Congress Abstracts

The Diagnostic Error in Medicine 13th Annual International Conference

October 19–21, 2020

Planning Committee
Andrew Olson, MD | Chair
Ashley Meyer, PhD | Co-chair
Gopi Astik, MD, FHM
Daniel Berg
Jeff Brady, MD, MPH
Christina Cifra, MD, MS
Robert El-Kareh, MD, MPH, MS
Amit Ghosh, MD
Kelly Gleason, PhD, RN
Harry Hoar III, MD
Abraham Jacob, MS, MHA
Becky Jones, MBA, BSN, RN, CPHRM
Michael Kanter, MD, CPPS
John Matulis, DO, MPH
David Meyers, MD, FACEP
Ruth Ryan, RN, BSN, MSW, CPHRM
Kathryn Schaefer, MSN, RN, CPHRM
Grant Shafer, MD, MA, FAAP

Abstract Selection Committee
Christina Cifra, MD, MS
Janice Kwan, MD, MPH
Benji Mathews, MD
Ashley Meyer, PhD
Assessment of Medical Student Diagnostic Decision-Making Performance in Electronic Health Record (EHR) Based Simulations

C. Clancy¹, I. Rasooly², N. Muthu¹
¹Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA
²Children’s Hospital of Philadelphia, Philadelphia, PA

Purpose/Problem: Electronic health record (EHR) based simulation provides an opportunity to teach and evaluate clinical reasoning and decision-making skills in a high-fidelity environment. We developed an EHR-simulation session for pediatric residents aimed at introducing diagnostic decision-making tools to overcome cognitive bias; participation was associated with persistent changes to EHR use patterns in clinical practice. Using the Situation Awareness Global Assessment Technique (SAGAT), we developed an instrument to stratify simulation performance across three domains: perception of relevant information, comprehension of information, and projection of potential outcomes. In this study, we conducted the session in a new population, clerkship medical students.

Methods: The simulation and assessment were administered to clerkship medical students at the Perelman School of Medicine, as part of a required online course. Students were directed to assume the role of an admitting resident and were given 15 minutes to review a simulated EHR chart of an infant with physiologic hyperbilirubinemia and complete a simplified SAGAT. While documentation suggests the patient is stable, embedded safety probes, including unstable vital signs and labs, suggest neonatal sepsis. The session concluded with a virtual debriefing session to review the case and emphasize the utility of data visualization and information retrieval tools in the EHR.

Outcomes: Five cohorts of clerkship medical students (n=106) participated in the simulation. The vast majority (96.2%) correctly perceived the heart rate. In the comprehension domain, 38.1% of students included abnormal vital signs and/or sepsis in a problem representation of the case. Responses on the projection item were mixed- 58.5% of students concluded that the patient was stable for transfer to the floor, while the remainder thought more discussion was needed (34%) or that the patient was not stable for transfer (7.5%).

Discussion: We implemented an EHR based simulation and decision-making assessment in clerkship medical students. Initial results indicate that medical students perform similarly to pediatrics residents in the perception and comprehension domains of the instrument. More detailed analysis is required to confirm and explore this trend. This study demonstrates that this simulation and assessment tool can be applied to different levels of learner.

Significance: EHR based simulation provides an opportunity to evaluate and address deficits in diagnostic decision-making. A simulation-based strategy for assessment of clinical decision-making in medical students can be used to enhance clinical reasoning and diagnostic error education.

The Diagnostic Statement: A Linguistic Analysis of How Clinicians Communicate Diagnosis

M. Dahm¹, C. Crock²
¹Australian National University, Canberra, AU
²Eye & Ear Hospital, Melbourne, AU

Background: The patient-centered definition of diagnostic error encompasses a failure to establish a timely explanation for a patient’s problem and failure to communicate that explanation to the patient. However, little is known from a linguistic perspective about how diagnostic statements are delivered to patients, i.e. how clinicians name, describe or explain the health problem to patients, and how these statements relate to diagnostic accuracy.
Methods: To identify temporal and discursive features in diagnostic statements, we analysed transcripts from 16 video-recorded interactions. Interactions were part of an objective structured clinical examination (OSCE) station conducted as part of a practice high-stakes exam for internationally trained clinicians (25% female, n=4) to gain accreditation to practice in Australia. The 8-minute OSCE revolved around a 3yo child with trouble settling at night, increased irritability and pulling at his ear. The ear examination depicted a bright blue object in the child’s ear canal. We analysed time spent on history-taking, examination and delivery of diagnostic statements. We then extracted and deductively analysed types of diagnostic statements informed by literature, ranging from a) plain assertion, to b) epistemic modality (modal verbs and personal judgements), c) evidentialised (alluding to diagnostic process or explicitly describing observations) and d) epidemiological generalisations.

Results: Fifty percent of participants communicated the accurate diagnosis ‘foreign body’ (FB) (n=8). Seven clinicians incorrectly diagnosed Otitis Media (OM) after ear examination and one diagnosed behavioural issues without ear examination. Total duration for history taking ranged from 85 to 210 seconds (s), and for diagnostic statements from 10 to 105s. On average, clinicians who made a diagnostic error took 30s less in history taking (140s, SD: 44s) and 30s more in providing diagnosis (68s, SD: 25s) than clinicians who accurately diagnosed FB (History: 170s, SD: 36s; Diagnosis: 37s, SD: 22s). The majority of diagnostic statements were evidentialised (describing specific observations (n=23) or alluding to diagnostic processes (n=7)), followed by epistemic modality (n=8), generalisations (n=6) and assertions (n=4). Overall, clinicians who misdiagnosed provided more specific observations (n=14) than those who diagnosed correctly (n=9).

Conclusion: Evidentialised diagnostic statements provide evidence to support a diagnosis and make diagnostic reasoning explicit to the patient. Our preliminary results suggest that in interactions where there is a diagnostic error, clinicians made longer diagnostic statements that featured more evidence. It appears clinicians might use more evidence statements to support uncertain diagnoses. Larger future studies are needed to explore links between evidentialised diagnostic statements and diagnostic uncertainty.

Bedside Clinicians’ Perceptions on the Contributing Role of Diagnostic Errors in Critically Ill Patient Presentation

C. Huang1, A. Barwise1, J. Soleimani1, Y. Dong1, S. Herasevich1, S. Anjum Khan2, S. Helgeson3, P. Moreno-Franco3, Y. Pinevich1, R. Kashyap1, V. Herasevich1, O. Gajic1, B. Pickering1

1Mayo Clinic, Rochester, MN
2Mayo Clinic Health System, Mankato
3Mayo Clinic, Jacksonville

Background: Diagnostic errors in critical illness remain poorly understood in patient safety research, which limits our understanding and ability to design future interventions aimed at reducing diagnostic errors in the acute care setting. The primary objectives of this study were to explore clinicians’ perceptions of the occurrence and factors associated with diagnostic errors in patients evaluated during a rapid response team (RRT) activation or unplanned admission to the intensive care unit (ICU).

Methods: A multicenter prospective survey study was conducted among multi-professional clinicians involved in the care of patients with RRT activations and/or unplanned ICU admissions at two academic medical centers and one community-based hospital between April 2019 and March 2020. A study investigator screened eligible patients every day. Within 24 hours of the event, a research coordinator administered the survey to clinicians, who were asked: whether diagnostic errors contributed to the reason for evaluation; whether any new diagnosis was made following evaluation; if there were any failures to communicate the diagnosis; and if involvement of specialists earlier would have benefited that patient. Patient clinical data were extracted from the electronic health record. All data were collected in REDCap and analyzed using JMP Pro 14.1.0 software.

Results: A total of 1815 patients experienced RRT activations and 1024 patients experienced unplanned ICU admissions, with 798 and 440 respectively deemed eligible. Surveys were sent out to clinicians in 963 patient care episodes in total. We received at least one survey on 522 patients (54.2%). Among completed surveys, clinicians reported that 18.2% (95/522) of patients experienced diagnostic errors; 8.0% (42/522) experienced a failure of communication; 16.7% (87/522) may have benefited from earlier involvement of specialists. Compared to academic settings, clinicians in the community hospital
were less likely to report diagnostic errors (7.0% vs 22.8%, P=0.002). When the analysis was restricted to perceptions from attending physicians only, the median SOFA score was significantly higher in patients considered to have errors compared to those who were not (7 vs 4, P=0.004). Clinicians reported that more patients with errors experienced failure of communication and failure to involve a specialist earlier (26.9% vs 5.2%, P<0.001 and 53.9% vs 5.1%, P<0.001, respectively), compared to patients without errors.

Conclusions: Critical care clinicians report a high rate of diagnostic errors in patients they evaluate during RRT activations or unplanned ICU admissions. Efforts that are likely to reduce diagnostic errors may benefit from focusing on improving communication and multidisciplinary team cooperation.

Developing a Data Stream to Detect Diagnostic Errors in the Pediatric ED

D. Lam1, F. Dominguez1, S. Deakyn-Davies1, L. Bajaj1, E. Nathaniel1, C. Barrett1, J. Grubenhoff1*

* (Presenting Author)

1 University of Colorado School of Medicine, Aurora, CO

Statement of Problem: The incidence and impact of diagnostic errors (Dxes) among children evaluated in emergency departments (EDs) and urgent cares (UCs) has not been described. Dxes arising during ED/UC visits rarely come to the attention of individual providers or hospital systems unless the error results in significant harm or medico-legal action. Defining the incidence of Dxes in the ED/UC will improve the ability to identify specific and recurrent vulnerabilities in the diagnostic process amenable to systematic improvement in diagnostic performance.

Description of the Intervention: Using the framework described by H. Singh and colleagues, we developed an e-trigger process to identify patients at high risk of Dxes evaluated in a large pediatric hospital system (5 satellite ED/UCs; 1 regional pediatric trauma center ED) with ~162,000 combined annual visits and ~12,000 admissions. Attention focused on episodes that included evaluation and discharge from an index ED/UC visit with an unplanned admission within the subsequent 14 days. A clinical research informaticist developed an algorithm to identify candidate encounters and extract demographic information and up to three diagnoses from the ED encounter and subsequent discharge summary. A physician or nurse screened candidate encounters for those in which the index visit diagnosis differed from the hospital discharge diagnosis (i.e. detection criteria for possible Dx). Encounters were then reviewed by a physician using the Revised SaferDx instrument to determine if the episode constituted a Dx.

Findings to Date: In 2018, the algorithm identified 926 cases (7.9% of all ED/UC admissions) of which 251 (27.1%) screened in for review using SaferDx. Reviewers identified 47 Dxes (5.07%); 25.5% required ICU care; one patient died.
None of the cases were identified through existing incident reporting structures or involved legal action. Themes characterizing DxEs include: 1) failure to consider past medical history; 2) failure to consider duration of symptoms; 3) attributing new symptoms to incompletely verified diagnoses. We also identified five cases in which children were misdiagnosed as having new onset migraines and subsequently found to have an intracranial infection or vascular event suggesting a systemic flaw in diagnostic reasoning.

Lessons Learned: An e-trigger applied to patients experiencing unplanned admissions following an ED encounter combined with a detailed chart review process identifies patients suffering DxEs not identified in existing reporting structures. We uncovered at least one systemic flaw in diagnostic reasoning involving children presenting with new onset headaches possibly amenable to a quality improvement initiative.

Impact of Teledermatology Program on Dermatology Resident Experience and Education: A Mixed Methods Analysis

A. Zakaria¹, T. Maurer², E. Ameson¹
¹University of California San Francisco School of Medicine, San Francisco, CA
²Indiana University School of Medicine, Indianapolis, IN

Purpose/Problem: Despite many dermatology residency programs incorporating teledermatology into their curriculums, few studies have analyzed its impact on resident experience and education. To address this gap, we evaluated the teledermatology program at the Zuckerberg San Francisco General Hospital and Trauma Center (ZSFG).

Description of Program, Assessment, or Study: Following teledermatology implementation in January 2015, referring providers were required to upload patient photographs and a brief history through a web-based telemedicine platform for all dermatology patient referrals at ZSFG. During a dedicated weekly session that averaged 100 minutes, a team of 3-4 dermatology residents and an attending dermatologist met to review an average of 70 teledermatology cases. Cases were first reviewed by a resident and then presented to the attending who helped finalize the assessment and plan in the telemedicine platform. For the qualitative analysis, we asked current and former UCSF dermatology residents to provide narrative comments about their teledermatology experiences at ZSFG. These responses were coded emergently and evaluated through inductive thematic analysis with focus on manifest content. For the quantitative analysis, we compared the number of patient cases managed per hour by residents at the in-person dermatology clinic and through teledermatology sessions between June 2017 and December 2017.

Outcomes: Fifteen out of twenty-one potential respondents (71%) provided narrative comments. Our coding unearthed five primary content areas about teledermatology providing high case volume, a low-stress learning environment, opportunities to consider a wider differential, focused teaching on visual skills and practice triaging cases. During our study period, residents managed 4.55 cases per hour in dermatology clinic and 11.49 cases per hour in teledermatology sessions.

Discussion: Our thematic analysis highlighted the features of teledermatology that residents found most beneficial. Pinpointing the aspects of teledermatology that are most educationally salient is an important first step towards establishing guidelines for teledermatology-based education best practices. Our finding that teledermatology enabled more than double the amount of patient cases to be evaluated per unit of resident time is novel, suggesting that teledermatology can be an effective tool in helping dermatology residents develop pattern recognition and visual diagnostic abilities. A potential limitation in the external validity of these findings relates to differences in electronic record systems and the process of implementing teledermatology software across different healthcare settings. Significance of findings: Teledermatology can be an effective component of dermatology residency curriculums by enabling residents to efficiently evaluate a high volume of cases while developing pattern recognition abilities.
Oral Abstracts

Monday, October 19
11:45 AM - 1:15 PM

Identifying Diagnostic Concerns Using ‘Safer Dx Patient': An Engagement Strategy Based on OpenNotes

T. Giardina¹, D. Choi², S. Korukonda³, D. Upadhyay¹, T. Scott¹, C. Spitzmueller⁴, C. Schuerch³, D. Torretti³, H. Singh⁵
¹Baylor College of Medicine, Houston, TX
²Michael E. DeBakey VA Medical Center, Houston, TX
³Geisinger Medical Center, Danville, PA
⁴University of Houston, Houston, TX
⁵Michael E. DeBakey VA Medical Center and Baylor College of Medicine, Houston, TX

Background: Providing patients access to OpenNotes presents a unique opportunity to engage them in diagnostic safety initiatives. We developed and tested a methodology to allow patients to identify diagnostic concerns by accessing OpenNotes to evaluate their progress notes from their recent visits. We also identified predictors of patient-identified diagnostic concerns.

Methods: We created the ‘Safer Dx Patient,’ a questionnaire adapted from the Safer Dx instrument, to help patients evaluate problems within the diagnostic process (i.e., accuracy of symptoms, physical exam, testing concerns, follow-up instructions, care plan, and diagnosis). Safer Dx Patient also collected variables related to trust in the provider and general feeling about the visit. We identified at-risk patients at Geisinger, a large integrated healthcare system, from October 2019 to April 2020 using an electronic algorithm based on patient visit patterns (index primary care visit followed by an unplanned return visit or admission within 14 days). Patients that were included were 18-85 years old and used Geisinger’s patient portal at least once to view their notes. On a daily basis, patients that met the inclusion criteria were sent the questionnaire link in a secure message with instructions to review their clinician’s note and a reminder at week 2. Patients were given a $25 gift card for their time. A multivariate logistic regression model was used to evaluate the association between main outcome, “I feel I was correctly diagnosed during my first visit,” and the dimensions of the diagnostic process and patient-level variables (trust and feeling about the visits).

Results: Of 418 patients, 12.2% (n=51) indicated they felt their diagnosis was incorrect. Multivariate logistic regression showed that patients who indicated that the care plan developed by the provider did not address all medical concerns were 2.8 times more likely to indicate an incorrect diagnosis (95% confidence interval [CI]=1.61-4.68). Additionally, patients who reported an incorrect diagnosis were 2.0 times more likely to not trust the provider (95% CI=1.21-3.36), and 1.4 times more likely to indicate they did not have a good feeling about the visit (95% CI=1.09-1.78).

Conclusion: The ‘Safer Dx Patient’ enables patients to identify diagnostic concerns based on their evaluation of visit notes. Perceptions of care planning, trust in the clinician, and general feeling about a visit play important roles in patients’ perception of a diagnosis. The ‘Safer Dx Patient’ engagement strategy has potential to improve transparency in the diagnostic process, highlight concerns, and lead to better patient-provider relationships.

Factors Associated with Optimal Follow-up Testing Completion in Patients with Incidental Pulmonary Nodules

R. Lacson, L. Cochon, N. Kapoor, S. Desai, R. Khorasani
Brigham and Woman’s Hospital, Boston, MA

Background: Optimal diagnostic follow-up is critical in order to address the Institute of Medicine’s (IOM) mandate for health care professionals to improve diagnostic testing processes. Standardized follow-up of pulmonary nodules often requires diagnostic imaging testing, but the diagnostic imaging process requires multiple steps prior to completion. A
better understanding of the follow-up of pulmonary nodules is needed, as well as identification of factors that might account for variation in the completion of follow-up. Therefore, we sought to assess factors associated with follow-up completion in patients with incidental pulmonary nodules (IPN) on CT scan.

Methods: We conducted a 12-month retrospective cohort study at an academic medical center upon IRB approval. All radiology reports of chest, abdomen and pelvis, or spine CT in 2016 were assessed to determine whether a report contained a pulmonary nodule (excluding those with lung cancer) using a previously-validated natural language processing tool. 278 patients with IPN were then randomly selected and follow-up completion (primary outcome measure), defined as either lung biopsy or chest CT performed within 1 year from the index imaging, was assessed using manual review. The institutional Research Data Warehouse was used to extract: 1) patient-specific features including social determinants of health; and 2) organization-related features (e.g. inpatient). Univariate analysis and multivariable logistic regression were used to determine features associated with follow-up completion.

Results: 278 study patients (chest=90, abdomen=92, spine=96) were randomly-sampled from 6,283 patients with incidental pulmonary nodules, and no known lung malignancy during the study period. The rate of follow-up completion was 80/278 (28.8%). On multivariable analysis, follow-up testing completion of orders for follow-up from the Emergency Department (ED) (OR: 0.16) was associated with decreased follow-up. Other patient-specific features were not associated with follow-up completion (Table 1).

Conclusion: Follow-up completion for IPN on CT scan remains low. Further initiatives should address transitions in care settings and patient hand-offs, especially for ED patients. Interventions that can promote follow-up completion in patients with pulmonary nodules should be addressed comprehensively.

