed acute pyelonephritis treatment. CT abdomen showed small amount of fluid in pelvis and was otherwise noncontributory. Neurosurgical team was consulted, since the VP revision was performed recently, and recommended against CSF evaluation. The patient continued to have high fever with increase in discomfort during abdominal exam. A repeat CT scan of the abdomen showed abscess formation around the tip of VP shunt while the head CT scan showed worsening hydrocephalus. At that time the neurosurgical team agreed to tap the VP shunt. CSF showed pleocytosis and culture was positive for Pseudomonas Aeruginosa.

Discussion: The primary team faced multiple difficulties while attempting to reach the correct diagnosis. The child recently moved to the area, had no medical home, was not able to communicate effectively, physical exam was limited, urine culture specimen was lost (thus, impossible to effectively evaluate presence of acute pyelonephritis), the consultant neurosurgery team advised against CSF collection, and the patient's spasticity combined with spinal fusion surgery made it impossible to obtain CSF specimen via lumbar puncture. The review of literature confirms that the presence of persistent fever and recent VP shunt revision are strong predictors of shunt infection.

#### Diagnostic Errors in Neonatal Intensive Care: A Case Series

G. J. D. C. Shafer<sup>1</sup>, G. K. Suresh<sup>2</sup>

<sup>1</sup>Baylor College of Medicine & Texas Children's Hospital, Houston, TX

<sup>2</sup>Baylor College of Medicine & Texas Children's Hospital

Learning objectives: Describe prototypical diagnostic errors in Neonatal Intensive Care Unit (NICU) patients, Identify components of the diagnostic process that lead to such errors. Classify the etiology of NICU diagnostic errors as cognitive, systems-based or no-fault.

Case information: We present ten cases of diagnostic error in the NICU at our institution. We reviewed each case in detail to identify which step/s in the diagnostic process broke down and classified the overall etiology of the diagnostic error as cognitive, systems-based or no-fault. We include a category of 'no-fault' diagnostic error – defined as a highly atypical presentation – to illustrate a diagnosis which would remain indiscernible despite a reasonable diagnostic evaluation. None of the patients died as a direct result of the diagnostic error.

Discussion: NICU patients are at high risk to experience diagnostic errors that result in significant short-term and long-term health consequences. However, there exist only a few studies on NICU diagnostic errors, and these are all autopsy-based evaluations, which did not assess diagnostic errors when the outcome was not death. The cases presented in this series constitute the first description of non-lethal diagnostic errors in the NICU, and highlight the need for further research in this field.

Case (n=10)	Initial Presentation	Missed Diagnosis	Missed Component of the Diagnostic Process	Etiology of Diagnostic Error
1	Preterm delivery at 31 weeks	HSV keratitis	Missed physical exam findings	Cognitive
2	Term delivery after low-risk pregnancy	Imperforate anus	Missed physical exam findings	Cognitive
3	Mistaken concern for hypospadias during routine circumcision	Incorrectly diagnosed hypospadias of anatomically- normal male genitalia	Missed physical exam findings	Cognitive
4	Preterm delivery at 34 weeks after maternal PPROM	NEC and Escherichia coli bacteremia	Incorrectly interpreted radiographic and missed laboratory findings	Cognitive
5	Preterm delivery at 24 weeks with abdominal distension	Pneumothorax	Incorrectly interpreted radiographic finding	Cognitive and systems- based
6	Term delivery with desaturations	13 ribs on chest radiograph	Incorrectly interpreted radiographic finding	Cognitive
7	CDH with hematuria	Renal calculus	Incorrectly interpreted radiographic finding	Cognitive
8	CDH with ventilator dependence	PAPVR	Incorrectly interpreted echocardiographic finding	Cognitive
9	Preterm delivery at 31 weeks	TEF	Inadequate team communication	Systems- based
10	Persistent hyperglycemia in a critically-ill neonate with multiple congenital abnormalities	Neonatal diabetes related to chromosome 6q24.1q24.3 triplication	No-fault	No-fault

#### Table Abbreviations

HSV: herpes simplex virus

PPROM: prolonged premature rupture of the membranes

NEC: necrotizing enterocolitis

CDH: congenital diaphragmatic hernia

PAPVR: partial anomalous pulmonary venous return

#### An Adolescent with Acute Anaphylaxis Secondary to Clostridium Difficile

D. R. Siegel, D. D. Hanba Wayne State University, Detroit, MI

Learning objectives: Recognize the extra-colonic manifestations of Clostridium difficile infection so that it can be diagnosed and treated promptly and appropriately.