Table 1: Multivariable Analysis of Factors Associated with Follow-up Completion in Patients with Incidental Pulmonary Nodules.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Odds Ratio</th>
<th>95% Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient-Specific Features</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.99</td>
<td>0.97</td>
</tr>
<tr>
<td>Sex, Female</td>
<td>0.64</td>
<td>0.32</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.73</td>
<td>0.18</td>
</tr>
<tr>
<td>Asian</td>
<td>1.09</td>
<td>0.17</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1.83</td>
<td>0.28</td>
</tr>
<tr>
<td>Other</td>
<td>1.57</td>
<td>0.37</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>1.92</td>
<td>0.21</td>
</tr>
<tr>
<td>Non-Hispanic</td>
<td>0.76</td>
<td>0.23</td>
</tr>
<tr>
<td>Insurance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Government</td>
<td>0.72</td>
<td>0.33</td>
</tr>
<tr>
<td>Other</td>
<td>0.29</td>
<td>0.06</td>
</tr>
<tr>
<td>Income</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quartile 1</td>
<td>0.98</td>
<td>0.34</td>
</tr>
<tr>
<td>Quartile 2</td>
<td>1.74</td>
<td>0.71</td>
</tr>
<tr>
<td>Quartile 3</td>
<td>0.89</td>
<td>0.34</td>
</tr>
<tr>
<td>Quartile 4</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Below High School</td>
<td>1.95</td>
<td>0.59</td>
</tr>
<tr>
<td>College</td>
<td>1.05</td>
<td>0.46</td>
</tr>
<tr>
<td>Grad School</td>
<td>0.81</td>
<td>0.15</td>
</tr>
<tr>
<td>Declined/Unavailable</td>
<td>0.28</td>
<td>0.07</td>
</tr>
<tr>
<td><strong>Organization-Related Features</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Care Setting</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Electronic Triggers to Study Diagnostic Errors in Pediatric Emergency Departments

1University of Michigan, Ann Arbor, MI
2Lurie Children’s Hospital of Chicago, Chicago, IL
3Children’s National Medical Center, Washington, DC
4Cincinnati Children’s Hospital Medical Center, Cincinnati, OH
5Children’s Hospital of Philadelphia, Philadelphia, PA
6Michael E. DeBakey VA Medical Center and Baylor College of Medicine, Houston, TX

Background: Diagnostic errors, framed as missed opportunities for improving diagnosis (MOIDs), are a source of risk in pediatric emergency departments (EDs). We sought to identify triggers to screen ED records for diagnostic safety.

Methods: We pilot tested three electronic triggers (eT) to study frequency and contributory factors of diagnostic errors in pediatric EDs: 1) return visits within 10 days resulting in admission (eT1), 2) care escalation or transfer from in-patient unit to intensive care unit within 24 hours of ED presentation (eT2), and 3) death within 24 hours of ED visit (eT3). We created a standardized electronic query and reporting template for the 3 eTs and applied them to electronic health record (EHR) systems of 5 pediatric EDs. All ED visits at each site for a single month (eT1, eT2) and for eT3 for 12 months to yield enough cases for review. After testing for accuracy and face validity of abstracted EHRs, we trained 2 clinicians from each ED to review the triggered cases. Each reviewer manually screened the ED diagnoses at initial and return visits (eT1) and admitting diagnosis and final discharge diagnosis (eT2, eT3) and initially categorized charts as “unlikely for MOIDs” or “unable to rule out MOIDs” without a detailed chart review. For the latter category, reviewers performed a detailed chart review using the validated instrument, Safer Dx, to categorize as MOIDs or no MOIDs, dichotomized as a score = 4 out of 7 on a Likert scale as suggestive of a MOID.

Results: Electronic queries identified 807 ED records, 52 (6.4%) were excluded after manual screen (Figure 1). 74% (559/755), 21% (156/755), and 5% (40/755) charts were triggered for eT1, eT2 and eT3 respectively. Among those, most records; 72.2% (404/559), 84% (131/156) and 95% (38/40) were categorized by the reviewers as unlikely for MOIDS. (Table 1) MOIDs was detected in 11.6% - 66.6% of eT1 and 16% - 100% for eT2 charts. The overall frequency of MOIDs in the e-triggered charts was 4.9% (37/755), 1.6% (12/755), and 0.13% (1/755) for eT1, eT2 and eT3 respectively.

Conclusions: We successfully pilot tested a methodology of using e-triggers and record review to screen large numbers of medical records to identify diagnostic errors in pediatric ED visits. Return visits with admission and transfer to higher level of care can be used as electronic triggers to screen for ED diagnostic errors while death within 24 hours does not appear to be useful.
Improving Communication to Enhance Diagnostic Safety

E. Tooley

1Children’s Hospital Association-Child Health Patient Safety Organization, Washington, DC

Problem Statement: Communication gaps or failures related to diagnostic safety. It is estimated that “80 percent of serious medical errors involve miscommunication between caregivers when responsibility for patients are transferred or handed-off.” This vulnerability also is a top safety concern from serious safety events reported to Child Health PSO. In fact, cases related to diagnostic errors caused by failure to communicate effectively were reported by half of children’s hospitals reporting to Child Health PSO and these cases frequently involved the most serious harm reported.

Description: Child Health PSO, has worked over 18 months with experts to create a patient safety learning tool kit that was disseminated in May 2020. The tool kit aides children’s hospitals, and potentially others, through case learning in building awareness about the problem, identifying and analyzing key vulnerabilities, and understanding the application of evidence-based mitigation strategies (e.g., Chronology of Present Illness, automatic clinical triggers, the take 2-Think Do framework, diagnostic timeout). A gap analysis and a diagnostic team timeout template provides sers with an actionable tool to assess their own internal gaps and mitigating strategies. A Patient Safety Alert emphasizes the fundamental issue and improvement opportunity found by Child Health PSO. Completing the tool kit is a robust listing of additional resources from across the industry. Access the toolkit from https://www.childrenshospitals.org/Quality-and-Performance/Patient-Safety/Resources/Diagnostic-Safety-Toolkit-Main

Results and Lessons Learned: Anticipated results include a potential shift in our understanding of causal factors in diagnostic errors and the generation of awareness about harm involving diagnostic errors. To date five international organizations have accessed the toolkit and Child Health PSO will follow-up on their adoption. While reporting to Child Health PSO is voluntary, we expect an increase in serious safety event reporting and this could lead to more understanding of associated problems and mitigating practices. The long-term anticipated result will be fewer patients experiencing a serious safety event related to diagnostic safety caused by failures in team communication, which potentially will contribute to decreases in each hospital’s overall Serious Safety Event Rate.

Developing a Repository of Diagnostic Errors and Learning Opportunities in a Health Care Organization

D. Upadhyay¹, D. Torretti¹, S. Korukonda¹, Y Mehta¹, C. Schuerch¹, A. Bansal¹, P. Lamphear¹, H. Singh²
¹Geisinger Medical Center, Danville, PA
²Michael E. DeBakey VA Medical Center and Baylor College of Medicine, Houston, TX

Problem: One reason healthcare organizations (HCOs) find it challenging to analyze and address diagnostic errors is lack of an easily available or well-catalogued database. Moreover, legal apprehensions and sensitivity regarding the data usually outweigh the need to be transparent. While quality improvement registries are impacting other fields, there is little similar traction for addressing diagnostic errors. Outside of malpractice claims databases, or research studies, HCOs do not maintain any diagnostic error registries for analysis and learning.

Program: We established a learning health system to better understand how HCOs can systematically identify and learn from diagnostic errors. For diagnostic safety surveillance, we developed an active on-site registry using a “Diagnostic opportunities intake form” (in REDCap) that enabled structured data collection for every case in anticipation of the organization’s operational, quality improvement, research and educational needs. We used a categorization scheme based on the Safer Dx framework (a 5-dimension socio-technical framework to analyze diagnostic process breakdowns). We solicited additional input from clinical leadership and aligned this scheme with elements of the DEER taxonomy.

Findings: Of 578 cases collected since Sept 2017, 223 have been entered in our database as of June 2020. We identified three essential areas of information: (Details in Table 1)

Case details:
• objective details of the diagnostic opportunity and how it was identified

Analysis/learning opportunities:
• brief descriptive details of the case, the diagnostic opportunity, lessons learned, categorization of the errors and harm scale

How we addressed it:
• details of the review, feedback and dissemination process involving departments and frontline clinicians and how case was closed, next steps.

Lessons learned: Categorizing diagnostic errors needs a standardized nomenclature and approach. Developing a diagnostic error database registry is an evolving, resource-intensive and time-consuming process but our tool has benefits in allowing a deep analysis of case-level details, lessons learned and aggregated data. This information provides actionable intelligence for improvement efforts such as identifying local patterns in missed opportunities. This registry and its components could be useful to other health systems and patient safety organizations as they start their journeys to become LEDE (Learning and Exploration of Diagnostic Excellence) organizations.

Table 1: Geisinger CICD Diagnostic Opportunity Case Intake Form.

<table>
<thead>
<tr>
<th>Section 1 (Objective data)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is this a diagnostic opportunity? Yes/No/No, but management issue/Other/Not yet determined</td>
</tr>
<tr>
<td>Medical Record Number: ______________________________</td>
</tr>
<tr>
<td>Age of patient: ______________________________</td>
</tr>
<tr>
<td>Gender of patient: Male/Female/Other</td>
</tr>
<tr>
<td>Study/Registry ID: ______________________________</td>
</tr>
<tr>
<td>Case reported by: Risk/Provider/Patient Experience/ Research Projects/Other</td>
</tr>
<tr>
<td>Name of reporter: ______________________________</td>
</tr>
<tr>
<td>If Provider, how was it shared: Hotline/EHR message/Page/E-Mail/Other</td>
</tr>
<tr>
<td>Date case was received by CICD: ______________________________</td>
</tr>
</tbody>
</table>
Patient Generated Research Priorities: A Systematic Prioritization Exercise

L. Zwaan¹, K. Smith², T. Giardina³, J. Hooftman¹, H. Singh³
¹Erasmus University Medical Center Rotterdam, Institute of Medical Education Research Rotterdam, Netherlands
²Medstar Health, Washington, DC
³Michael E. DeBakey VA Medical Center and Baylor College of Medicine, Houston, TX

Background: Identifying research priorities for diagnostic error reduction is essential. A recent study systematically identified top research priorities for diagnostic safety focusing on topics generated by researchers. However, input from patients and families could complement this researcher perspective and identify additional areas for consideration. We thus aimed to systematically identify patient generated research priorities for diagnostic error reduction.

Methods: We used a systematic prioritization method based on the Child Health and Nutrition Research Initiative (CHNRI) methodology. Patients (including patient advocates) were invited to submit research questions they considered important for diagnostic error reduction. In addition, we invited patients from a Facebook page (Mothers Against Medical Error). The submitted questions were prioritized at an in-person meeting in Washington DC in November 2019 and included 8 patients and 4 researchers. The remaining questions were subsequently re-prioritized by patients scoring the questions on five prioritization criteria 1) usefulness 2) answerability 3) effectiveness 4) potential for translation and 5) maximal potential for effect on diagnostic safety. Previously determined weights were assigned by stakeholders (including risk-managers, funders, educators) and applied to the five criteria to adjust the final prioritization score for each question.

Results: Thirty patients submitted 130 research questions. Eleven members of the Facebook group submitted 41 questions. After consolidation and the in-person meeting, 30 questions were prioritized and scored by the patients who attended the in-person meeting. The stakeholders assigned the following weights to the prioritization criteria (1 being a
neutral weight): Maximum potential for effect on diagnostic safety $\times 1.3$; Effectiveness $\times 1.11$; Potential for translation $\times 1.06$; Usefulness $\times 0.95$; and Answerability $\times 0.75$. Based on the weighted scores, we identified the 10 highest priority questions (Table 1).

**Conclusions:** Using a systematic prioritization exercise, we identified the top-10 patient generated priorities for advancing diagnostic safety research using transparent and objective methods. Patients prioritized not just communication and transition issues but also better system-wide tracking and reporting of diagnostic errors, including patient reporting systems. In addition, they prioritized study of areas of vulnerability where patients are at higher risk for diagnostic errors due to race, gender, or mental health problems. These priorities should be integrated within current research initiatives.

**Table 1:** Overview of the research priorities.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Research question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>How do we implement better integration, coordination, and communication between clinical teams and patients/caregivers to improve the accuracy and efficiency of the diagnostic continuum?</td>
</tr>
<tr>
<td>2</td>
<td>How to accurately track and report diagnostic errors at a health system level?</td>
</tr>
<tr>
<td>3</td>
<td>How do clinician documentation requirements affect the diagnostic process and outcomes?</td>
</tr>
<tr>
<td>4</td>
<td>What specific solutions would address the common contributing factors that affect the diagnostic process for at risk patients such as rural and low health literacy.</td>
</tr>
<tr>
<td>5</td>
<td>How do we identify and decrease gaps in diagnostic care across care transitions?</td>
</tr>
<tr>
<td>6</td>
<td>How do we develop and use a comprehensive public database of misdiagnoses to create learning opportunities to prevent future diagnostic errors?</td>
</tr>
<tr>
<td>7</td>
<td>What does the ideal patient reporting system look like for patients and family members to report diagnostic errors?</td>
</tr>
<tr>
<td>8</td>
<td>What are the best strategies to overcome implicit bias in the clinical encounter. Implicit biases could relate to Race, Ethnicity, Age, Mental Health, language and Gender?</td>
</tr>
<tr>
<td>9</td>
<td>What are the barriers and facilitators patients face while seeking effective second opinions?</td>
</tr>
<tr>
<td>10</td>
<td>What is the frequency of diagnostic issues in failure-to-rescue events post-surgery? For example, someone missed the signs and symptoms to diagnose post-surgery complications/infection sooner.</td>
</tr>
</tbody>
</table>

**Oral Abstracts**

**Tuesday, October 20**
1:30 - 3:00 PM

**Improving Diagnostic Performance Through Feedback - The Diagnosis Learning Cycle**

C. Fernandez Branson¹, C. Friedman², M. Williams², A. Olson¹*
*(Presenting Author)*

¹University of Minnesota, Minneapolis, MN
²University of Michigan, Ann Arbor, MI

**Background:** Diagnostic error is a leading cause of morbidity and mortality, and routine feedback for clinicians is one solution to address this problem. However, work to date has not developed a model of feedback about diagnostic decision-making that is ready to be translated from theory and implemented into practice. To close this gap, we developed the Diagnosis Learning Cycle to describe how clinicians and teams can use specific types of actionable feedback to improve diagnostic decision-making.

**Methods:** Using a 6-step qualitative research process, we surveyed, interviewed, and collaborated with professionals in the health sciences, and in fields such as aviation, meteorology, and team sports, to understand how medical professionals could learn from other fields about how to use feedback to improve performance. We then developed this comprehensive model.
Results: We discovered that improving diagnostic performance is currently hindered by key missing elements that enable performance improvement in other fields. Specifically, we identified the fundamental importance of developing representations of the reasoning process, representations of confidence in diagnosis, and the ability to compare diagnostic hypotheses and outcomes, so that these elements may be used as feedback. Following this discovery, we developed a model that proposes the best way to capture these representations in order to generate feedback to providers. The model also illuminates how this specific type of feedback can contribute to diagnostic learning and calibration, and ultimately improve providers’ diagnostic performance and calibration.

Conclusion: Through discussing feedback and performance improvement with other fields and health sciences professionals, we discovered that healthcare providers may be able to improve their performance by being given feedback about their diagnoses that includes representations of abstract concepts such as reasoning and confidence. In order to have access to this type of feedback, these elements need to be captured in a physical space and a standardized method to compare diagnostic hypotheses and outcomes needs to be developed. Next steps will be to use the model to inform the development of a physical space where these abstractions and outcomes can be stored and test the model to see if it yields improvement in diagnostic decision-making and outcomes.

Use of Chart-Stimulated Recall to Examine Uncertainty in Decision-Making Among Senior Internal Medicine Residents

M. Mutter1, J. Kyle2, E. Yecies2, D. DiNardo2
1University of Colorado School of Medicine, Aurora, CO
2University of Pittsburgh, Pittsburgh, PA

Background: Errors in medical decision making are common and have been linked to adverse events, particularly when uncertainty exists. Prior studies have explored uncertainty but were limited by use of hypothetical cases or by hindsight and/or recall bias. Chart-stimulated recall (CSR) has potential for use in this setting, both to explore uncertainty in real-time and to promote reflection as a means of improving decision-making. In this study, we utilized CSR-based interviews to examine uncertainty in medical decision-making among senior internal medicine (IM) residents. We also sought to evaluate the utility of CSR as an educational tool to promote reflection about uncertainty in decision-making.

Methods: We invited senior IM residents on night float rotations at the University of Pittsburgh Medical Center from February to September 2019 to participate. Each participant completed one, 20-minute CSR session based on a self-selected case from the prior evening’s shift in which there was uncertainty in medical decision-making. Interviews explored the sources of and approaches to uncertainty and were audio-recorded, de-identified, and transcribed. Two independent investigators reviewed and coded the transcripts and identified themes in sources and approaches to uncertainty.
Results: Between February and September 2019, 41/45 (91%) eligible residents participated. The most commonly identified sources of uncertainty were: 1) uncertainty due to incomplete/conflicting information, 2) diagnostic reasoning uncertainty and 3) management uncertainty. In general, approach to uncertainty was driven by the perceived patient acuity rather than degree of uncertainty or source of uncertainty. Some residents utilized a stepwise approach, only attempting to resolve uncertainty when under the impression that doing so would impact immediate overnight care. Others adopted a broader approach, attempting to resolve any uncertainty, even when perceived acuity/urgency was low. Discussion with an attending provided comfort, especially when the attending also reflected uncertainty. Residents rated the exercise as being both comfortable and valuable.

Conclusions: We demonstrated a novel approach to the exploration of uncertainty in night-float residents through CSR. Themes from resident interviews underscore the important role for supervising attendings in this setting, particularly as role models in reflections about and comfort with uncertainty. Additionally, the variability in approach to uncertainty by our residents suggests a need for education regarding ideal approach in various scenarios. In an era when a large portion of new patient admissions are occurring overnight, these findings raise areas of potential focus for both faculty development (role modeling approach to uncertainty) and resident education within IM training programs.

Identifying the Sick Child: Heuristics of the “Doorway Exam” to Predict Hospital Admission

L. O'Neil¹, P. Bhansali¹, J. Bost¹, J. Chamberlain¹, M. Ottolini²
¹Children’s National Hospital, Washington, DC
²Barbara Bush Children’s Hospital, Portland, ME

Background: The ability to rapidly identify seriously ill pediatric patients is important to prevent morbidity and mortality. Physicians rely on intuitive or heuristic decision-making to identify ill patients, but little research has included pediatric patients. Our objective was to evaluate the performance of an intuitive illness rating score, as provided by experienced pediatricians, for predicting admission to the hospital.

Methods: Prospective observational study of attending pediatric emergency medicine physicians at a tertiary academic children’s hospital emergency department. Physicians completed an illness rating score for patients after brief review of the medical record and a rapid visual evaluation from the doorway. The illness rating score was a continuous scale anchored by 0 (totally well) and 10 (critically ill). Chart review was then conducted to extract outcome variables. Statistical analyses compared the illness rating scores for predicting the primary outcome of admission and the secondary outcomes of respiratory support, intravenous (IV) antibiotics, and resuscitative IV fluids by constructing area under the curve (AUC) of the receiver operating characteristics curves. Optimal illness rating cut-off scores for predicting the outcomes were derived using the Youden index method.

Results: Twenty-two physicians with 13 +/- 8 (SD) mean years of experience (range 1-30) completed a total of 146 ratings on eligible patients (mean age 82 months +/- 66 (SD), 59% male, 36% hospitalized). The overall average illness rating score was 3.64 (SD 2.37, range 0-10). Patients requiring hospitalization had significantly higher illness rating scores (not admitted: mean 3.16, 95% CI 2.74-3.59; admitted: mean 4.48, 95% CI 3.76-5.21; p = .001). The AUC for admission was 0.65, with an optimal illness rating cut-off score of 3.75. The illness rating score also predicted the administration of respiratory support (AUC 0.76, optimal cut-off 4.20), IV antibiotics (AUC 0.67, optimal cut-off 5.75), and IV fluids (AUC 0.68, optimal cut-off 2.85) within 24 hours of presentation.

Conclusion: Rapid, intuitive doorway assessments reflected in the illness rating score performed moderately well to differentiate those patients who required admission. Higher illness scores were also associated with markers of illness, including need for respiratory support and intravenous resuscitative fluids and antibiotics.
Table 1: Performance of the intuitive illness rating score for the primary and secondary outcomes based on the Youden cut-off selection criteria and 80% sensitivity.

<table>
<thead>
<tr>
<th>Outcome variable</th>
<th>Youden selection criteria*</th>
<th>At 80% sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cut-off score</td>
<td>Sensitivity</td>
</tr>
<tr>
<td>Admission</td>
<td>3.75</td>
<td>0.62</td>
</tr>
<tr>
<td>Respiratory support</td>
<td>4.20</td>
<td>1.00</td>
</tr>
<tr>
<td>IV antibiotics</td>
<td>5.75</td>
<td>0.50</td>
</tr>
<tr>
<td>IV resuscitative fluids</td>
<td>2.85</td>
<td>0.74</td>
</tr>
</tbody>
</table>

*maximizes sensitivity + specificity, IV = intravenous

Post Handoff Report of Outcomes to Facilitate Patient Follow-up and Reflection for Emergency Medicine Residents

F. Rudolf, E. Pott, L. Oyama, R. El-Kareh
University of California San Diego, San Diego, CA

Purpose/Problem: Self-assessment and reflection of patient outcomes are key components of emergency medicine (EM) training. While EM residents value this outcome feedback, the realities of rapid patient turnover and lack of continuity make this challenging. Our objective was to develop an educational program using an EHR-generated patient report and a facilitated chart review exercise to promote an informed process to patient follow-up and self-assessment for EM residents.

Description of Program: Our EM residents have access to the Post Handoff Reports of Outcomes (PHAROS) report. The PHAROS report is integrated into the EHR and generates a list of patients a particular provider cared for within the prior 14 days with links to patient charts. It also flags important “trigger events” including upgraded level of care, rapid response, and return visit to the emergency department (ED) within 72 hours. We introduced this report with a short video and a series of emails. Residents reviewed their reports and completed a self-reflection worksheet. They then participated in a faculty-facilitated debriefing exercise via video conferencing. Residents completed a survey to assess the effectiveness of the intervention.

Outcomes: 18 residents participated in the patient follow-up exercise. 17 or 94.4% completed the survey. Survey responses are presented in Table 1. Notably, 100% of respondents agreed that the follow-up exercise was a valuable learning experience, and 82.4% felt they learned something that would change their clinical practice.

Discussion: Residents overwhelmingly reported that the PHAROS report made it easier to follow-up on their patients, and 100% of residents felt the structured reflection exercise was a valuable learning experience. Within EM, it is difficult to find a balance between service-related activities and personalized education. A small group faculty led debrief led to rich discussions on patient outcomes and self-reflection. We believe these sessions will help residents perform self-assessment and identify areas for continued self-improvement.

Significance of Findings: Our findings demonstrate that having a structured methodology for evaluation of patient care practices overall makes it easier for residents to perform self-assessment and practice-based improvement, aspects of learning that are required by the Accreditation Council for Graduate Medical Education (ACGME).

An Innovative Virtual Rounding Solution to Focus on Clinical Reasoning

M. Sakumoto, S. Sukumar, A. Zakaria, R. Khanna, C. Lai, N. Choi
University of California San Francisco, San Francisco, CA

Purpose/Problem: Due to COVID-19’s impact on the hospital and medical education system, third year medical students (M3s) were temporarily excluded from on-site clinical rotations. The Internal Medicine (IM) clerkship at UCSF created a Virtual Rounding (VR) curriculum focused on helping students develop skills in oral presentation and clinical reasoning.
**Program Description:** The VR tele-team structure included one IM attending, two teaching assistants (IM resident and/or fourth year medical student), and three M3s. The teams utilized video conferencing (Zoom, Version 4.6.10, San Jose) and remote electronic health record (EHR) access. M3s followed the clinical course of one hospitalized patient per week but were not directly involved in care. They first “pre-rounding” electronically, and later in the morning, listened to the intern’s presentation on hospital rounds. Finally, in the afternoon, they remotely met with the VR team for one hour, during which each M3 delivered an oral presentation on their patient. The tele-attending and teaching assistants provided real-time feedback and clinical teaching. In addition to VR, the IM curriculum included daily virtual attendance at Morning Report (clinical problem solving cases) and Noon Conference (didactics on medicine topics). Initial program assessment was conducted via semi-structured interviews (via email and teleconference) with the tele-attendings, teaching assistants, and students.

**Outcomes:** The initial 2-week pilot period included 10 students, 8 TAs, and 5 tele-attendings. A few themes emerged after reviewing initial feedback including 1) appreciation from attendings and students for the extra time devoted to teaching and coaching, 2) the fact that the VR curriculum complemented other virtual didactic sessions they were receiving, and 3) increased medical student confidence around clinical reasoning. Constructive feedback included building more time for orientation to the EHR and the ability to directly video or audio conference with the patient.

**Discussion:** VR allowed the tele-attending and teaching assistants to focus on M3 skill development around acquisition and interpretation of clinical information. This included extensive focus on forming a problem representation, prioritizing a problem list, generating appropriate differentials, and refining oral presentation skills.

**Significance:** VR offered M3s an opportunity to practice clinical data synthesis and reasoning during a time when learning through direct patient care was not possible.

---

**Dx Reasoning Challenge: An Online Platform for Education, Teamwork, & Friendly Competition between Practicing Physicians**

K. Tran¹, M. Arteaga-Garavito¹, A. Patel², B. Wells², M. Hugh¹, J. Mangun³, G. Castro³

¹Southern California Permanente Medical Group, Pasadena, CA
²Human Dx
³Society to Improve Diagnosis in Medicine, Evanston, IL

While resources aimed towards increasing awareness and reducing the rate of diagnostic errors are available for physicians in training, few collaborative educational programs are available for practicing clinicians.

The Southern California Permanente Medical Group (SCPMG) partnered with the Human Diagnosis Project (Human Dx) and the Society to Improve Diagnosis in Medicine (SIDM) to implement a Quality Improvement (QI) Education Intervention for physicians: The Diagnostic Reasoning Challenge.

This project aimed to assess outcomes in diagnostic accuracy and collaboration in an innovative virtual educational format. We recruited physicians from three Kaiser Permanente regions. Participants were randomly assigned to 3 teams (2 intervention and 1 control). In the pre-intervention phase, participants solved a set of Human Dx Global Morning Report (GMR) cases. In the 6-week intervention phase, each team received weekly educational materials using an online collaboration platform. The intervention groups received content focused on diagnostic errors in cancer, infectious diseases, and cardiology (areas with highest rates for diagnostic errors according to the literature). The control group received educational materials on other topics. Educational materials were provided in the forms of videos, articles, surveys, and clinical cases. In the post-intervention phase, participants were asked to solve another set of GMR cases.

Performance and collaboration were monitored throughout the project. Physicians who completed all phases of the program were eligible to obtain CME and MOC credits. SIDM developed an implementation package with the learnings of this intervention for replicability and scaling.
Outcomes:
- Created a culture of learning to help improve diagnostic reasoning skills
- Measured diagnostic quotient
- Facilitated collaboration among practicing physicians from different geographical regions
- Documented the benefits of combining various virtual learning tools
- Demonstrated high participation interest and engagement for virtual educational content
- Developed successful friendly competition environment

As in any research education program, a challenge encountered was the recruitment and retention of participants. The benefit of having a flexible virtual participation opportunity allowed participants to engage and collaborate on an innovative platform. Although this has proven to be a successful educational intervention, additional initiatives with diagnostic reasoning focus will help reinforce the outcomes of this project. This intervention demonstrated it is possible to successfully implement a virtual education program targeting busy practicing physicians. The structure of the project is flexible, allowing physicians to participate on their own time and engage in meaningful collaborative discussions. These types of interventions are applicable to multiple settings and promote diversity of practicing clinicians.

**Poster Session 1**

**Monday, October 19**

**2:15 PM - 3:15 PM**

**Scientific**

**A Framework to Implement SPADE Using Deep Learning**

N. Brestoff
Intraspeixon LLC, Sequim, WA

**Background:** In 2018, Drs. Liberman and Newman-Toker co-authored “Symptom-Disease Pair Analysis of Diagnostic Error (SPADE): A Conceptual Framework and Methodological Approach for Unearthing Misdiagnosis-Related Harms Using Big Data.” Previously, in 2016, my colleagues and I built an end-to-end software system using a deep learning model to identify and provide an early warning of potential litigation in the category of Civil Rights-Employment (“employment discrimination”). In 2017, after receiving five related patents, we received a patent for “Using Classified Text and Deep Learning Algorithms to Identify Medical Risk and Provide Early Warning,” No. 9,754,220 (Priority Date: July 1, 2016). In that patent, Claim 1 mentions “one or more training datasets including but not limited to textual data corresponding to one or more specific medical diagnoses of interest.”

**Methods:** The deep learning model for “discrimination” was trained with positive and negative documents. The positive documents consisted of 400 sets of factual allegations from complaints for “employment discrimination.” The negative documents consisted of any other text. We tested the model on subsets of Enron emails.

**Results:** Initially, the model was assessed by Indico Data Systems, Inc. (Boston MA) as “strong” because the AUROC score was 0.967 out of 1.000. Later, when we used “transfer learning” and added examples of text in emails, we obtained an AUROC score of 0.99+. A t-distributed Stochastic Neighbor Embedding (t-SNE) analysis (Fig. 1) showed why: The training data had resulted in a clear “decision boundary.” When we input “test” data consisting of 20,401 Enron emails, the model (in a Graphics Processing Unit inside AWS) scored, ranked and reported only 25 emails as being “related” to the discrimination risk, a fraction equal to 0.0012254. Of the 25 emails related to “discrimination,” four were True Positives. When True Positives are saved, each symptom-disease model may be re-trained and will get sharper.

**Conclusion:** Facts alleged in complaints are to categories of litigation as patient symptoms are to categories of disease. With AWS for speed and parallel processing, and NLP advances like BEHRT and BERT-XML for EHRs, if broadly available, we hope to work with others to more fully implement the SPADE framework.
How Clerkship Students Learn Unfamiliar Diseases

E. Cabral¹, G. Dhaliwal², A. Teherani¹, D. Connor²
¹University of California San Francisco, San Francisco, CA
²UCSF, SF VA Medical Center, San Francisco, CA

Background: In response to the persistent problem of diagnostic error, the National Academies of Sciences’ report ‘Improving Diagnosis in Health Care’ calls on medical schools to enhance diagnosis education. Building effective illness scripts is essential for the development of diagnostic competence, yet the learning approaches medical students use to build illness scripts for unfamiliar diagnoses they encounter in the clinical environment are unknown.

Methods: We performed an exploratory qualitative study using thematic analysis. Ten clerkship students from a single institution participated in a learning task and semi-structured in-person interview. To simulate a clerkship setting, participants were prompted to imagine that they were about to see a new clinic patient with Gaucher disease (interviewers first confirmed that students were unfamiliar with this disease). While using self-selected online resources, participants were asked to verbalize their thinking as they learned about Gaucher disease for 15 minutes. Participants then described their approach to learning about Gaucher disease. A code book was developed through open coding by two investigators, followed by comparison and consensus on codes, using memos to document evolution of codes over time.

Results: Ten participants completed the study. We identified common preferences regarding resources, tools, strategies, prioritized information, and motivations for learning. Resources used included UpToDate, Google search, Google images, Wikipedia, YouTube, PubMed, GeneReviews, and disease foundation websites. Participants sought information about signs and symptoms, complications, management, and basic pathophysiology, while deprioritizing epidemiology, genetics, differential diagnosis, medications, and unusual presentations. Motivations for learning included the desire to be useful in the clinical encounter, to avoid embarrassment, to prevent patient harm, and to appear knowledgeable.

Conclusion: When learning about an unfamiliar disease, clerkship students seek resources that provide concise explanations. They prioritize information they perceive as most clinically relevant, tailoring their learning to their understanding of the clinical task at hand. Students’ motivations include both mastery (desire to learn) and performance (desire to appear competent) orientations as described in goal orientation theory. When caring for patients with established diagnoses, students deprioritize learning about differential diagnoses, overlooking comparative learning. This learning approach represents a missed opportunity to build contrastive illness scripts for a given clinical problem, leaving students without the tools to differentiate between related diagnoses in future encounters where diagnostic
uncertainty exists. Coaching students on best practices for structuring their learning during clerkships may enhance the development of diagnostic competence.

References to Clinical Reasoning in Teaching Evaluations are Associated with Higher Teaching Effectiveness Scores

C. Clancy1, N. Etherington2, M. Barrosse-Antle1, C. J. Dine3
1Perelman School of Medicine University of Pennsylvania, Philadelphia, PA
2Temple University School of Medicine, Philadelphia, PA

Background: A critical task of the clinical teacher is to facilitate development of trainee clinical reasoning skills. While there have been multiple studies examining characteristics of ideal clinical teachers, and the relationship between attributes and teacher ratings, no studies have explored whether references to clinical reasoning teaching skills are associated with teaching effectiveness ratings.

Methods: We performed mixed-methods retrospective analysis of evaluations of medicine faculty members at a single large academic medical center, consisting of 7,424 unique evaluations. We first performed qualitative analysis to sort comments based on whether they contained positive, negative or no reference to clinical reasoning. Coding was performed by three coders in an overlapping design, and interrater reliability was assessed with kappa testing. As the primary analysis, multivariate logistic regression was performed to test for association between positive and negative references to clinical reasoning and perfect scores on numeric ratings of teaching effectiveness, controlling for faculty gender, trainee gender, and level of training. In addition, thematic analysis was performed to identify themes in comments containing references to clinical reasoning.

Results: Positive and negative references to clinical reasoning were identified in 1037 (14%) and 96 (1.3%) of evaluations, respectively, with substantial agreement between raters (k=0.796). In our multivariate logistic regression controlling for trainee and faculty gender and training level, positive references to clinical reasoning were positively associated with perfect teaching effectiveness score (OR=1.89, p<0.001, 95% CI 1.55 - 2.31). Negative references to clinical reasoning were negatively associated with perfect teaching effectiveness scores (OR=0.58, p<0.001, 95% CI 0.032 - 0.106). Thematic analysis of comments containing references to clinical reasoning identified four behaviors that trainees identified as positive clinical reasoning education: explicitly teaching clinical reasoning skills/frameworks, eliciting clinical reasoning from trainees, allowing trainees to express their own clinical reasoning, and demonstrating their own clinical reasoning.

Conclusions: Our single-center, mixed-methods study of clinical teaching evaluations demonstrated a significant relationship between references to clinical reasoning in narrative comments and teaching effectiveness scores. This suggests that ability to demonstrate and teach clinical reasoning is a valued skill for clinical teachers and should be a target of faculty development. Qualitative analysis of evaluations identified a range of behaviors that represent potential strategies for incorporating clinical reasoning into clinical teaching.

Improving Diagnostic Safety in Health Care Organizations: Insights from Safety Leaders

T. Giardina1, U. Shahid1, U. Mushtaq1, D. Upadhyay2, A. Marinez1, H. Singh3
1Baylor College of Medicine, Houston, TX
2Geisinger Health, Danville, PA
3Baylor College of Medicine and Michael E. DeBakey Houston VA, Houston, TX

Background: Most health care organization (HCO) leaders find diagnostic errors harder to address compared to other safety concerns. We identified main organizational challenges and potential best practices to ensure diagnostic safety.

Methods: We interviewed 32 health system safety leaders across the US as a first step to develop organizational best practices in diagnostic safety. Participants were recruited through email and represented geographically diverse academic and non-academic settings caring for adults and/or children. The interview guide included questions on culture of
reporting and learning, quality and safety, training and education, and engagement of clinical leadership/staff and patients & families. We conducted a content analysis of interview transcripts.

**Results:** Of the 32 respondents, 12 reported having a program in place to address diagnostic errors. HCOs face two main reported barriers to diagnostic safety: 1) organizational culture and 2) measuring and monitoring (Figure 1). Within the theme of organizational culture, three main subcategories of challenges emerged: psychological safety, HCO stakeholder engagement, and diagnostic error identification and learning opportunities. Despite use of incident reporting, lack of a learning environment created fear and stress among frontline clinicians for acknowledging and reporting diagnostic errors. Additionally, incident reporting favored confidentiality, limiting feedback to clinicians about opportunities for learning and improvement. Leadership commitment was particularly viewed as essential for allocation of resources to move from ideology to action. Second, measuring and monitoring was a significant challenge, with concerns such as lack of an operational definition of diagnostic errors, resource and infrastructure constraints, and addressing cognitive biases that contribute to diagnostic error. Suggestions for improvement included pragmatically defining and operationalizing the term ‘diagnostic error’ as a first step towards measurement. Respondents also emphasized documentation of differential diagnoses to reinforce transparency by capturing clinicians’ critical thinking within the medical record. Other suggestions included clinician education/feedback and support, adopting OpenNotes, and investing in AI to improve diagnostic accuracy and to support clinician cognitive skills.

**Conclusions:** Leading organizations identified specific barriers for diagnostic safety and suggested actions to bolster improvement. Next steps should include development of organizations that learn and explore diagnostic excellence.

**Creating a Rectal Bleeding Safety Net to Improve Colorectal Cancer Detection and Prevention**

T. Imley, M. Kanter, A. Adams, N. Gin, T. Ho, M. Pruitt, R. Timmins, R. Chandiramani
Southern California Permanente Medical Group, Pasadena, CA

**Background:** Rectal bleeding is a common symptom in adults yet follow up of this finding is not always consistent. Colonoscopy is advised for older patients with rectal bleeding due to higher risk of bleeding caused by colorectal cancer. We know that less than half of patients presenting with rectal bleeding receive colonoscopies potentially delaying or missing the diagnosis of colon cancer or clinically significant adenomas. In our system we wanted design a safety net to capture this population, and ensure they had access to colonoscopies.

**Methods:** In an integrated healthcare delivery system serving >4.6 million patients, we developed a centralized system (Sure Net) leveraging our electronic medical records (EMR) to screen for members between the ages of 45-80 with a recent diagnosis of rectal bleeding, but no colonoscopy referral sent at the time of rectal bleeding diagnosis and no colonoscopy

---

**Figure 1:** Main Barriers to Diagnostic Safety Identification by Safety Leaders from 32 Health Care Organization.
in the last 5 years. Gastroenterologists reviewed these medical records, and if a colonoscopy was indicated, a nurse engaged the patient, and placed a referral for colonoscopy to be cosigned by the primary care physician.

**Results:** From December 2014 to December 2019, we identified 1433 patients aged 45-81 years, who met our inclusion criteria. These members agreed to a colonoscopy after being in contact with our safety net system. The colonoscopy findings of this group showed that 9.6% of this population had cancer or advanced adenomas and 34.3% had polyps on colonoscopy that warranted more frequent surveillance than routine screening. Two patients were found to have Lynch Syndrome and 1 was BRCA-positive after being captured in our system. Of the 18 colorectal cancers detected, 83% were pre-metastatic at diagnosis.

**Conclusion:** Leveraging EMR data and centralized processes, we developed a safety net system that effectively identified and screened 1433 people with rectal bleeding with no recent colonoscopy. Using this program, we detected colorectal cancer 18 patients, and non-colorectal cancer in an additional 2 patients. In addition, we detected significant abnormalities in 34.3% of this population. This process effectively ensured appropriate colonoscopy testing in those with rectal bleeding. Ongoing improvements will include designing a feedback loop to physicians who diagnose rectal bleeding, to ensure colonoscopies are ordered when appropriate.