Case information: A 14-year-old female presented to the emergency department with an acute onset of diffuse urticarial rash and angioedema. The patient had accompanying hypotension and diarrhea but no airway comprise. The patient and her mother denied having any history of allergies, history of an insect bite or animal contact, current medications, or any changes to her diet or environment. The patient was promptly treated for anaphylaxis with stabilization of her vital signs but persistence of urticaria and angioedema for over 24 hours. Initial infectious serology work-up was unremarkable. Upon further questioning, the patient admitted to a recent course of clindamycin approximately three weeks prior to presentation. Clostridium difficile toxin PCR was ordered and found to be positive. The patient was immediately started on metronidazole along with prednisone, hydrocortisone cream and hydroxyzine and her symptoms resolved over the next few days. **Discussion:** This patient presented with acute anaphylaxis likely secondary to *Clostridium difficile* infection. *C difficile* is a gram-positive, spore-forming anaerobic bacillus. Symptoms usually begin during or shortly after antibiotic therapy, but presentation may be delayed for several weeks. C difficile infection can range from asymptomatic carriers to mild or severe diarrhea. Although the typical clinical presentation of C difficile infection is limited to the colonic symptoms, more recently C difficile infection has been implicated to cause disease in a variety of other organ systems. Despite this unusual presentation of C difficile infection, clinicians must always maintain a high index of suspicion for this highly transmissible infection. Early diagnosis allows for proper contact precautions to be initiated, limiting the spread of this hospitalacquired infection while potentially shortening the length of hospital stay and decreasing the utilization of resources.

#### Failing to Consider a Fall: Anchoring on Recent Procedures

M. G. Simonson

University of Pittsburgh Medical Center, Pittsburgh, PA

Learning objectives: Understand how anchoring bias and the Semmelweis reflex contribute to diagnostic error. Recognize that patients with heavy alcohol and tobacco use are at increased risk for acute fragility fractures.

Case information: A 50-year-old female with a history of heavy tobacco and alcohol use was admitted with refractory seizures. Work up included a CT head without acute pathology, lumbar puncture with unremarkable CSF studies, and unremarkable serum chemistries except for sodium 106mEq/L. The patient received hypertonic saline with rapid clinical improvement. During the hospitalization, she developed new low back pain without fecal or urinary incontinence. Severe pain was elicited with palpation over the lumbosacral spine. Strength and sensation remained intact. These findings were attributed to post-lumbar puncture discomfort and imaging was deferred. Due to persistent pain, lumbar MRI was obtained to rule out spinal hematoma. The study demonstrated incidental finding of acute sacral insufficiency fracture. Upon further questioning, the patient admitted she had fallen to her knees while transitioning to bedside commode without nursing assistance. The patient was taken for surgery for trans-iliac trans-sacral screw fixation.

**Discussion:** Spinal hematoma is a known but rare complication of lumbar puncture occurring in <2% of patients, primarily in patients with thrombocytopenia or coagulopathy. In this patient, a history of smoking, heavy alcohol use, and hyponatremia all increase her risk for osteoporosis and subsequent fracture. By anchoring on the possible complications of a recent procedure, the medical team failed to consider the possibility of a fall or fragility fracture. Furthermore, the timing and severity of pain in this case are inconsistent with the considered procedural complications. Rejection of information that contradicts a favored diagnosis is an example of the Semmelweis reflex. Recognition of how these cognitive biases can manifest in patients following invasive procedures may help prevent diagnostic delays for future patients.