**The Identification and Characterization of Diagnostic Process Errors among Patients with Bacteremia Using a Trigger**

A. Li\(^1\), A. Lim\(^2\), M. Mirica\(^2\), D. Kobewka\(^1\), J. Ranshoff\(^2\), N. Carlile\(^2\), S. Shah\(^2\), G. Schiff\(^2\)

\(^1\)The Ottawa Hospital, Ottawa, ON, Canada

\(^2\)Brigham and Women’s Hospital, Boston, MA

**Background:** Accurate and timely diagnosis of sepsis is a critical clinical problem, and identifying sepsis related diagnostic errors and delays represents an important challenge. We aimed to evaluate a method for identifying and characterizing diagnostic error among patients with bacteremia using a trigger definition of a positive blood cultures and failure to initiate antibiotics within the ensuing 24 hours, suggesting a potential missed or delayed diagnosis.

**Methods:** We used hospital electronic medical records (EMRs) to create a cohort of all patients with positive blood cultures drawn within 12 hours of presentation at two academic hospitals (Hospital A and Hospital B). We excluded cases of probable-contaminated blood cultures. In our study, the trigger definition was the absence of antibiotic administration within 24 hours of presentation. We applied this trigger definition by matching the cases in our study with data on antibiotic ordering and administration. For cases meeting the trigger definition, we performed a chart review to identify diagnostic process errors (DPEs) using the Diagnostic Error Evaluation and Research (DEER) tool. Chart review for each case was performed independently by two members of our team and conflicts in scoring were adjudicated by a third member.

**Results:** At Hospital A, we identified 1955 of positive blood cultures excluding contaminants in the year 2016 and 2017. Of these cases, 44 met the trigger definition. After chart review, we identified DPEs in 30 of cases. The yield of the trigger in identifying DPEs at Hospital A was 68.2% (30/44). At Hospital B, we identified 1163 positive blood cultures excluding contaminants in the year 2016, of which 73 met the trigger definition. After chart review, we identified DPEs in 68 of the cases. The yield of the trigger in identifying DPEs at Hospital B was 93.2% (68/73). DPEs occurred most commonly in the category of “Assessment” at both hospitals.

**Conclusion:** In this study, we demonstrated the feasibility and yield of a method to identify DPEs by linking data on bacteremia and medications from hospital databases and applying a trigger definition. Review of these cases demonstrated a variety of diagnostic process errors, as well as limitations of this trigger—mainly trigger positive patients where clinicians were appropriately withholding initiation of antibiotics. The approach was helpful in understanding the burden and characteristics of bacteremia related DPEs. This method can be applied to other important diagnoses if an appropriate trigger definition exists.
A Simulation-Based Approach to Estimate the Incidence, Duration, and Risk Factors for Diagnostic Delays

A. Miller, P. Polgreen
University of Iowa, Iowa City, IA

**Background:** Diagnostic delays are an important type of diagnostic error and a major contributor to morbidity and mortality. Investigating delays can be challenging or costly using existing methods (e.g., retrospective chart reviews, autopsy studies); these methods often cannot estimate population-level incidence and may not be feasible for some diseases. We develop a methodological framework to evaluate “likely” diagnostic delays associated with infectious diseases using a large administrative data source. We evaluate our approach applied to delays for tuberculosis (TB) and herpes simplex encephalitis (HSE).

**Methods:** We used data from the IBM Marketscan Research Databases from 2001-2017; these represent one of the large sources of inpatient, outpatient and emergency department data in the United States. We identify symptomatic healthcare visits occurring prior to the index diagnosis, including symptoms, diseases or syndromes that suggest infection is present (e.g., cough or pneumonia prior to a TB diagnosis). We use a time-series change-point algorithm to identify the point in time when the frequency of symptoms increases and then estimate the incidence of delays prior to diagnosis. We also consider “control” conditions that are not symptomatically similar (e.g., fractures). Finally, we develop simulation models to evaluate individual-level outcomes including delay duration and risk factors.

**Results:** We find a dramatic spike in symptomatic visits prior to TB and HSE (see figure) but do not find an increase in care for unrelated “control” conditions. We evaluated 3,158 and 4,470 cases of TB and HSE, respectively. Of these, 80.1% (95% CI 77.9-82.2%) and 24.7% (CI 22.8-26.8) of patients with TB and HSE experienced a delay. Our simulation results identify an average delay duration of 41.06 days (CI 36.83-45.24) for TB and 10.8 days (CI 8.64-13.51) for HSE. We found that diagnostic delays are more common in areas where diseases are less prevalent, or in patients with a history of related symptoms (e.g., migraines before HSE, asthma before TB).

**Conclusions:** For the infections we evaluate, our results were consistent with existing literature regarding delays for these conditions. Our methodological framework can be applied to a range of diseases using large existing data sources.

Exploring Uncertainty and the Growth Mindset Among Senior Internal Medicine Residents

M. Mutter, E. Yecies, D. DiNardo
1University of Colorado School of Medicine, Aurora, CO
2University of Pittsburgh, Pittsburgh, PA

**Background:** Errors in medical decision making have been linked to adverse events and patient harm, particularly when diagnostic or therapeutic uncertainty exists. Prior research has attempted to categorize uncertainty and to associate physicians’ tolerance of uncertainty with various factors such as patient-provider communication and diagnostic testing. The “growth mindset,” or the perceived ability to change one’s own intelligence level, has been well studied in the...
psychology literature but its utility in medical decision making has not been well-described. In this study, we sought to determine senior internal medicine (IM) residents’ reactions to uncertainty and the association of tolerance of uncertainty with the “growth mindset.” We also sought to determine the association of resident demographic characteristics with tolerance of uncertainty and the growth mindset.

**Methods:** We conducted a study to explore uncertainty in medical decision-making with senior IM residents at the University of Pittsburgh Medical Center from February through September 2019. Participants completed a demographic survey, a Physicians’ Reactions to Uncertainty (PRU) scale, and a Revised Implicit Theories of Intelligence or “growth mindset” scale. Spearman’s rho, Wilcoxon Rank-Sum Test and Kruskal-Wallis were calculated as measures of association, where appropriate.

**Results:** Between February and September 2019, 41 out of 45 eligible residents participated in the study. Residents had a moderate overall “anxiety due to uncertainty” (average 18.4/27), though did demonstrate less “reluctance to disclose uncertainty to patients” (average 19.1/24) and more “reluctance to disclose mistakes to physicians” (average 4.73/10), with higher scores on these sub-scales suggestive of greater affective responses to uncertainty (i.e. more anxiety or reluctance). Residents scored an average of 2.6 out of 6.0 on the “growth mindset” scale, with scores less than 3.1 suggesting a growth mindset. No association was found between age, gender and level of training on tolerance of uncertainty or having a growth mindset and no association was found between greater tolerance of uncertainty and having a growth mindset.

**Conclusions:** Our results demonstrate that residents had a moderate comfort level with uncertainty overall, though the ideal amount of uncertainty to provide optimal patient care is unknown. Relatively lower scores on the “reluctance to disclose mistakes to physicians” sub-scale suggest a greater need for educational efforts in error disclosure in order to ultimately reduce adverse events. In addition, though no association was found between greater tolerance of uncertainty and the “growth” mindset, other trainee factors may be related to tolerance of uncertainty and should be explored.

**Patient Journey Mapping: Visualizing Diagnostic Error Experiences**

T. Scott¹, H. Haskell², H. Singh¹, G. True³, T. Giardina¹  
¹Baylor College of Medicine and Houston VA, Houston, TX  
²Mothers Against Medical Error  
³New Orleans VA, New Orleans, LA

**Background:** Patients’ experiences of the diagnostic process, unlike other patient experience data, is more difficult to collect and monitor as patients continuously transition between clinicians and care settings. Patient journey mapping, a visualization of a patient’s experience interacting with the healthcare system, may highlight problem areas that would otherwise be challenging to capture. Our study aimed to explore the use of patient journey mapping in diagnostic error cases to reveal critical gaps in care processes.

**Methods:** We conducted narrative interviews with patients and family members who experienced a diagnostic error. Participants were recruited via a patient advocacy email listserv and social media. Interviews were conducted by telephone (range: 40-120 minutes) and participants were compensated for their time. The interviews were unstructured, and participants were asked to share their diagnostic error experience as a ‘story’. Probing questions were asked for clarification purposes. Interviews were audio recorded and transcribed verbatim. Final transcriptions were approved by each participant. Using narrative analysis, we created timelines of events for each story and identified key characters and points of interaction with the healthcare system. We then created visual maps of events within the diagnostic process and looked for salient and divergent themes across cases.

**Results:** Of six interviews analyzed thus far, final diagnoses included: dysautonomia-related diagnoses, celiac disease, and a missed tumor. Clinicians commonly attributed symptoms to mental health and communication problems issues were reported in every story. Despite diminished trust, participants continued to seek a diagnosis with multiple clinicians over multiple visits (Figure). A salient theme across the maps was the action/intervention of a doctor who supported the patient journey and helped change the trajectory of the diagnostic journey by refocusing on symptoms, ordering appropriate testing, and making referrals. Supporting the patient journey included clinicians responding with empathy.
and acknowledging trauma, and expressing curiosity about symptoms, transparency about uncertainty, and tenacity toward a diagnosis.

**Conclusion:** Our findings highlight the potential of patient journey mapping to reveal barriers, facilitators, and gaps in care across patients’ experiences of the diagnostic process. This methodology can identify areas of improvement in the diagnostic process most valued by patients.

---

**Figure 1:** Patient Journey Map Example.

**Patient Harm due to Diagnostic Error of Neuro-Ophthalmologic Conditions**

L. Stunkel\(^1\), R. Sharma\(^2\), D. Mackay\(^3\), B. Wilson\(^1\), G. Van Stavern\(^1\), N. Newman\(^2\), V. Biousse\(^2\)

\(^1\) Washington University in St. Louis, St. Louis, MO

\(^2\) Emory University, Atlanta, GA

\(^3\) Indiana University, Bloomington, IN

**Background:** Neuro-ophthalmic conditions are frequently misdiagnosed, but little is known about patient harm resulting from these misdiagnoses. Our aim is to prospectively examine diagnostic error of neuro-ophthalmic conditions and resultant harm at multiple sites.

**Methods:** Prospective cohort study of 437 consecutive adult patients seen in consultation at three university-based neuro-ophthalmology clinics in the United States. Referral documentation was reviewed to collect details regarding prior care, which were confirmed with patients at the time of the office visit. Demographics, final diagnosis, diagnostic testing, treatment, patient disposition and impact of the neuro-ophthalmologic encounter were collected. For patients whose referral and final diagnosis were different, we identified the cause of error using the Diagnosis Error Evaluation and Research (DEER) taxonomy tool and assessed whether the patient suffered harm due to the misdiagnosis.

**Results:** Referral diagnosis was incorrect or incomplete in 47% of cases. 29% of misdiagnosed patients suffered harm, which could have been prevented by earlier referral to neuro-ophthalmology in 97%. 22% of patients had experienced either inappropriate laboratory testing, diagnostic imaging, or treatment prior to referral, with higher rates for patients misdiagnosed prior to referral (33% of patients compared to 13% with a correct referral diagnosis, \(p<0.0001\)). 89% of referrals were appropriate (defined by whether they asked a neuro-ophthalmic question). 75% of inappropriate referrals were misdiagnosed, compared to 43% of appropriate referrals (\(p<0.0001\)). Median number of providers seen prior to neuro-ophthalmology consultation was 2 (IQR 2-3), with range from 0-22. The most common reasons for referral were
idiopathic intracranial hypertension (18%), diplopia or cranial nerve palsy (16%), optic atrophy (11%), and unspecified vision loss (10%). The most common reasons for diagnostic error were incorrect performance or interpretation of the physical examination (37%), incorrect generation of the differential diagnosis or diagnostic reasoning (24%), incorrect performance or interpretation of the history (23%), and incorrect utilization or interpretation of diagnostic testing (11%).

**Conclusion:** This is the first prospective, multisite study of diagnostic error of neuro-ophthalmic conditions, and the first to directly evaluate harm due to misdiagnosis. Misdiagnosis of neuro-ophthalmic conditions and preventable harm due to misdiagnosis were common. Mismanagement prior to referral (unnecessary laboratory tests, imaging, and treatments) was common, exposing patients to unnecessary risk and medical costs. Early, appropriate referral to a neuro-ophthalmologist has the potential to prevent patient harm and make efficient use of diagnostic resources.

**Prediction of High Risk of Death amongst Elderly Patients Presenting to Emergency Departments in the United States**

D. Su, Y. Chen, K. He, X. Zhang
University of Michigan, Ann Arbor, MI

**Background:** Early and accurate prediction of high risk of death and overcoming the limitations of the Emergency Severity Index (ESI) system for elderly patients presenting to emergency departments (EDs) may help optimize the utilization efficiency of medical resources in the ED and the clinical outcomes. We aimed to use natural language processing (NLP) and machine learning (ML) techniques to examine the association between high mortality risk and demographic information, visit information, and clinical factors for elderly ED patients, and to compare these techniques with ESI-based reference models to support early classification diagnosis.

**Methods:** We included 36,283 elderly ED patients in the National Hospital Ambulatory Medical Care Survey datasets from 2009 to 2017. Three variables were recoded to reflect the high risk of death for elderly ED patients (outcome variable). We developed four machine learning models (lasso regression, gradient-boosted decision tree, support vector machine, and deep neural network) based on medical record information (in which text information was processed using NLP). We then compared them with ESI-based reference models. Prediction performance was measured by computing C statistics, prospective prediction results (sensitivity, specificity, accuracy, positive predictive value, and negative predictive value), and decision curve analysis (DCA).

**Results:** Amongst the 36,283 elderly ED patients between 2009 and 2017, 1,776 (4.9%) had high risk of death. The AUC value of the reference model was the lowest (AUC=0.723), and the AUC values of logistic regression with lasso regularization (AUC=0.820; p-value<0.001), gradient-boosted decision tree (AUC=0.815; p-value<0.001), and deep neural network (AUC=0.802; p-value<0.001) were significantly higher than those of the reference model. In the DCA, the net benefit of lasso regression and gradient-boosted decision tree were larger than those of the other models within the 10% threshold probability. At the ESI levels of urgent, semi-urgent, and non-urgent, the predictive appearance of the reference model was lower than those of the ML models. Some important variables affecting predictability such as arrival by ambulance, triage level, and temperature were identified.

**Conclusion:** The ML model based on medical record and NLP technology can effectively improve the prediction performance of high risk of death for elderly ED patients based on ESI. These models may be integrated into the ED classification to quickly identify elderly ED patients who may be at high risk of death and to assist medical decisions.
A Portfolio of e-Triggers to Identify Diagnostic Errors in Emergency Departments: A Prioritization Exercise

V. Vaghani1, D. Murphy2, S. Memon2, A. Zimolzak2, D. Subramanian3, D. Upadhyay4, H. Singh1
1Baylor College of Medicine & Michael E. DeBakey VA Medical Center, Houston, TX
2Baylor College of Medicine, Houston, TX
3Rice University, Houston, TX
4Geisinger Health System, Danville, PA

Background: Frequency of diagnostic errors in the emergency department (ED) is unknown but likely higher than outpatient care. Rules-based e-trigger algorithms have been used to identify selected patients likely to have diagnostic error, but trigger precision is limited. As a first step to determine frequency of ED diagnostic error, we aimed to create a portfolio of e-triggers to identify ED records likely to involve diagnostic error.

Methods: We are using the Safer Dx Trigger Tools Framework to develop and validate e-triggers listed below. The current project focuses on Steps 1 and 3.

1. Identification and prioritization of diagnostic error of interest
2. Operational definition of all the criteria for e-trigger
3. Determination of potential data sources
4. Construction of e-trigger algorithm
5. Testing of e-trigger tools on the data sources
6. Assessment of e-trigger tool performance
7. Iterative refinement of e-trigger algorithm.

We first consulted a panel of 6 experts, whose roles included physicians, administrators, informaticians, and/or researchers, for identification and prioritization of high-risk areas of diagnostic errors. In initial consultations, the expert panel helped identify possible areas for e-trigger development using large electronic health record data repositories. In subsequent consultations, the panel ranked each potential e-trigger concept identified on value and impact, feasibility, expected performance, replicability, and novelty using 10-point Likert-type scale for prioritization.

Results: After updated literature searches and initial consultations with experts, we identified 20 possible areas for e-trigger development. The panel prioritized 6 potential e-trigger concepts (Table 1) based on scoring (Steps 1-2 of framework). We identified 2 large data repositories as electronic data sources for the triggers (Step 3): Veterans Affairs (VA) Health Care System and Geisinger and are now operationalizing and matching concepts to data in order to translate them into formal e-trigger algorithms. This is an essential step before further application and testing (Steps 4-7).

Conclusion: An expert panel provided prioritization for development of e-triggers to identify and measure ED diagnostic error using comprehensive EHR data. Newly developed e-triggers will be pilot tested and implemented in order to calculate frequency of diagnostic errors and to create actionable safety-related insights.

Title: A Portfolio of e-Triggers to Identify Diagnostic Errors In Emergency Departments: A Prioritization Exercise

Table 1: Final e-Trigger Concepts Prioritized by Expert Panel

1. Hospitalization after ED visit for high risk presentation (e.g. new onset of abdominal pain with fever)
2. Unexpected transfer from hospital general floor to ICU within 24 hours of ED admission
3. Specific test results without appropriate follow up action (e.g. blood culture)
4. Treat and release ED visit followed by unscheduled return to ED or hospitalization.
5. Treat and release ED visit followed by primary care visit with different diagnosis within 7 days (e.g. Pneumonia preceded by URI)
6. Treat and release ED visit followed by hospitalization for symptom-disease dyad (e.g. acute appendicitis preceded by abdominal pain)
Physicians' Reflection of Diagnostic Error in Internal Medicine: A Cross-Sectional Web Survey in Japan

T. Watari
Shimane University Hospital, Japan

**Background:** In Japan, diagnostic errors have not been well investigated. In addition, there is little useful information on these, due to which cognitive biases tend to be the causes of wrong, delayed, and missed diagnoses. Therefore, we conducted a cross-sectional web survey on Japan’s largest website (Nikkei Medical); this survey included physicians who could recollect their diagnostic error cases that they could most vividly recall, along with the detailed situation and psychological condition.

**Methods:** The recruitment period was 10 days, from January 21 to 31, 2019. We collected detailed information on physicians’ demographic characteristics and background characteristics of diagnostic error cases. To minimize recall bias, we defined the diagnostic error cases as 1) those in which the internist could most vividly recall the situation and factors of the error, and 2) cases in which the internist is primarily involved in the error.

**Results:** We included a total of 2,220 participants over the 10 days, after excluding patients based on our criteria. We analyzed a total of 687 internal medicine physician-related diagnostic error cases. The median age of the participating internists was 50 (IQR 40-58) years, 20% were female. The median post-graduate age was 25 (IQR 14-31) years, and approximately 70% work at the hospital. Two-thirds of the overall cases analyzed were encountered in the general outpatient office or emergency department. Wrong diagnosis (52.0%), delayed diagnosis (35.4%), missed diagnosis (29.5%) were classified through the internist’s reflection. Malignancies (n = 139, 20.2%) and cardiovascular diseases including acute coronary syndrome (n = 125, 18.2%) and stroke (n = 52, 7.6%) were common final diagnoses. A multivariate analysis after adjusting situational factor, data gathering factor, and cognitive bias showed 1) rule bias and confirmation bias for wrong diagnoses; 2) delayed diagnoses also have a strong influence on base rate neglect, and anchoring bias significantly contributes to diagnostic errors; and 3) crowded patient or emergency room as a situational factor and over-confidence bias and Hassle bias were more likely to lead to missed diagnoses.

**Conclusion:** The types of cognitive biases vary in their contribution to wrong, delayed, and missed diagnoses.

### Table: A multivariate analysis after adjusting situational factor, data gathering factor, and cognitive bias.

<table>
<thead>
<tr>
<th></th>
<th>OR</th>
<th>95CI</th>
<th>p-value</th>
<th>adjusted OR</th>
<th>95CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wrong diagnosis 520% (95%/0: 482%-55.7%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rule bias</td>
<td>3.08</td>
<td>0.99-9.54</td>
<td>0.051</td>
<td>322</td>
<td>1.004-10.40</td>
<td>0.049</td>
</tr>
<tr>
<td>Confirmation bias</td>
<td>1.51</td>
<td>1.02-2.24</td>
<td>0.039</td>
<td>1.51</td>
<td>1.00-2.30</td>
<td>0.053</td>
</tr>
<tr>
<td>Delayed diagnosis 35.4% (95%/0: 319%39.0%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base-rate neglect</td>
<td>2.10</td>
<td>1.24-3.54</td>
<td>0.006</td>
<td>1.86</td>
<td>1.03-3.17</td>
<td>0.040</td>
</tr>
<tr>
<td>Anchoring bias</td>
<td>1.10</td>
<td>0.63-1.95</td>
<td>0.728</td>
<td>2.35</td>
<td>1.02-5.42</td>
<td>0.046</td>
</tr>
<tr>
<td>Missed diagnosis 29.5% (95%/0: 26.2%-33.2%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crowded situation</td>
<td>1.68</td>
<td>1.14-2.47</td>
<td>0.009</td>
<td>1.53</td>
<td>1.0-2.35</td>
<td>0.051</td>
</tr>
<tr>
<td>Over-confidence bias</td>
<td>1.65</td>
<td>1.14-2.40</td>
<td>0.008</td>
<td>1.66</td>
<td>1.1-2.52</td>
<td>0.016</td>
</tr>
<tr>
<td>Hassle bias</td>
<td>2.07</td>
<td>1.20-3.58</td>
<td>0.009</td>
<td>3.08</td>
<td>1.42-6.69</td>
<td>0.004</td>
</tr>
</tbody>
</table>
Clinical Vignettes

When the Script Doesn't Fit

J. Alexander¹, C. Jackson¹, D. Burroughs-Ray²
¹UTHSC-Internal Medicine Residency, Memphis, TN
²UTHSC-IM-Pediatrics Residency, Memphis, TN

Case Information: A 76-year-old male with COPD, CAD one month after STEMI secondary to LAD in-stent thrombosis, presenting for the 5th time in 4 weeks for shortness of breath. He had contemplated suicide before arrival citing persistent dyspnea. Before recent STEMI he was independent without home O2. Now he complained of non-exertional right sided chest pain. CXR showed vascular congestion. Troponin, CBC, CMP were unremarkable. Cardiovascular exam revealed elevated JVP, basilar rales, and tachypnea. TTE showed LVEF 30-35%. For each prior encounter, he was treated for CHF and COPD exacerbation simultaneously. On his 3rd presentation to the hospital, CTA of the chest was negative for PE. On current admission, telemetry showed sinus tachycardia, which correlated to the patient’s nightly walks. Wells score estimated a moderate risk for PE. Age adjusted D-Dimer was elevated. CTA was repeated and showed a right middle lobe PE. Patient was treated with apixaban and discharged home with no subsequent readmission for dyspnea.

Discussion: Our case exemplifies several forms of cognitive bias: diagnostic momentum, technology bias, and satisfied search bias. Over reliance on one piece of information, such as a previously negative CT, or multiple admissions for the same problem does not obviate consideration of all likely diagnosis until they are sufficiently excluded. For patients presenting multiple times for similar complaints, this statement is especially true. To minimize cognitive bias, it is necessary to perform a diagnostic time out to avoid premature closure. Considering that “old patients get new problems” before anchoring on a diagnosis is essential. Concern over utilization of diagnostic testing can also delay diagnosis, despite its noble intent. Consistent application of proven risk stratification tools such as the YEARS, Well’s, and PERC scores to may help minimize cognitive bias in clinical decision making around the probability PE, particularly in patients at high risk of delayed or missed diagnosis.

IgG-4 Retroperitoneal Fibrosis: A Deceiving Disease

L. Diab, R. Singh, K. Bhairavarasu
Baylor College of Medicine, Houston, TX

Learning Objectives: Review a case of a delayed diagnosis in the setting anchoring bias and premature closure Review the utility of a cognitive autopsy in mitigating cognitive biases Discuss the features of IgG-4 related disease.

Case Information: A 41-year-old man with a reported history of a pancreatic and retroperitoneal (RP) mass status post chemotherapy and radiation twelve years prior presented with three weeks of worsening back pain. CT scan of his abdomen/pelvis revealed a 9.5cm RP mass encasing his aorta and inferior vena cava with bilateral hydronephrosis. Mass biopsy showed fibro-inflammatory spindle cell proliferation, favoring sarcoma. Although the pathology and staining was equivocal, his history suggested possible malignancy. Three months later, he returned with worsening symptoms and was found to have interval growth of the mass with increased lymphadenopathy. Given progression and equivocal biopsy, repeat biopsy was pursued with additional special studies to evaluate non-malignant processes. Immunohistopathology showed areas of IgG4 positive plasma cells coupled with elevated serum IgG4. He was initiated on prednisone given concern for IgG4-related (IgG4RD) RP fibrosis (RPF) and had significant improvement and interval decrease in mass size.

Discussion: Our case highlights a delayed diagnosis due to diagnostic errors that resulted in late therapy and progression of disease. A diagnostic autopsy identified anchoring, premature closure, and limitations of knowledge base as factors that impacted delay in diagnosis. Biopsy is the gold standard for diagnosing RPF etiology. Having an idea of the prevalence and pretest probability for potential diagnoses is ideal since biopsies can be subjected to sampling error. This patient’s oncological history anchored the team to favor malignancy as the etiology, despite equivocal biopsy.
Malignancy, however, only accounts for 10% of RPF, while idiopathic conditions account for 50-70% of RPF cases, of which IgG4RD accounts for almost half. Since IgG4RD is an emerging, novel entity, lack of knowledge of this condition may have narrowed the diagnostic focus and contributed to premature closure. By understanding our cognitive processes, we can avoid diagnostic pitfalls which can lead to earlier diagnosis to help reduce morbidity.

**Clinical Care Pathways May Be A Source of Diagnostic Error: Saved by an ED Handoff**

J. Grubenhoff1, L. Anne Bakel2, E. Ficco3, L. Bajaj2
1University of Colorado School of Medicine, Aurora, CO
2University of Colorado, Aurora, CO
3Children’s Hospital Colorado, Aurora, CO

**Learning Objectives:**
1) Diagnosis-oriented clinical care pathways may constrain diagnostic reasoning limiting sufficient consideration of competing diagnostic possibilities. 2) Emergency department hand-offs provide opportunities to reassess the working diagnosis.

**Case Information:** A 16 y/o healthy female presented with 1 day of acute non-migratory right lower quadrant (RLQ) abdominal pain without fever, nausea, vomiting, diarrhea, anorexia or dysuria. Last menstrual period was 1 week prior. Vital signs were normal. Physical exam revealed focal RLQ tenderness with guarding but no rebound tenderness. Urine pregnancy test was negative; urinalysis did not suggest infection. WBC count was 6.9 x 10³ cell/microL with 30% neutrophils. RLQ ultrasound demonstrated “solid inhomogeneous mass in RLQ with central tubular structure possibly representing ruptured appendix.” Surgeon recommended serial exams overnight. At hand-off, the departing emergency physician relayed a plan for oral fluid challenge if patient had improved. Tenderness resolved with ketorolac on repeat exam and surgeon recommended discharge. The arriving emergency physician obtained an abdomen/pelvis CT to further investigate the mass which revealed the appendix terminating at the site of an ovarian teratoma. The tumor was resected by OB/GYN 1 week later.

**Discussion:** Clinical risk scores (e.g. Alvarado score) incorporated in clinical pathways aim to use evidence to reduce practice variability and optimize resource use. This patient’s Alvarado score (2/10) placed her at low risk; an RLQ ultrasound to evaluate for appendicitis was not indicated. The initial team, however, reframed their thinking from “What is causing this girl’s pain?” to “Is this an appendicitis?” The progression of illness did not fit appendicitis, so the latter question was answered ‘no’ and the search for the etiology almost ceased. A fresh perspective at hand-off permitted reassessment of the diagnostic process and pursuit of alternative explanations. Our institution is beginning to incorporate “diagnostic timeouts” and pathway “offramps” into our clinical care pathways program to reduce future instances.

**A Case of Delayed Diagnosis of Cervical Vertebral Osteomyelitis Mimicking Crowned Dens Syndrome**

Y. Hasegawa, S. Kosugi, T. Matono, R. Yamate, D. Mawatari, T. Hara, H. Imura
Iizuka Hospital, Japan

**Learning Objectives:** In acute neck pain, it is important to remain aware of symptoms to identify abnormalities in case presentation and reconsider initial diagnosis, particularly based on System 1 decision-making.

**Case Information:** An 88-year-old man with chronic kidney disease presented with a 2-day history of posterior neck pain upon movement, particularly during rotation, without fever. On admission, blood test showed a C-reactive protein level of 115 mg/L, and cervical computed tomography scan revealed calcification around the odontoid process. He was initially diagnosed with crowned dens syndrome. Although we initiated nonsteroidal anti-inflammatory drugs, his neck pain exacerbated. Two days after admission, Enterococcus faecalis was detected in two sets of blood cultures. As persistent bacteremia was revealed, we conducted a transthoracic echocardiogram and found a vegetation measuring 5.2×2.3 mm on the aortic valve. Four days after admission, cervical magnetic resonance imaging scan showed vertebral osteomyelitis
in the 5th-6th cervical spine. The patient fully recovered from the endocarditis with vertebral osteomyelitis after a total of six weeks of intravenous ampicillin and ceftriaxone.

**Discussion:** Anchoring bias and premature closure caused diagnostic delay in this case. We believe that the rare disease diagnosed using System 1 and the unusual afebrile clinical presentation led to the bias. Unusual severe neck pain in a patient is one of the red flags, and blood culture is the key element to diagnose the infection. Additionally, it is vital to notice the uncommon clinical course after treatment initiation to reconsider the initial diagnosis and to arrive at an accurate diagnosis.

---

**A Case of Anchoring and Ascites**

C. Inchaustegu, R. Rana

Baylor College of Medicine, Houston, TX

**Objectives:** Exemplify the pitfalls of pattern recognition and anchoring.

**Case:** A 63 year-old homeless man with HCV and alcohol use disorder presented with chronic diarrhea, abdominal pain and ascites. 12 months ago, he developed watery diarrhea, abdominal pain and intermittent fevers. 4 months ago, he developed abdominal distention with ascites. He underwent paracentesis, was told he had cirrhosis and was discharged. 2 weeks ago, his ascites recurred prompting repeat paracentesis. He appeared cachectic with a distended abdomen but lacked cirrhotic stigmata. Cultures exhibited polymicrobial growth including anaerobes. CT abdomen revealed pneumoperitoneum prompting laparotomy, which discovered no gross perforation, a thickened peritoneum and non-cirrhotic liver. His working diagnosis was gas-forming anaerobic spontaneous bacterial peritonitis (SBP) and decompensated cirrhosis. His diarrhea persisted, and fecal calprotectin was markedly elevated. HIV serologies were negative.

**Discussion:** Gas-forming SBP was a problematic diagnosis as it is rare without appropriate risk factors. This patient’s “cirrhotic” label -despite a lack of stigmata and non-cirrhotic liver appearance- likely facilitated anchoring. Pattern recognition would dictate that ascites in a patient with HCV and alcohol use disorder may originate from portal
hypertension from cirrhosis. Review of prior ascitic fluid studies revealed a lymphocyte-predominant exudate with serum-ascites albumin gradient <1.1, inconsistent with portal hypertension. His diarrhea was also unexplained by SBP. If not from SBP, his pneumoperitoneum could be explained by perforated bowel, likely covered by omentum by the time of laparotomy. Reframing the case, our patient was a non-cirrhotic homeless man with chronic inflammatory diarrhea, lymphocyte-predominant low-SAAG ascites and thickened peritoneum with suspected bowel perforation. His differential now included peritoneal carcinomatosis, serosis in connective tissue diseases (e.g. IBD) or tuberculous peritonitis. CT chest revealed a right lung cavitary lesion, and subsequent biopsy revealed caseating granulomas with growth of Mycobacterium tuberculosis. He was diagnosed with disseminated tuberculosis and improved on RIPE therapy.

From Bagels to Beer: Autobrewery Syndrome as a Cause of Recurrent Pancreatitis

V. Jaspan, D. Kudlowitz
NYU School of Medicine, New York City, NY

Learning Objectives:
1) Develop an illness script for autobrewery syndrome
2) Review rare causes of pancreatitis

Case Presentation: This is a case of a 41-year old male with past medical history notable only for frequent sinus infections requiring antibiotic treatment, who presented with 9 months of recurrent episodes of acute pancreatitis. In the week prior to his first episode of pancreatitis, he consumed 4 beers, increased from his baseline of 1-2 drinks per month. Following discharge, he stopped all alcohol consumption. At home, he had several episodes of confusion, which he attributed to fatigue. He underwent laparoscopic cholecystectomy and testing for genetic and autoimmune causes of pancreatitis, which were non-revealing. He was hospitalized 10 more times during that 9-month period for acute pancreatitis with elevated transaminases. During 3 of these admissions, he had elevated triglycerides requiring an insulin drip. During one admission, he had an elevated alcohol level despite abstaining from alcohol for the prior 8 months. He underwent a glucose challenge in the hospital and a carbohydrate challenge as an outpatient, with measurement of alcohol levels, both of which were positive. This confirmed the diagnosis of autobrewery syndrome. Stool cultures were negative for yeast and fermenting bacteria. His alcohol levels became undetectable while on an empiric course of nystatin and low-carbohydrate diet.

Discussion: Autobrewery syndrome is a rare, but likely underdiagnosed, condition in which gut flora ferment glucose, producing ethanol. It most frequently presents with unexplained episodes of inebriation. It should be considered as a diagnosis in cases of recurrent pancreatitis and elevated transaminases after ruling out more common causes. Diagnosis can be established using a carbohydrate challenge. Empiric therapy with antifungals may be effective, even in the absence of positive stool cultures. More research is necessary to clarify the role of endogenous alcohol production in hypertriglyceridemia and elevated transaminases.

Renal Failure and the Old Foe

M. John, E. Sokolova
Icahn School of Medicine at the Mount Sinai Hospital, New York City, NY

Learning Objectives: Recognize active tuberculosis as a cause of secondary amyloidosis leading to renal failure.

Case Information: A 60-year-old South Asian man with no prior history presented with three days of fevers and cough. Initial exam was only notable for a fever of 38.4°C. Chest x-ray showed increased markings at the left base. Labs revealed BUN of 43 and creatinine of 7.7 (previously 1.4 one year ago). He received empiric CAP coverage, which resolved his presenting symptoms. His worsening creatinine prompted further workup. Urinalysis showed >1000 protein. Total urine protein was elevated. Serum protein electrophoresis, Bence jones protein, and urine immunofixation and kappa and lambda levels were unremarkable. Kidney biopsy revealed a new diagnosis of AA amyloidosis. CT imaging showed a 2.7cm mass posterior to the right mainstem bronchus, left lung peripheral nodular opacities, and mediastinal,
mesenteric, and retroperitoneal lymphadenopathy, read as characteristic of nodular parenchymal amyloidosis. Initial quantiferon gold was indeterminate. Hepatitis, HIV, RPR, and rheumatological workup were negative. Colonoscopy revealed gastric amyloidosis. A repeat quantiferon gold, obtained weeks after initial presentation, was positive. Sputum AFB testing revealed Mycobacterium tuberculosis complex. His final diagnosis was active Pulmonary Mycobacterium Tuberculosis causing AA amyloidosis and renal failure.

**Discussion:** Tuberculosis and other chronic infections were historically common causes of AA amyloidosis; prevalence decreased as access to effective treatment improved. In the United States, inflammatory arthritides and IBD are common causes of AA amyloidosis. In TB-endemic countries, however, tuberculosis remains a prevalent cause of AA amyloidosis, sometimes leading to renal failure. As this patient clinically improved with empiric CAP treatment, there was a delay to rule-out tuberculosis, though he previously lived in a TB-endemic country. While this is an unusual case, practitioners should be aware that untreated active tuberculosis can present with non-specific symptoms and can lead to secondary amyloidosis and eventual renal failure.

**Be Aware of Retro-Orbital Pain and Double Vision; Delayed Diagnosis of Subarachnoid Hemorrhage**

Tokyo Nishi Tokushukai Hospital, Japan

**Learning Objectives:** A retro-orbital pain and double vision can be a presentation of subarachnoid hemorrhage. Once you suspect subarachnoid hemorrhage, you should rule it out thoroughly even when brain computed tomography is normal.

**Case Information:** An 80-year-old woman presented to our hospital with complaints of two-day history of left retroorbital pain and double vision. On examination, she was alert, afebrile, and well oriented, with stable vital signs. Cranial nerve examination revealed that the left eye could not move laterally but she had normal pupil size and light reflex. Other neurological examination was unremarkable. Considering the possibility of intracranial abnormalities such as subarachnoid hemorrhage (SAH), unruptured intracranial aneurysm, and pituitary apoplexy, the emergency physician performed emergent radiological evaluation. However, there was no evidence of hemorrhage or an obvious aneurysm in noncontrasted computed tomography (CT) and magnetic resonance imaging of the head. The emergency physician referred the patient to an optometrist, but he denied an acute attack of angle-closure glaucoma. Although she visited a neurosurgeon and a neurologist on the next morning, she went home without any treatment because both of them determined that she did not have emergent problems and continuous follow-up was needed. However, she returned to the hospital that night because of unconsciousness. Emergent brain CT revealed SAH. She was intubated and transferred to another hospital for further treatment.

**Discussion:** A subarachnoid hemorrhage is a medical emergency. However, the diagnosis of SAH can be challenging because its initial presentation may be limited to common symptoms and subtle signs. The rate of missed or delayed diagnosis has been reported to vary from 12% to 51%. Although it is not sure when her bleeding occurred, her symptoms were associated with SAH taking into account her clinical course. Head CT has a high sensitivity but not perfect predictive value. We should consider other approaches such as a lumbar puncture or hospital observation with high suspicion of SAH, but we did not and missed the life-threatening status. There were some diagnostic biases such as rule bias and premature closure. Even if your patients show mild signs and symptoms or normal head CT, if he or she shows a possible presentation of SAH, physicians need to rule it out thoroughly.
Mary-ed to Anchoring Bias, Delayed Diagnosis of a Sister Mary Joseph’s Nodule

T. Kim, D. McCurry
University of Pittsburgh Medical Center, Pittsburgh, PA

Learning Objectives: Recognize the Clinical Features of Sister Mary Joseph’s Nodule and How Cognitive Biases Lead to Delayed Diagnosis of Malignancy

Case Information: A 76-year-old gentleman with a history of thyroid and prostate cancer was found to have an asymptomatic erythematous umbilical nodule. He was initially given bacitracin ointment. Over subsequent visits in the intervening months, he was prescribed doxycycline, clotrimazole/betamethasone, and several silver nitrate and hydrogen peroxide formulations for cellulitis, fungal infection, and warts, respectively, without improvement. He also reported abdominal pain, malaise, and 30-lbs weight loss during this time. During a subsequent hospitalization for new severe emesis, a Sister Mary Joseph’s Nodule was identified. Imaging demonstrated peritoneal carcinomatosis, metastatic umbilical nodule, and the diagnosis of end-stage pancreatic adenocarcinoma was made.