#### Cystocerebral Syndrome: A Confusing Presentation of Encephalopathy.

H. Tabaja, A. Kanj, D. L. Levine

Wayne State University, School of Medicine; Detroit Medical Center, Detroit, MI

Learning objectives: Avoid representativeness heuristics in elderly patients presenting with acute encephalopathy

Case information: An 89-year-old woman with ischemic cardiomyopathy was brought to the emergency department by her son for worsening shortness of breath and orthopnea, typical of her previous heart failure exacerbations. The patient has advanced dementia and depression, for which she is on donepezil, memantine and quetiapine. She was afebrile and hemodynamically stable. She was obtunded, but minimally arousable to painful stimuli. Her exam was otherwise consistent with fluid overload. Her son claimed that despite her dementia she is usually more alert and that this change in mentation only started few hours ago. The patient had normal electrolytes, TSH, BUN, and ammonia levels.

No infiltrates were seen on chest X-ray. A head computed tomography (CT) was ordered. On the medical ward, she had urinary incontinence and suprapubic tenderness and fullness. A Foley catheter was inserted and 1 liter of urine was drained. Within seconds of catheterization, the patient became fully alert and went back to her baseline mentation. Her urinalysis was sterile ruling out urinary tract infection. Further workup for encephalopathy, including head CT, was discontinued. Her retention was attributed to quetiapine which was held. Her retention did not recur, and she continued to be alert.

Discussion: Cystocerebral syndrome describes delirium secondary to urinary retention that immediately responds to bladder decompression. It is most commonly encountered in the geriatric population, particularly in patients with dementia. It is possibly related to the increased sympathetic tone and catecholamine surge induced by bladder wall tension. Although infections remain the most common cause of encephalopathy in this population, physicians, especially at the emergency department, should avoid representativeness heuristics and rule out urinary retention in patients with advanced dementia presenting with encephalopathy. This only requires a simple bedside examination and helps prevent unnecessary, costly, and sophisticated workup.

#### Well Controlled HIV Patient with Kaposi's Sarcoma

H. Takahashi, T. Naito Juntendo University Hospital, Tokyo, Japan

Learning objectives: After attending this lecture the participant will demonstrate the ability to recognize the importance of small illogical medical story to reach true diagnosis. If your logical story and the result of test were contradiction, you have to look for another story to reach true diagnosis.

Case information: A 49-year-old man with a history of 9-year history of HIV infection and a medical history of long-standing asthma treated with inhaled corticosteroids. His HIV treatment was well controlled by Raltegravir and Darunavir ethanolate/Cobicistat. He had a medical history of compression fracture 2 years before admission. He admitted our hospital because Kaposi's sarcoma and edema has been seen during HIV treatment. His CD4 counts were approximately over 700/µl until 1 year before admission. However CD4 count was 318 /µl and HIV-RNA was undetectable when he admitted. GCS:15, BT:37.3, HR:84 /min, BP:126/84 mmHg, RR:12 /min, SpO2:98%. There are several Kaposi's sarcoma like tumor at his legs and 6 compression fractures at lumber spine. His blood ACTH level and cortisol level were decreased. Short ACTH, TRH, CRH and LHRH stimulation test were all negative. We diagnosed him as pseudo adrenal insufficiency based on exogenous steroid. After stopping his inhaled corticosteroid his cortisol level and CD4 count were improved.

Discussion: Kaposi's sarcoma is well known as an one of AIDS related diseases. Although HIV patient with Kaposi's sarcoma is reasonable, it is not in well controlled HIV patient. Cushing's disease also cause Kaposi's sarcoma and osteoporosis at the same time. Adrenal insufficiency was demonstrated in patient's blood test despite medical condition indicated steroid excessive status. He has never taken exogenous steroids, and stopping the causative inhaled steroid improved the cortisol in the blood. Steroid excessive condition by inhaled steroid in HIV patient has been reported 3 cases and the cause were drug interaction between inhaled steroid and antiretroviral drug.