Discussion: Cancer misdiagnosis is among the most frequent diagnostic errors in primary care based on studies of ambulatory malpractice claims. Failure to obtain appropriate diagnostic studies is a frequently cited preventable factor. On the patient’s initial presentation, imaging was reasonably deferred for a presumed mild skin infection. However, the lack of clinical response to topical therapy, antibiotics, and antifungals did not lead to diagnostic evaluation of other clinical etiologies. We hypothesize that anchoring to the diagnosis of an uncomplicated skin lesion, in spite of discordant data, led to diagnostic delay and failure to consider alternative diagnoses. Given the rarity of SMJN in all intra-abdominal or pelvic malignancies, availability bias was also a likely contributor.

A Rheumatic Schematic

E. Leven1, A. Linker2
1Icahn School of Medicine at Mount Sinai, New York City, NY
2Mount Sinai Hospital, New York City, NY

Learning Objectives: 1. Evaluate factors contributing to delayed diagnosis of an uncommon illness. 2. Differentiate among causes of polyarticular joint pain in a young patient from a foreign country.

Case Information: A 36 year-old man with recurrent migratory joint pain on chronic prednisone presented with palpitations and dyspnea. He was found to have rapid atrial fibrillation with volume overload. Transthoracic echocardiogram showed findings specific for rheumatic heart disease (RHD) with severe mitral regurgitation, mitral stenosis and decompensated heart failure. Review of the patient’s history revealed multiple prior evaluations for low-grade fever and joint pain, initially diagnosed as atypical polymyalgia rheumatica and treated with glucocorticoids. His pain recurred years later, accompanied by hematuria. Second evaluation showed elevated ESR and antistreptolysin O titers; however, he maintained the diagnosis of “nonspecific inflammatory arthritis” and resumed steroids. There was no documented cardiac review of systems or discussion of RHD screening. Further history revealed that the patient grew up in Eastern Europe, experienced “frequent pneumonias” in childhood, and had young children. He was diagnosed with RHD with migratory polyarthralgia and severe valvular dysfunction. Valve surgery is planned.

Discussion: This patient was misdiagnosed with primary inflammatory arthritis, though his arthralgias were likely related to recurrent rheumatic fever. The arthritis associated with acute rheumatic fever is time-limited and salicylate-responsive, though case series show that symptoms can linger after the acute period. The history of febrile illness, arthritis and hematuria should have suggested a diagnosis of past acute rheumatic fever, which would have prompted earlier evaluation and intervention for RHD to improve his cardiac outcome and avoided the risks associated with chronic glucocorticoids. Providers training in developed countries have little exposure to RHD but this disease remains prevalent globally. This patient’s atypical presentation of a disease that is uncommon in the United States increased the risk of misdiagnosis.
The Role of Autopsy in Diagnostic Error

A. McHenry, D. McGuone, V. Parkash
Yale University School of Medicine, New Haven, CT

Learning Objectives: Demonstrate the role of autopsy in understanding diagnostic error

Case Information: A 51-year-old non-verbal, mentally disabled, group-home resident presented with abdominal discomfort with several-days of obstipation. He had history of pica, seizure disorder and remote abdominal surgery. Abdominal CT showed fecal impaction with proximal air-fluid levels, collapsed small bowel and hiatal hernia with distended stomach. He was started on antimicrobials for possible infectious colitis and admitted for “bowel cleanse”. Within hours of admission, he was found pulseless and apneic with copious vomitus. A 21 x 10 cm fabric aggregate obstructing the left colon and a 4.5 x 1 cm one obstructing the pylorus, with massive pulmonary aspiration of gastric contents, secondary to obstruction from impaction consequent upon pica from mental disability was the cause of death.

Discussion: This case demonstrates the role of autopsy in the understanding of diagnostic error. Note-bloat, “cut and paste”, and undue reliance on the completeness of the medical record led to the oversight of the patient’s known history of surgery for obstruction from fabric ingestion. Observer bias led to a presumptive diagnosis of fecal impaction. That imaging cannot discriminate between food and fabric created further confirmation bias. The patients mental and psychological incompetency coupled with bias against the mentally disabled contributed to inadequate monitoring of the patient in the home and emergency room. This case led to a change in the group homes policies where stool examination was added to the routine monitoring of patients with known history of pica.

Unique Risks for Diagnostic Error in Long-Term Care Settings

J. Pederson
Stratis Health, Bloomington, MN

Long-term care (LTC) which includes skilled nursing facilities, assisted living and other housing with services are unique sites for health care delivery as the care is provided in the individuals residence. Clinicians often come onsite rather than seeing patients in more traditional health care settings. Much of the literature regarding error in long-term care involve medication administration errors, however the opportunity for error in diagnosis exists due to workflows, differences in electronic health record documentation, staffing and other factors. The goal of this presentation is to create discussion using short case examples that highlight potential sources of diagnostic error in order to begin describing and classifying
error risk areas. The ability to identify and classify error could serve to raise awareness among LTC clinicians and caregivers. Ideally, this could lead to more study of how to improve diagnostic safety in LTC settings.

**Many Stigma: One Missed Diagnosis**

M. Qureshi, J. Swinden  
Geisinger Medical Center, Danville, PA

We present a case of a 45-year-old female with morbid obesity, obstructive sleep apnea, hypertension, generalized anxiety, depression, and post-traumatic stress disorder who presented to an emergency department in February of 2020 with abdominal distension, lower extremity swelling, shortness of breath, orthopnea & weight gain. An “elevated troponin” prompted non-urgent outpatient cardiology referral given the absence of chest pain. She was diagnosed with constipation and discharged. She returned to the emergency department with an unchanged constellation of symptoms which on this encounter were attributed to menorrhagia. She was again discharged home. Subsequent evaluation by 2 outpatient providers for the same clinical syndrome resulted in diagnosis of abdominal panniculitis which was treated with 3 courses of different oral antibiotics with no improvement. In April of 2020, two months after initial presentation and in the midst of the global coronavirus pandemic, she was admitted to our facility with overt clinical heart failure and 74 pounds above baseline weight from 3 months prior. Echocardiogram showed biventricular heart failure and ejection fraction of 35%. Coronary angiogram showed no obstructive coronary disease. Diagnosis of non-ischemic cardiomyopathy was made. She had rapid symptomatic improvement and return to baseline weight with intravenous diuresis. Despite the textbook presentation of heart failure, the diagnosis was repeatedly missed. The existence of implicit bias and stigma against the obese, females, and patients with mental health disorders is well documented though in this case these factors were likely further complicated by the potential impact of an unprecedented global pandemic. This case illustrates the need to pause, digest the available information and reflect on one’s own potential implicit bias before continuing on a path of diagnostic certainty, especially in the face of data contradictory to the working clinical hypothesis.

**Implications of a Missed Vancomycin and Piperacillin-Tazobactam Related Acute Kidney Injury**

K. Reid  
Cleveland Clinic Martin Health, Cleveland, OH

**Case Information:** A 32-year-old male with history of intravenous drug use presented with left thumb redness, swelling and pain after puncture with a metal cable. Incision and drainage was performed and vancomycin and piperacillin-tazobactam was started. Creatinine on day 2 of admission was 1.32mg/dL from baseline 1.0mg/dL and vancomycin trough was 15.7mcg/dL. On day 5, when his creatinine was next ed, it was 6.89mg/dL and vancomycin trough was 80.6mcg/dL. Vancomycin and piperacillin-tazobactam were immediately discontinued and hemodialysis commenced. Cystatin C was ed to help predict renal sequelae, this was 2.8mg/L. Unfortunately, his clinical course is unknown as he eloped.  
**Discussion:** This case illustrates a patient safety issue from a missed diagnosis of vancomycin nephrotoxicity and AKI in a patient without prior kidney disease. Nephrotoxicity can occur within four days after receiving vancomycin. The risk is increased with elevated trough levels. Additionally, a three-fold increased risk of AKI exists with vancomycin and piperacillin-tazobactam combination and the number needed to harm is 11. The rise in creatinine and vancomycin level on day 2 should have prompted closer monitoring. It is imperative to consider the implications of using this combination including kidney failure, increased health care costs and prolong hospitalization. Exercise caution in using both antibiotics. If ordering both antibiotics, consider monitoring vancomycin trough and renal function more closely. This may contradict high value care efforts, but infrequent monitoring has implications for delayed AKI detection.
Myocardial Ischemia and Gastrointestinal Bleeding: A Diagnostic Dilemma

M. Shukla, R. Xiao, R. El-Serag
Baylor College of Medicine, Houston, TX

Learning Objectives: Gastrointestinal bleeding can cause myocardial infarction due to hypovolemia and reduced myocardial perfusion. Management for such patients with suspected myocardial ischemia can be difficult as the treatment for a possible acute coronary syndrome can exacerbate gastrointestinal bleeding. Understanding the underlying pathophysiology of the elevated troponin I is crucial for appropriate treatment.

Case Summary: 74-year-old male with a past medical history of multi-vessel coronary artery disease who presented with acute onset chest pain with activity, requiring sublingual nitroglycerin. The patient was admitted for monitoring of high-risk chest pain. Of note, he was not started on treatment-dose anticoagulation as his presenting symptoms seemed more consistent with stable angina. His serial troponins were negative and his EKG did not show new ischemic changes. On hospital day three, the patient had a bloody bowel movement. CT Abdomen and Pelvis showed findings concerning for ischemic colitis, however his hemoglobin was stable, and no intervention was taken. On hospital day four, the patient had an episode of acute chest pain; subsequent troponin was elevated to 0.22 while an EKG showed non-contiguous mild ST changes for which he was started on a heparin drip due to concern for NSTEMI. Later that day, the patient had two more bloody bowel movements and was eventually transferred to the Coronary Care Unit for further monitoring and potential cardiac intervention.

Discussion: The care of a patient with suspected acute coronary syndrome in the presence of a gastrointestinal bleed can be complicated, as treatment of the former may exacerbate the latter. In this patient, his underlying coronary artery disease led to a vicious cycle of hypoperfusion, gastrointestinal bleeding, demand ischemia, initiation of anticoagulation and worsening gastrointestinal bleeding. This case demonstrates the importance of balancing aggressive empiric treatment, understanding comorbid patient factors and ultimately treating the underlying exacerbating factor.

The Role of Clinical Laboratory Data in Multi-Disciplinary Healthcare Team Communication and Education Strategies

E. Tiffany, S. Grant
COLA Inc., Columbia, MD

Learning Objectives:
(1) To demonstrate the importance of healthcare team interdisciplinary communication and education strategies to prevent diagnostic testing error and to support the integration of patient, physician, nurse, and laboratory professional information
(2) To provide a general list of concepts and potential solutions to address common diagnostic errors in patient care

Case Information: This clinical vignette presents a potential diagnostic laboratory testing error in a hospital emergency department. The error demonstrates a lack of communication, knowledge, and documentation of the patient’s herbal supplement use with a known negative side effect when taken with a commonly prescribed medication. A patient with a history of congestive heart failure presents to the emergency department (ED) with a false reading of a high level of digoxin in the blood due to the patient’s voluntary use of a common Chinese herbal remedy for an acute sore throat episode. This was not communicated to nurses/clinicians of the patient and was not added to the patient’s ED medical record, creating unaddressed patient harm.
**Discussion:** There are communication challenges and a lack of understanding between clinicians, nurses, patients, and laboratory professionals in various healthcare settings. Improving the quality of patient care can have a positive impact on diagnostic accuracy and patient outcomes. The poster proposes high-level solutions for physicians, nurses, laboratory professionals, and other clinicians in a patient care team to apply to their healthcare system dynamic. It lists general solutions and supporting references that may be used to improve the understanding of the role of each healthcare discipline, to organize regular meetings to review current risk management topics and resolve actual and potential patient diagnostic errors, and to improve healthcare team communications and educational activities between the disciplines.

**Tracheomalacia: A Difficulty to Diagnose Cause of Shortness of Breath**

J. Welkie, N. Olsen  
Penn State College of Medicine, Hershey, PA

**Learning Objectives:** Understand common and uncommon causes of cough and difficulty breathing. Assess potential consequences resulting from delayed diagnosis of tracheomalacia. Distinguish when expiratory CT and bronchoscopy should be utilized in suspected tracheomalacia.

**Case Information:** Ms. F is a 72-year-old female with a history of psoriatic arthritis who in December developed cough and difficulty breathing. Her primary physician and pulmonologist performed an extensive yet inconclusive workup including chest CT and pulmonary function testing (PFT). In February, she was hospitalized for worsening difficulty breathing. During hospitalization, her methotrexate (MTX) was held and diagnostic testing including CT and V/Q scans were normal and a CT scan with expiratory images was interpreted as unremarkable. Subsequently, an outpatient sleep study revealed mild obstructive sleep apnea and an otolaryngology evaluation with a sinus CT scan was relatively unremarkable. Finally, after seven months, the pulmonologist and radiologist re-evaluated the expiratory CT study from February and concluded that Ms. F had tracheomalacia. The delayed diagnosis resulted in unnecessary diagnostic tests, multiple rounds of antibiotics and corticosteroids, radiation exposure, and interruption of MTX, resulting in hastening of her disease process, and excessive costs to the patient and healthcare system.

**Discussion:** Tracheomalacia is an underrecognized condition characterized by softening of the tracheal cartilage resulting in dynamic airway collapse. Given the non-specific symptoms, tracheomalacia is often misdiagnosed as asthma, COPD, or other cardiorespiratory conditions. As highlighted in this case, delayed diagnosis often results in unnecessary corticosteroid therapy which hastens the disease progression and has numerous undesirable side effects. Tracheomalacia is easily missed with routine lung imaging and PFTs and therefore, if suspected, patients should undergo bronchoscopy or expiratory CT. Although bronchoscopy is currently the gold standard for diagnosis, expiratory CT is less invasive, and therefore may be preferred. Expiratory CT has demonstrated high accuracy for diagnosing tracheomalacia, however sensitivity and specificity are not well established.

**Missed Metastases: The Importance of Reviewing CT Imaging**

E. Ye  
Baylor College of Medicine, Houston, TX

**Learning Objectives:** Recognize how anchoring bias can affect timely diagnosis and emphasize the importance of utilizing all radiology images.

**Case Information:** A 59 year old female presented with symptoms of nausea, poor oral intake, and unintended weight loss. Workup revealed abnormal liver enzymes and a new liver mass. CT chest imaging was also done for staging and showed a spiculated lung mass with multiple nodules. However, the CT chest was incorrectly attached to the abdominal CT, and no separate read was provided. The CT abdomen only showed atelectasis, and the lung abnormalities went unnoticed by all teams. The patient initially underwent endoscopy for biopsy, which was negative. A second biopsy by IR ultimately showed cholangiocarcinoma. During this time, patient became more hypoxic, and a CT PE was completed. It was then found that the initial CT chest and lung masses were missed. Because the lung mass appeared radiographically more like a primary lung cancer, lung biopsy was recommended. Bronchoscopy with biopsy ultimately showed metastatic cholangiocarcinoma.
**Discussion:** This case demonstrates how anchoring bias can impact a patient’s care. Although there was an unexpected technological error, the medical team also played a role in missing the metastases. The CT chest images were always available for review, but because initial workup suggested abnormal liver pathology, the team limited their focus to the liver. Because of this limitation, the patient ultimately underwent unnecessary biopsies and had a delay in her cancer treatment. The case demonstrates the importance of physically reviewing all imaging and recognizing the risks of anchoring on initial workup results.

![CT Chest Images](image1)

**COVID19 Induced Saddle Pulmonary Embolus and Bilateral Multifocal Pneumonia**

E. Yousef  
Cleveland Clinic Foundation, Cleveland, OH

**Learning Objectives:**

(1) Recognize clinical presentation of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (COVID 19) pneumonia and pulmonary embolism (PE)

(2) Describe chest CT pattern of COVID19 related PE and pneumonia

**Case Information and Discussion:** A previously healthy 54 year old male was brought to the ED by EMS after having a pre-syncopal episode and falling at home, his BP 70s/40s and hypoxic 83% on room air, tachycardic, tachypneic with unremarkable physical examination. Blood work up was normal except for mildly elevated lactate and markedly elevated D-dimer; 17,910 (<500 ng/mL FEU). CXR showed patchy bilateral opacities suggestive of multifocal pneumonia. Naso-pharyngeal PCR was positive for SARS-CoV-2. CT PE showed: Saddle pulmonary embolus extends into the right, left, lobar, segmental and subsegmental pulmonary arteries (PA). Enlarged right heart chambers and mild interventricular septal bowing are suggestive of right ventricular strain. Bilateral diffuse peripheral groundglass opacities, with lower lobes predominance. Patient underwent pulmonary angiogram with bilateral aspiration thrombectomy. Pre intervention, PA pressure (PAP) 44/13 mmHg (15-25/8-15 mmHg), with global hypoperfusion to bilateral lower, right middle and partially the right upper lobes. Post-intervention, PAP 27/8 mmHg, with contrast perfusion of bilateral lower and right middle lobes. PE associated with COVID-19 pneumonia is a life-threatening condition and its etiology remains unclear. High clinical suspicion is needed especially for patients presenting with hypoxemia and markedly elevated D.Dimer. CT signs were those of multifocal predominantly peripheral ground-glass opacities and filling defects in the entire PAs tree. Pulmonary angiography provided both diagnostic and therapeutic value.

![CT PE Images](image2)
Education

The Utility of the Patient Portal

R. Artrip\textsuperscript{1}, C. Dougherty\textsuperscript{1}, P. Milianta\textsuperscript{1}, S. Hardy\textsuperscript{2}, J. Beatty-Chadha\textsuperscript{2}, Jonelle Thomas\textsuperscript{2}

\textsuperscript{1}Penn State College of Medicine, Hershey, PA
\textsuperscript{2}Penn State Medical Center, Hershey, PA

Many hospital systems are rapidly developing patient portals. The Penn State Hershey Medical Center (HMC) established its portal years ago, with radiology reports accessible since October 2013. The online patient portal, which provides patients access to health information such as laboratory data, radiology, and pathology results, is an efficient way to deliver information to patients. It also serves as a platform that patients can use to schedule appointments and speak with physicians directly. We are interested in assessing how the patients at our institution use the patient portal and we aim to identify barriers to access and learn how patients prefer to receive their results and propose interventions to increase usability. To characterize the usage at our home institution, we sampled patients in the various radiology waiting areas in the Hershey Medical Center. Areas sampled included: Nuclear Medicine, General Radiology, and MRI waiting areas. Interviews were conducted with a 6-question survey designed by the third-year medical students to assess patient attitude toward and characterize usage of the patient portal. Demographics of portal usage were measured, and it was found that the mean age of portal users was 60 years, the median age was 65 years with a standard deviation of 14.2 years. 59\% of women surveyed showed that they use the portal compared to 37\% of men sampled. 36\% of patients prefer to use the portal to obtain their radiology results, 32\% prefer multiple modalities of obtaining results, most notably a combination of the portal and face-to-face conversations, 20\% of patients prefer a phone call, and 4\% of patients prefer mail. 91\% of patients felt the portal was useful stating it allowed them to schedule appointments online, message physicians, and forward images to non-affiliated providers. A majority (44\%) of patients first heard about the portal from the hospital, and a majority (85\%) said they would use the portal to view their radiology results. A bipolar distribution of patient users was identified, a group of heavy utilizers and non-utilizers. We hypothesized that we could increase the usability of the patient portal through workshops at HMC targeting patient populations that do not use the portal as identified by the survey, adopt automatic patient portal enrollment, informative marketing highlighting the benefits of the patient portal, and a customizable patient portal which would increase the usability and retention across age groups.

Variability in Measuring Competencies in Diagnostic Error: A Novel Digital Concept Mapping Tool

D. Chartash\textsuperscript{1}, M. Rosenman\textsuperscript{2}, T. Wijesekera\textsuperscript{3}

\textsuperscript{1}Yale University School of Medicine, New Haven, CT
\textsuperscript{2}Lurie Hospital for Children, Chicago, IL
\textsuperscript{3}Yale University, New Haven, CT

Focusing on three competencies from the Society to Improve Medicine’s Consensus Curriculum on Diagnostic Error (collecting key clinical findings; use the core clinical skills appropriately; explain and justify the prioritization of differential diagnosis), we developed a tool to perform three aligned tasks (information gathering; differential diagnosis; information integration) in order to assess the diagnostic competencies. The tool was intended to measure the appropriate selection of clinical findings, the use of meta-cognition, and a structural approach to causal logic. Four cases were selected, two each from adult and pediatric medicine. For each case, participants were asked to perform the following: highlight pertinent information within the case, connect each of those highlighted text spans in a directed network representative of clinical entailment, and create a differential including five ranked diagnoses. Assessing the information encoded in the clinical reasoning task, we compared information from each participant’s network using Bar-Ilan’s M, which computes the weighted similarity between two non-conjoint sets on the interval \([0,1]\). We analyzed the differential diagnosis lists created by students, which were normalized into Unified Medical Language System Concept Unique
Identifiers (CUIs). Figure 1 Comparing the median Bar-Ilan’s M across cases between students, we found greater agreement in Case IV (0.598) than Case I (0.575), Case II (0.241), and Case III (0.0575). For Cases IV and I, we see almost uniform identification of the first ranked diagnosis (Obstructive Sleep Apnea and Teratoma respectively), while for Cases II and III there was much more variability. In Case IV, there was also less variability observed of the second, third and fourth-ranked diagnosis compared to in Case I or III. In our qualitative evaluation of the students’ information integration, we observed higher levels of perceived difficulty when students were selecting a first-ranked diagnosis in cases I and IV than in cases II and III. Furthermore, it was clear that premature closure occurred in Case III (which resulted in most students’ ranking a diagnosis of gastroenteritis, rather than pneumonia, first). Ultimately, this tool allows for the measurement of student diagnostic performance consistent with the consensus competencies.

Cases:
(1) 10.1016/j.amjmed.2017.10.007
(2) 10.1016/j.amjmed.2016.04.005

Cultivating Participation through Competition: An Innovation for Clinical Pathological Conferences

J. Chun1, A. Parikh2, K. Luoma3, M. LoCurcio3
1Stanford University, Stanford, CA
2Vanderbilt University, Nashville, TN
3New York University, New York, NY

Purpose/Problem: The Clinical Pathological Conference (CPC), illustrates a clinical vignette and the logical reasoning process to establish a reasonable differential diagnosis. At New York University’s Internal Medicine residency we ask the entire community of learners to submit their most likely diagnosis and their reasoning. The conference itself attended with great interest by much of the house staff and faculty, who usually fill the room to capacity. However, this stands in contrast to the number of diagnoses submitted by those same groups. Predominant theories of learning suggest
improvements in recall when learners commit to their own answer and are provided with corrective feedback. We sought to stimulate such commitment, and participation through the creation of low stakes competition.

**Description of the Program:** Participants are asked to submit their unifying diagnosis, next diagnostic step, and diagnostic reasoning through an online portal. For the purposes of this competition a chief resident would review the responses and award points to each of the four training sites for attending submissions and to the resident group. These point totals were broadcast with the teaching point highlights from the conference. In subsequent years, a tournament style trophy was awarded to the group with the most points.

**Outcomes:** Prior to initiation of this competition faculty and resident responses averaged 1-2 per month. After initiation, the number of responses increased to around 23 per month. Some narrative comments from each of the sites also demonstrate some of the intangible changes that were fostered by the competition. Most of these comments could be summarized as the competition brought people together for greater discussion of the case. Some selected examples include “The level of faculty engagement in our monthly CPC is astounding. We have developed another layer of camaraderie around the idea of learning and deciphering a complicated case together.” Also “While the initiation of the competition was the genesis of broader faculty engagement in the CPC, the richness of the discussions during our faculty meetings is what kept faculty engaged and always eager to participate”. And finally, “They have served both as sources of friendly competition and as springboards for thoughtful discussion”.

**Discussion:** The addition of low stakes competition increases educational buy in with minimal added stress. Newly engaged participants spontaneously developed discussion groups which in and of themselves were major drivers of engagement. Other programs who already run CPCs may benefit from such a change to their conferences.

---

**Establishing the Educational Foundation for Admission to Medical School**

M. Gusack  
MANX Enterprises, Ltd., Huntington, WV

**Purpose/Problem:** The present academic requirements for admission to medical school are decades old. Yet, during this period, the growth in diagnostic categories, acceleration in the complexity of the diagnostic process, and the growth in medical liability has placed an enormous responsibility on the shoulders of the prospective clinician. The result? Ever more Diagnostic Error in Medicine.

**Description of Program, Assessment, or Study:** Proposed is a comprehensive, integrated premedical program that brings together critical knowledge, critical thinking, and critical problem-solving methods to lay the foundation for future work in the knowledge-based world of medicine. The more specific purpose includes but is not limited to:  
• Preparing the prospective clinician for the academic and real-world challenges of medical school  
• Assessing the effectiveness of the subjects chosen and how they are presented to the student  
• Deferring applicants who, despite high grades and standardized test results, are not suited to be physicians.

**Outcomes:** During the course teaching at the premedical school, medical school, and post medical school level over a forty-year period, I have had the opportunity to observe the inadequacy of preparation of the prospective clinician to achieve a high degree of diagnostic acumen. This experience has exposed numerous serious deficiencies including but not limited to the absence of exposure to:  
• Reading comprehension  
• Logic and epistemology  
• Mathematics, especially statistics and epidemiology  
• Human history, and psychology  
• Advanced principles of biochemistry and physiology  
• Information technology and artificial intelligence.

**Discussion:** Personal efforts to alter the course of the premedical education has met with uniform resistance amplified by the demotion of teaching activities relative to research and sports that bring in large sums of money to the college or university. The result has been a growing inadequacy of properly educated and dedicated medical school applicants.

**Significance of Findings:** Premedical education does not prepare the medical student to:  
• Understand pathophysiology and the relationship of clinical findings to diagnostic categories  
• Read complex scientific papers critically to determine their significance or lack there of  
• Determine if the statistical methodology employed and the way in which it is applied to support a published scientific paper is appropriate  
• Apply valid scientific knowledge and underlying principles in problem identification and problem solving
Redesigning the Medical School Experience to Reduce Diagnostic Error in Medicine

M. Gusack
MANX Enterprises, Ltd., Huntington, WV

Purpose/Problem: Today medical school education does not vary significantly from the century old Johns Hopkins model. Yet, during this century, the field of diagnostic medicine has been transformed by the intersection of scientific, technical, and managerial knowledge that impacts everything the clinician does. The result? Ever more Diagnostic Error in Medicine.

Description of Program, Assessment, or Study: Proposed is a comprehensive, integrated medical ‘educational’ program that applies part of the cognitive cascade: Priming: Exposing the medical student to the technical aspects of healthcare using a preceptor system Training: Developing a high degree of competency in history and physical taking as well as patient management Educating: Building upon the Priming and Training experience to lay down a solid knowledge foundation Experiencing: Then, returning the student to the diagnostic ring under appropriate supervision

The more specific purpose includes but is not limited to: • Eliminating rote learning that is forgotten due to the absence of priming • Using priming and training to establishing the significance of what is being taught leading to better retention • Reverse the traditional educational approach from the disease-to-finding model to finding-to-diseases model.

Outcomes: I have had the opportunity to teach at five different medical schools over a thirty-five-year period, including acting as an academic coordinator. During this time, I have observed and participated in the design and implementation of medical education programs. This experience has revealed a uniformity in approach to teaching that does not properly prepare the medical student to be an effective diagnostician or problem solver under clinical conditions. Furthermore, research in the field of education, training, and competency reveal a dearth of solutions and few experiments in medical school education that vary significantly from the century old Johns Hopkins model.

Discussion: Personal efforts to alter the course of the medical education have been blocked due to the focus on research and service work that bring in large sums of money to the medical school and university. The result has been a growing inadequacy of properly educated and dedicated physicians entering their residency training.

Significance of Findings: Medical students are: • Forced to waste time memorizing questionable knowledge out of clinical context that is mostly forgotten • Expected to become full-fledged diagnosticians by tagging along with residents after two wasted years • Confused and demoralized when faced with a cluster of ill-defined clinical findings instead of a textbook case.

Redesigning Post Graduate Training to Reduce Diagnostic Error in Medicine

M. Gusack
MANX Enterprises, Ltd., Huntington, WV

Purpose/Problem: Today residency programs do not vary significantly from the century old Johns Hopkins model. Yet, over the last century, the field of diagnostic medicine has been transformed by the intersection of scientific, technical, and managerial knowledge. In addition, the growing threat of liability and the focus on hospital length of stay has reduced the amount of experience a resident can get. The result? Ever more Diagnostic Error in Medicine.

Description of Program, Assessment, or Study: Proposed is a comprehensive, integrated medical ‘training’ program that continues the cognitive cascade of: Priming: Exposing the fledgling resident to the operational aspects of healthcare using a preceptor system Training: Further developing competency in history and physical taking as well as patient management Educating: Building upon the Priming and Training experience to lay down good diagnostic habits Experiencing: Increasing the opportunity to do everything needed to be a competent physician before graduating Judgement: Developing the capacity to learn from their and other’s mistakes to achieve a high degree of self-improvement.

Outcomes: The experience of overseeing residents over a thirty-five-year period has provided an opportunity to observe and participate in the design and implementation of residency training programs. The sameness in approach to training I
have found does not properly prepare the resident to be an effective diagnostician nor a professional capable of both improving the diagnostic process and passing on their experience and judgement to those coming up from behind.

**Discussion:** Personal efforts to alter the course of the residency training have been hampered due to the focus on service work that bring in large sums of money to the medical school and university, the threat of malpractice suits, as well as the refusal to invest in residents who may soon go elsewhere to practice medicine. The result has been a growing inadequacy of professionally trained and dedicated physicians entering medical practice.

**Significance of Findings:** Residents: • Have no adequate diagnostic model to learn and master due to the absence of appropriate preceptors • Cannot gain adequate experience in diagnosis due to lack of adequate time to read and be thoughtful • Are not expected to think critically and challenge what they hear and read. Just the opposite.

**Diagnostic Reasoning Competencies for Nurse Practitioner Education Programs: A Process for Consensus Building and Validation**

S. Smith
University of Minnesota, Minneapolis, MN

**Purpose:** The purpose of this initiative is to activate development, confirm, and achieve support for diagnostic reasoning competencies in Nurse Practitioner (NP) education. While clinical reasoning in general and some components of diagnostic reasoning are included in NP education programs, intentional teaching of diagnostic reasoning as a specialized competency is less frequently included. As a result, NP educator consensus around competencies for diagnostic reasoning has not yet been achieved. The aim of this initiative is to address that gap by: 1) proposing diagnostic reasoning competencies for NP education programs; and 2) developing a process for validating the competencies.

**Description of the Program:** To meet the need for more robust diagnostic reasoning curriculum in NP education programs, an initial set of NP diagnostic reasoning competencies has been proposed. The proposed competencies, derived from the SIDM consensus curriculum individual competencies (Olson, et al., 2019), focus on key components of NP practice and required diagnostic reasoning skills: collecting and interpreting clinical findings; formulating and expressing problem representations; producing, prioritizing, and justifying the differential; appropriate use of clinical guidelines and information tools; and use of metacognition to improve diagnostic performance. A combined focus group-Delphi-Q sort method of consensus development by expert NP faculty and practitioners will be used to refine, build consensus for, and validate the competencies.

**Outcomes:** An integrative review will be conducted of the scientific and educational literature on diagnostic reasoning in NP education and practice. A Delphi/focus group of expert NP faculty and practitioners will be established to review the resulting literature. Abstracts will be analyzed for key diagnostic reasoning knowledge, skills, and attitudes relevant to NP education and practice. Potential competencies will be categorized and compared to the proposed NP education draft competencies and those of the SIDM consensus curriculum. A Q-sort process will be used to identify highest ranked competencies. Re-evaluation for definitional clarity and specificity will be completed by a small group of reviewers and sent to the full group for finalization and approval.

**Discussion:** The proposed process for establishing diagnostic reasoning competencies for NP education programs builds upon and mirrors the approach used by SIDM/Macy in building the diagnostic reasoning consensus curriculum. The expected outcome of the project is the development of consensus around diagnostic reasoning competencies for graduates of NP education programs.

**Significance of Findings:** Resulting competencies will be submitted to NP education regulatory organizations for inclusion in curricular requirements for NP education programs.
Practice Improvement

Rapid Cycle Deliberate Practice Simulations to Decrease Variability in Care Delivery in Pediatric Sepsis

J. Corboy, K. Denicolo, E. Alpern, K. Mangold, M. Adler
Ann & Robert H. Lurie Children’s Hospital of Chicago, Chicago, IL

The processes of early recognition and treatment of pediatric sepsis patients are ever changing within our Emergency Department (ED). Decreasing variability and sustaining improvement in the practice of sepsis care delivery has been challenging and a major focus for meeting the timeliness goals for fluid and antibiotics of the Surviving Sepsis Campaign. To address knowledge gaps among staff and identify new barriers, a Rapid-Cycle Deliberate Practice (RCDP) simulation strategy was used to improve performance. RCDP focuses on rapid acquisition of skills through a practice-pause-teach-repeat model (Fig 1). Teams worked through the sepsis care algorithm until all steps were accomplished without error and within time goals. Prior to this intervention, barriers to efficient team performance were identified by staff during multiple sepsis huddles and a key driver diagram was developed to address these issues. A simulated sepsis patient scenario was developed for training purposes. Using the RCDP method, a disciplined approach to care delivery was introduced. The need for additional communication surrounding team expectations, continued gaps in knowledge surrounding the vital-sign based electronic sepsis alert and clarification of roles and responsibilities involved in sepsis care were highlighted by this exercise. Through simulated team-based training, variability in care was decreased and marked improvements were noted in staff performance from the initial to final simulation for all teams. In addition, information gained was used to inform further interventions and plan-do-study-act (PDSA) cycles, improving the process of care for pediatric sepsis patients. Conclusion/Implication Rapid cycle deliberate practice simulation was a novel approach in our emergency department. This environment engaged the multidisciplinary team members to learn from each other and improve team performance through repetition and shared, goal-directed objectives. By decreasing the variability of the process of care delivery, our team was able to improve treatment timeliness for both fluid and antibiotic delivery in pediatric sepsis patients.

Figure 1: Conceptual Model of Rapid Cycle Deliberate Practice training for sepsis care process improvement.
Development of eMeasures to Study Missed and Delayed Diagnosis of Lung and Colorectal Cancer

P. Kapadia¹, A. Zimolzak², D. Murphy², G. Lyrazopoulos³, G. Abel⁴, D. Upadhyay⁵, T. Scott², H. Singh³
¹Baylor College of Medicine, Houston, TX
²Michael E. DeBakey VAMC, Houston, TX
³University College London, London, England
⁴University of Exeter, Exeter, England
⁵Geisinger, Danville, PA

**Statement of Problem:** Accurate measurement of healthcare processes and patient outcomes is essential to improving quality. While many efforts have focused on measuring quality of treatment processes, measurement of diagnostic quality has lagged. Measures for diagnostic quality in cancer diagnosis are needed to recognize and prevent delayed and missed diagnoses.

**Description of the Program:** We will develop and validate algorithmically-calculated “e-measures” to quantify delays in cancer diagnosis. We are in the process of consulting an expert panel of physicians, administrators, researchers, and patients for input and will use the Safer Dx Trigger Tools Framework to develop e-measures. This framework has seven steps (outlined below) for constructing e-trigger algorithms, which will guide development of e-measures for cancer diagnosis applied to “testing sets” of sampled EHR data. 1. Identification and prioritization of diagnostic errors of interest 2. Operational definition of all the criteria for e-measure 3. Determination of potential data sources 4. Construction of e-measure algorithm 5. Testing of e-measure tools on the data sources 6. Assessment of e-measure tool performance 7. Iterative refinement of e-measure algorithm

**Findings to Date:** In collaboration with technical experts, we initially identified two clinical areas of interest (lung and colorectal cancer) and three preliminary e-measures: 1. Missed Red Flag e-measure: Process measure of missed opportunities to act on test results that may be early signs of cancer (e.g., lack of response to a positive stool test for colorectal cancer or follow-up to an incidentally found lung nodule). 2. Emergency Presentation e-measure: Outcome measure of the proportion of new cancers diagnosed through emergent presentations (e.g., colon cancers presenting with acute obstruction or perforation) rather than screening or non-emergent signs/symptoms. 3. Late Diagnosis e-measure: Outcome measure of the proportion of new cancers diagnosed at a late stage (3 or 4), as opposed to early-stage (1 or 2). We will now test these concepts in 4 unique data repositories in 4 disparate healthcare systems: Baylor College of Medicine, the US Department of Veterans Affairs, Geisinger, and the Clinical Practice Research Datalink (United Kingdom).

**Lessons Learned:** We have applied framework steps 1-5 to two of the measures in one health system as proof-of-concept. While standardizing e-measures across multiple data sources with different data elements is challenging, we are iteratively developing measure concepts for pragmatic application and use by many types of health systems. After application and pilot-testing, e-measures could be ready for wider implementation and evaluation in additional clinical settings.

Tele-Ultrasound: A Virtual Training, Supervision and Consultation

B. Mathews, O. Dickinson, J. Verner
HealthPartners, Bloomington, MN

**Statement of Problem:** Point-of-care ultrasound (POCUS), or bedside ultrasound performed by a clinician caring for the patient, is being used to support the diagnosis and serially monitor patients with COVID-19. There have been concerns about use of personal protective equipment, the ability to receive hands-on supervision and support for interpretations of ultrasound in this era. Some ultrasound manufacturers have added tele-ultrasound software that allows remote training of novice POCUS users and remote guidance in actual patient care. Tele-ultrasound can be utilized to share images in real time with consultants or expert providers.
**Description of the Intervention or Program:** 5 handheld ultrasound devices were obtained with tele-ultrasound features and were available for attending physicians with internal medicine residents and non-teaching services at Regions Hospital in St. Paul, MN. The program reviewed the use of tele ultrasound with remote scanning, training and supervision.

**Findings to Date:** These are preliminary findings in the COVID-19 era. We are gathering data on the frequency of use of an integrated ultrasound telemedicine platform in its use with trainees and with peers in a consultative role in our hospital. In the early data so far, 7 patients have been scanned with an ultrasound device in the hands of a novice learner while the faculty supervised remotely. 3 consultations were provided remotely for peers in a similar manner. The bedside ultrasound image can be remotely adjusted with an intuitive visual overlay the supports the training and supervision during the care of a patient.

**Lessons Learned:** Technology is key in improving diagnosis and a tool such as POCUS brings the provider back to the bedside to help in the diagnosis and management of a patient. The COVID-19 pandemic has led to concerns for the amount of people in a patient’s room, the use of PPE and limited hands-on training options for learners. The ability to use tele-ultrasound to remotely supervise and guide a learner or a peer is a key tool to provide virtual training and support for practitioners, safely guiding with a real-time scan at the bedside.

**Illness Script as a Tool for Improve Knowledge Organization in Medical Students**

P. Srichan, C. Chayangsu, P. Sangkla, P. Yodnopaklow
Medical Education Center Surin Hospital, Surin, Thailand

**Background:** Teaching clinical reasoning is a real challenge. Illness script was developed to improve clinical reasoning skills based on problem representation. The aim of this study is to evaluate perception of medical students to use illness script.