# Geography & Sick Contacts: How a Tiered Approach to the Diagnosis of Meningitis Can Save Time, Money and **Unnecessary Treatment**

B. E. Ueberroth

Wayne State University School of Medicine, Detroit, MI

Learning objectives: 1. Knowledge of endemic organisms in infection of unknown origin can provide focused diagnostic recommendations and limit unnecessary testing

Case information: A 32 year-old woman presented to the emergency department with headache, neck stiffness, nausea and photophobia for two days. She returned from vacation in Cancun, Mexico two weeks prior. Ten days ago, her son was admitted to the local children's hospital and was treated for viral meningitis. Neurology was consulted. In addition to the CSF profile already pending from lumbar puncture, they recommended a wide panel of infectious CSF workup, including enterovirus, HSV, Borrelia, VDRL, Zika virus, West Nile virus, cryptococcus, coccidioides and mycobacterium. The patient's CSF tested positive for enterovirus, all IV treatments were discontinued and she was discharged in stable condition.

Discussion: Infectious disease was consulted after neurology and reviewed the case, ultimately recommending against many of the tests. While coccidioides is endemic to Mexico, it is not endemic to Cancun, and therefore would be highly unlikely. The patient had no recent travel to an endemic borrelia area. Zika and West Nile viruses were also highly unlikely given the endemic distribution and low prevalence. When there is question of what organism is at play, early involvement of infectious disease experts can help guide diagnostic tests and limit unnecessary assays that are unlikely to yield a causative organism based on geography or other exposure. Utilizing a likelihood-based approach derived from the patient's history, one can prioritize higher yield tests and reach a diagnosis with fewer tests, shorter delay and lower cost.

#### 18 Intubations and 1 Missed Diagnosis

A. Wadehra<sup>1</sup>, C. Kraus<sup>2</sup>, D. L. Levine<sup>3</sup>, A. R. Peralta<sup>2</sup>

<sup>1</sup>Wayne State University/Detroit Medical Center, Detroit, MI

<sup>2</sup>Henry Ford Hospital, Detroit, MI

<sup>3</sup>Wayne State University, School of Medicine; Detroit Medical Center, Detroit, MI

Learning objectives: 1) Understand the importance of maintaining a broad differential diagnosis in patients with shortness of breath, 2) Explain the impact of not properly documenting medical histories in the electronic health record.

Case information: A 50 year old woman with asthma, anxiety, bipolar disorder, and 18 previous intubations presented with shortness of breath, non-productive cough, and conversational dyspnea. On exam, she was using accessory muscles and tachypneic. Expiratory phase was prolonged and wheezing was noted. Cardiac exam, chest x-ray and EKG were unremarkable. B-natriuretic peptide and troponins were not elevated. Patient was placed on BiPAP, but eventually required intubation. She was transferred to the ICU with a diagnosis of acute exacerbation of asthma and treated accordingly. Peak inspiratory pressures on the ventilator were normal. A detailed chart review showed a laryngoscopy, performed 7 months earlier, was significant for Paradoxical Vocal Fold Motion. The diagnosis was not added to the patient's history in the electronic health record and was not noted on any of the admissions that occurred post-procedure.

**Discussion:** Paradoxical Vocal Fold Motion (PVFM) is a condition in which there is inappropriate adduction of the true vocal folds during inspiration. Patients present with shortness of breath, cough, and dysphonia. They are often misdiagnosed and therefore receive suboptimal or inappropriate treatment. This case illustrates the importance of having a broad differential diagnosis and importance of documenting medical histories appropriately. PVFM is a cause of obstructive airway disease and can be mistaken for asthma. The level of obstruction, however, is extra-thoracic and it is treated acutely with breathing maneuvers. An association with psychiatric disorders is observed, but PVFM is also misdiagnosed as anxiety. It is important to have a high index of suspicion to ensure appropriate care, prevent unnecessary intubation, and potentially, unnecessary admissions.

#### No Doubt!! Certainly Pneumothorax!! Diagnostic Error By Intuitive Visual Diagnosis.

T. Watari<sup>1</sup>, K. Kishimoto<sup>2</sup>, S. Okazaki<sup>3</sup>, K. Onigata<sup>4</sup>

<sup>1</sup>Shimane University Hospital, Shimane, Japan

<sup>2</sup>Shimane University Hospital, Izumo-shi, Japan

<sup>3</sup>Shimane University Hospital, Izumo, Japan

<sup>4</sup>Shimane University Hospital, Shimane prefecture, Japan

Learning objectives: Skinfold shadow is a frequent artifact on the anterior-posterior portable chest radiographs of patients in the critical care setting Pneumothorax and skinfold might be difficult to differentiate upon chest radiography. In those cases, lung ultrasonography or CT facilitates are helpful to make a correct diagnosis

Case information: A 79-year-old hospitalized woman underwent central venous catheterization for high-calorie infusion. Chest radiography to check the position of the catheter showed a curvilinear line at the lateral zone of the right lung (Figure 1A). We suspected pneumothorax and performed computed tomography (CT) of the chest. However, it revealed the absence of pneumothorax; the curvilinear mimicked pneumothorax was revealed to be a skinfold (Figure 1B). The lateral margin of the skinfold artifact, caused by compression of flaccid back skin against the film, can be sharply outlined by a lucent line. This line is an edge enhancement phenomenon at the margin between objects with different contrast or luminance (Mach band effect). The area lateral to the skinfold artifact may be perceived darker than the hilar side because of the added density of the skinfold. The lung marking sign is useful for differentiate pneumothorax from the other conditions, but in the emphysema and bullous disease the lung marking can be vanished. And in the case with hubbly collapse of pneumothorax, the lung marking can be seen beyond the outside of the collapsed lung.

Discussion: In some situations, pneumothorax and skinfold might be difficult to differentiate upon chest radiography. In those cases, lung ultrasonography or CT facilitates are helpful to make a correct diagnosis.

#### Poor Follow-up Leads to a Delay in Diagnosis of Malignant Otitis Externa

R. Wojcik<sup>1</sup>, D. DiNardo<sup>2</sup>

<sup>1</sup>University of Pittsburgh Medical Center, Pittsburgh, PA

<sup>2</sup>University of Pittsburgh School of Medicine, Pittsburgh, PA

#### Learning objectives:

- Identify the role for clinical follow-up as a means for feedback regarding working diagnoses in the outpatient setting.
- Define patient characteristics that should prompt consideration of malignant otitis externa (MOE) as a possible diagnosis.

Case information: 90 year-old woman with well-controlled diabetes, hypertension, and history of seizure disorder presented to the ER after her family noticed a left-sided facial droop. Imaging demonstrated widespread malignant otitis externa with a cranial nerve VII palsy. Notably, she had called her primary care clinic reporting an "ear infection" approximately five months prior. Her symptoms at that time included bloody nasal discharge and ear pain with associated severe headaches. She was given a short course of antibiotics over the phone without clinical evaluation and with no specific plan for follow-up.

Discussion: Diagnosis of MOE is often delayed (mean time to diagnosis of 6.9 months), contributing to mortality rates, up to 50%. Early consideration in patients at risk is key for appropriate diagnosis and treatment. The diagnosis should be considered in elderly patients with otalgia who also have diabetes, regardless of level of insulin resistance, and in patients who are immunosuppressed for any reason. In this case, MOE was not initially considered. Anchoring on the patient's chief compliant and failure to seek additional information through history and physical limited the differential diagnosis. Furthermore, the lack of follow-up after treatment for presumed ear infection likely contributed to the delayed diagnosis of MOE. Follow-up after the establishment of a diagnosis and management is essential for assessment of response to treatment, and for the consideration of alternative diagnoses. In the outpatient setting, scheduled follow-up can allow clinicians to receive otherwise unobtainable feedback regarding both successful and unsuccessful diagnoses, and to calibrate diagnostic reasoning going forward.