**Method:** Seventeen fourth year medical students were included in one session workshop using illness script. After that they practice with real patients and get feedback from preceptors. The questionnaire was used to evaluate the perception of students. Results All medical students agree that an illness script is useful to recall memory of classical clinical features and understanding diagnostic process. They agree that this tool can be applied to diagnose disease with the same and general characteristics 94% and 82.4% respectively. Eighty eight percent of them accept concept of illness script may help to organize knowledge and improve clinical reasoning.

**Discussion:** Illness script has some benefit to improve clinical problem solving. We conducted a workshop for novice medical students. Most of them can be process problem list, prioritize and contrast differential diagnostic list. Furthermore, they can apply this concept in clinical practice. We plan to implement this illness script in our clinical reasoning curriculum.

**Conclusion:** Illness script is a key to improve diagnostic accuracy and should be a tool in teaching clinical reasoning.

**Poster Session 2**

**Tuesday, October 20**
11:30 AM - 12:30 PM

**Scientific**

**Subadditivity and Overestimation of Diagnostic Probabilities, A Study of Undergraduate and Postgraduate Medical Students**

A. Almothaaffar, B. Yassin
College of Medicine /University of Baghdad, Baghdad, Iraq

**Background:** Doctors use diseases probabilities often during their clinical decision making, but in many instances the understanding of the concept of probability discloses misunderstanding. Identifying Subadditivity (i.e., the sum of probabilities concerning a single case scenario exceeding 100%) is a way that can uncover this problem.
Methods: This pilot study was conducted in the College of Medicine /University of Baghdad from January to April 2020. It included a group of undergraduate year 4 students (UG) and postgraduate haematology board trainees (PG). A clinical haematology case of bleeding tendency was constructed and distributed online; the participants were asked to estimate the probability of each component of the differential diagnosis. The sum of probabilities of each participant was recorded. Subadditivity was identified if the sum of probabilities concerning this case scenario exceeded 100%.

Results: The participants included 22 PG trainees (12 males and 10 females) and 37 UG students (25 females and 12 males). Regarding UG students, the range of sum of probabilities was 98-435%, with 26 (70%) showed subadditivity. While for the PG trainees, the range of sum of probabilities was 82-410%, with 9 of them (40%) showed subadditivity. Comparing both groups using chi-square showed statistically significant association between the graduation level and subadditivity (P value 0.026). When the 2 groups are summed up, 10 (40%) of males and 25 (71%) of females have subadditivity with statistically significant association (P value 0.022).

Conclusion: Substantial subadditivity is present in both UG medical students and PG medical trainees. Females are more likely to show subadditivity. Further study is needed for better understanding of this phenomenon that may affect decision making of doctors.

Experience, Time, and Testing. Impact of Three Variables on Diagnostic Accuracy in the Emergency Department

K. Barhaugh¹, K. Coleman¹, M. Jesso², S. Anderson²
¹Virginia Tech Carilion School of Medicine, Roanoke, VA
²Carilion Clinic, Roanoke, VA

Background: Diagnosis formulation is a longitudinal exercise. It is a trajectory defined by factors that include the presentation, an accurate history, clinician education, experience, available diagnostics and interpretation of tests. In Emergency Medicine (EM) a strategy is formulated using the most predictive testing to forward the inclusion or exclusion of life and limb threats. Clinician experience has been identified as an important factor in diagnosis. The goal of this study was to examine diagnostic accuracy and its relationship to the length of initial patient interaction, clinician experience, and diagnostic testing.

Methods: This was an observational study in the emergency department of a tertiary academic center. Emergency Medicine residents were evaluated on their assessment of patients with chest or abdominal pain. The residents were blinded to the fact that their initial evaluation was timed. After the initial evaluation, the EM resident was asked to provide 3 differential diagnoses and a leading “most likely” diagnosis prior to speaking with an attending physician. These were compared to the final admitting diagnosis or final discharge diagnosis for each encounter. The diagnostic studies ordered during the ED visit were also recorded. A total of 101 patient encounters were included. Five interactions were excluded.
Four were incomplete, and 1 did not have a chief complaint of either chest pain or abdominal pain. The encounters were collated based on the residents’ experience levels for analysis.

**Results:** The study demonstrated that post-graduate year one (PGY1) residents spend the greatest amount of time initially evaluating the patient compared to their PGY2 and PGY3 colleagues (11.6 minutes vs 7.58 minutes and 7.09 minutes respectively). The number of imaging and lab studies ordered was also greatest for PGY1 residents. Regarding diagnostic accuracy, the PGY3 group demonstrated greater accuracy when comparing their initial “most likely” diagnosis to the final EM diagnosis (33.90% vs 23.53% for PGY1 and 16.0% for PGY2). For admitted patients, PGY3 residents did not demonstrate greater diagnostic accuracy than PGY1 residents (42.8% vs 43.48%). Both groups were more accurate than PGY2 residents (20%).

**Conclusion:** Resident experience plays an important role in diagnostic accuracy. Legal risk, medical coding specificity, availability of diagnostic testing in the Emergency Department, and specialty practice patterns impact the decision making of medical learners. More study is required to further elucidate the factors impacting EM diagnostic accuracy.

**What Influences Agreement about Patients’ Diagnoses? - A Prospective Observational Study of Intensive Care Unit Teams**

P. Bergl, N. Shukla, J. Patel, R. Nanchal
Medical College of Wisconsin, Milwaukee, WI

**Aims:** The aim of this study was to establish how often clinicians on the ICU team agree on their patients’ primary diagnoses and to characterize how other variables, such as ICU length of stay and perceived patient complexity, correlated with the level of agreement.

**Methods:** This study was a prospective observational cohort study conducted in a medical ICU between June and August 2019. Subjects included attending physicians, fellows, residents, advanced practice providers (APPs), and nurses. For the first three days of their patients’ ICU admissions, subjects were asked about the patient’s primary diagnosis, confidence level, perceived complexity, clinician stress level, level of nurse involvement in the care plan, and perceived patient prognosis. Primary diagnoses were collected as open-ended responses, and all other variables were ordinal data on 5-point Likert scales. Each set of clinicians’ responses for a given patient on a given date (henceforth “patient-days”) were reviewed and level of agreement scored by two members of the investigation team. Descriptive statistical analyses were performed on both ordinal and dichotomized data (e.g. clinician’s confidence of “4 or 5” considered “high” while clinician’s confidence of “1, 2, or 3” considered “low”).

**Results:** We reviewed 480 unique patient-days and found good agreement about the primary diagnosis in 72% of patient-days. The two investigators’ ratings of team agreement were strongly correlated (rho = 0.747) and demonstrated modest agreement (Cohen’s kappa of 0.34 on ordinal data, kappa = 0.49 on dichotomized data). There was a modest and significant correlation for both APPs/residents’ and fellows’ level of confidence in diagnosis and average level of agreement (APP/resident: rho = 0.357 with p<0.0001, fellows: rho = 0.335 with p<0.0001). There was also a weak but significant correlation between the average level of agreement on the team and the resident/APP/fellow’s perceived likelihood of survival of the patient.

**Conclusions:** These preliminary results suggest that ICU clinicians generally agree about their patients’ leading diagnosis. The level of agreement is positively associated with clinicians’ confidence and the perceived prognosis of the patient. These findings may indicate that clinicians’ confidence in the diagnosis is a marker for communication between members of the team.
Understanding the Relationship between Patient Factors and Mental Health Clinical Decisions: A Mixed Methods Study

L. Burns, A. Silva, A. John
Swansea University, Swansea, Wales

Background: Patient factors such as gender can unknowingly impact mental health-related clinical decisions. Understanding the extent to which these non-clinical patient factors (NCpF) influence clinical decision-making is imperative to minimizing errors in diagnosis or treatment caused from internal biases. This research uses a combination of focus groups, mixed-methods experimental vignettes co-produced with clinical practitioners and population-level health data. It is one of few studies to use anonymised patient health records to understand the impact of NCpF on mental health-related diagnosis, treatment and referral decisions.

Methods: i) Firstly, service-users with mental ill-health were involved in two focus groups, conducted to identify any NCpF not previously found in the literature. ii) Secondly, Welsh Health Survey records were linked with healthcare records. Subsequent analysis aimed to identify patterns in diagnostic, treatment and referral decisions when compared to different NCpF of individuals with mental ill-health. iii) Finally, clinical vignettes displaying three mental illnesses were created and sent to general practitioners as the participants. The three cases depicted Bulimia Nervosa (BN), Anxiety Disorder, and Bipolar Disorder. Each disorder is paired with a NCpF, BN with “sex” for example. Participants were randomly allocated to variable 1 (male pronouns for BN), variable 2 (female pronouns for BN) or the control variable (gender neutral pronouns for BN), for each of the mental health illnesses, in order to see whether any biases were present in the clinical decisions made.

Results: The results from each of the three projects will be presented and discussed together, to relay the impact of NCpF on mental health-related clinical decisions. The strengths and limitations of each project will be identified, as well the methodological challenges of using each research method in clinical decision-making.

Discussion: This study uses a combination of designs to gain a complete view on clinical decision-making and to understand the impact of NCpF on mental health clinical decisions. It demonstrates the potential for using patient data to further the research in this field. The results will be used to: better understand patients’ perceptions on the clinical encounter; identify where inconsistencies between decisions lie; determine which individuals may be more at risk to not be diagnosed, treated or referred as a result of their NCpF; and assist clinicians to make better informed mental health-related decisions. By gaining a better understanding of the influence of NCpF on clinical decisions, it is possible to transform practice and prevent inequity in mental health care.

Quality of Imaging Requisitions Using the RI-RADS Criteria: A Single-Institution Analysis

K. Chiang1, C. Lee1, N. DeAngelo1, J. Beatty-Chadha1, J. Thomas2
1Penn State University College of Medicine, Hershey, PA
2Hershey Medical Center, Hershey, PA

Background: Diagnostic imaging requisitions ensure accuracy and efficiency of imaging study analyses and interpretations. Requisitions lacking sufficient information can contribute to diagnostic errors and potential harm to patients. We assessed the quality of our institution’s imaging requisitions using the Reason for exam Imaging Reporting and Data System (RI-RADS) criteria.

Methods: Using a picture archiving and communication system, we retrospectively assessed 102 outpatient imaging requisitions from 2018 to 2020. Evaluations of 51 CT scans of the brain/head with or without contrast and 51 MRI scans of the brain with or without contrast were performed utilizing RI-RADS criteria. Scores were determined based on the presence or lack of each of three key categories in the requisition: impression, clinical findings, and diagnostic question (Table 1). Each requisition was assigned a score of ‘A’ (all key categories included), ‘B’ (all key categories included, some clinical findings missing), ‘C’ (two categories included), or ‘D’ (one or no categories included). Grade comparisons were made between years and between CT and MRI scans.
Results: None of the requisitions received a grade of ‘A’. 10.8% received a grade of ‘B’, 31.4% received ‘C’, and 57.8% received ‘D’. Of the 102 imaging requisitions, 57.8% contained an impression, 51.0% contained clinical findings, and 33.3% contained diagnostic questions. In the CT group, 15.7% received a grade of ‘B’, 29.4% received ‘C’, and 54.9% received ‘D’. In the MRI group, 5.9% received a grade of ‘B’, 33.3% received ‘C’, and 60.8% received ‘D’. No statistically significant associations between RI-RADS grades and years (p = 0.16) or RI-RADS grades and imaging modalities (p = 0.28) were found.

Conclusion: This study highlights the potential utility of RI-RADS in providing a framework for grading radiology requisitions. However, improvement in its criteria is necessary to improve user efficiency and accuracy. In particular, the presence of certain clinical findings may be difficult to assess without an understanding of how pertinent they are to a patient’s situation. Developing a standardized framework for requisitions can enhance provider communication and reduce diagnostic errors. Future applications may lead to institutional quality improvement initiatives and efforts to incorporate requisitions into medical education.

Table 1: Reason for exam Imaging Reporting and Data System (RI-RADS) Key Categories

<table>
<thead>
<tr>
<th>Key RI-RADS Categories</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impression</td>
<td>Working or differential diagnoses</td>
</tr>
<tr>
<td>Clinical Findings</td>
<td>Signs and symptoms, chronicity of current episode, location of signs and symptoms, pertinent past medical/surgical history, pertinent laboratory findings, previous imaging reports (when available)</td>
</tr>
<tr>
<td>Diagnostic Question</td>
<td>Confirmation/exclusion of diagnosis, grading/staging, pre-operative planning, follow-up of progress, response to treatment</td>
</tr>
</tbody>
</table>

Diagnosis Documentation of Critically Ill Children on Admission to a Pediatric Intensive Care Unit

C. Cifra1, V. Vvitcharenko1, S. Ramesh1, K. Dukes1, H. Singh2, L. Herwaldt1, H. S. Reisinger1

1University of Iowa Carver College of Medicine, Iowa City, IA
2M.E. DeBakey VA Medical Center, Houston, TX

Background: Clinician notes play an essential role in establishing a shared understanding of the patient’s diagnosis among members of the healthcare team. Shared mental models are important for well-coordinated team efforts to provide accurate diagnoses and appropriate care, especially in high-risk settings like the pediatric intensive care unit (PICU). Our objective is to describe how clinicians document patients’ diagnoses on admission to the PICU with the long-term goal of improving diagnosis communication in clinical documentation to prevent diagnostic error.

Methods: We conducted a retrospective cohort mixed methods study including 100 non-elective admissions to a tertiary referral academic PICU in 2017. We performed stratified random selection to select an equal number of patients admitted during the day and night shifts by 9 attending physicians. We collected demographic/clinical data and data on word count and format of the note. We reviewed the assessment and plan sections of each note. We performed descriptive quantitative data analysis and conducted initial qualitative analysis of the assessment and plan sections, using MAXQDA software. We are currently identifying qualitative themes.

Results: We reviewed 100 PICU admission notes (Table). Both attending physicians and residents/advanced practice providers (APPs) documented the primary diagnosis in 87% of the notes, 13% were documented by residents/APPs alone. Most diagnoses (72%) were written as narrative free text, 11% were documented as problem lists/billing codes, and 17% used both formats. Diagnoses were also commonly documented outside of the assessment section and included in the plan (54%). At least one rationale for the primary diagnosis was documented in 91% of notes. Diagnostic uncertainty was expressed in 52%, most commonly via words/expressions indicating uncertainty (34%) and the presence of differential diagnoses (31%). Aside from the primary diagnosis (reason for PICU admission), clinicians also documented patients’ pre-existing complex chronic conditions related to the current problem in their assessments (35%).
Conclusions: Our preliminary results showed that aside from documenting patients’ diagnoses, clinicians included elements such as rationales and indicators of uncertainty that clarify their diagnostic reasoning. Results of qualitative analysis will help us further examine how clinicians structure diagnostic narratives and perhaps show ways to improve diagnostic documentation.

Table: Characteristics of Diagnosis Documentation

<table>
<thead>
<tr>
<th>Characteristics of Documentation</th>
<th>n* / median (IQR) (n=100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinician documenting the primary diagnosis** in the admission note</td>
<td></td>
</tr>
<tr>
<td>Attending physician alone, n</td>
<td>0</td>
</tr>
<tr>
<td>Resident or advanced practice provider alone, n</td>
<td>13</td>
</tr>
<tr>
<td>Both, n</td>
<td>87</td>
</tr>
<tr>
<td>Word count of entire H&amp;P note, median (IOR)</td>
<td>378 (276-645)</td>
</tr>
<tr>
<td>Word count of assessment field***, median (IOR)</td>
<td>44(26-71)</td>
</tr>
<tr>
<td>Admission note format</td>
<td></td>
</tr>
<tr>
<td>Traditional S O A P, chronology, n</td>
<td>84</td>
</tr>
<tr>
<td>Non-traditional format, n</td>
<td>16</td>
</tr>
<tr>
<td>Manner of diagnosis documentation of the primary diagnosis**</td>
<td></td>
</tr>
<tr>
<td>Narrative, n</td>
<td>72</td>
</tr>
<tr>
<td>Problem list or billing codes, n</td>
<td>11</td>
</tr>
<tr>
<td>Both narrative and problem list/billing codes, n</td>
<td>17</td>
</tr>
<tr>
<td>Other diagnoses documented in the assessment field</td>
<td></td>
</tr>
<tr>
<td>Additional acute diagnoses related to the primary diagnosis, n</td>
<td>10</td>
</tr>
<tr>
<td>Additional acute diagnoses not related to the primary diagnosis, n</td>
<td>9</td>
</tr>
<tr>
<td>Diagnosis documentation outside of assessment section (included in the plan), n</td>
<td>54</td>
</tr>
<tr>
<td>Rationales provided for primary diagnosis****</td>
<td></td>
</tr>
<tr>
<td>At least one rationale was provided for the primary diagnosis, n</td>
<td>91</td>
</tr>
<tr>
<td>Types of rationales provided per primary diagnosis, n</td>
<td></td>
</tr>
<tr>
<td>Rationale from patient history</td>
<td>79</td>
</tr>
<tr>
<td>Rationale from physical examination</td>
<td>38</td>
</tr>
<tr>
<td>Rationale from test results</td>
<td>28</td>
</tr>
<tr>
<td>Rationale from patient’s clinical course anchor response to therapy</td>
<td>40</td>
</tr>
<tr>
<td>Presence of uncertainty in diagnosis documentation****</td>
<td></td>
</tr>
<tr>
<td>At least one indicator of uncertainty in the primary diagnosis, n</td>
<td>52</td>
</tr>
<tr>
<td>Direct indicators of uncertainty</td>
<td></td>
</tr>
<tr>
<td>Presence of words/expressions indicating uncertainty, n</td>
<td>34</td>
</tr>
<tr>
<td>Only a symptomatic diagnosis was documented, n</td>
<td>8</td>
</tr>
<tr>
<td>Differential diagnoses were documented, n</td>
<td>31</td>
</tr>
<tr>
<td>Number of differential diagnoses provided, median (IOR)</td>
<td>1 (1-2)</td>
</tr>
<tr>
<td>Indirect indicators of uncertainty</td>
<td></td>
</tr>
<tr>
<td>Diagnostic tests were ordered to help resolve uncertainty, n</td>
<td>10</td>
</tr>
<tr>
<td>Subspeciality services were consulted to help resolve uncertainty, n</td>
<td>24</td>
</tr>
<tr>
<td>“Test-of-treatment” applied to help resolve uncertainty</td>
<td>2</td>
</tr>
<tr>
<td>Other information documented in the assessment</td>
<td></td>
</tr>
<tr>
<td>Complex chronic conditions related to the current problem, n</td>
<td>35</td>
</tr>
<tr>
<td>Complex chronic conditions not related to the current problem, n</td>
<td>9</td>
</tr>
<tr>
<td>Past treatments for previous conditions or chronic disease, n</td>
<td>22</td>
</tr>
<tr>
<td>Clinical course, n</td>
<td>22</td>
</tr>
<tr>
<td>Anticipated problems/complications and/or plans to monitor these, n</td>
<td>27</td>
</tr>
<tr>
<td>Diagnostic testing plans and/or results, n</td>
<td>12</td>
</tr>
</tbody>
</table>

PICU - pediatric intensive care unit; IQR - interquartile range S O A P. Subjective Objective Assessment. Plan *For all values, n is equivalent to n% since the denominator is 100 for all variables (total pattern admissions included) **The primary diagnosis is the explanation given for the patients current presenting problem requiring PICU admission. ***The assessment field is the section of the note wherein the clinician usually documents the primary diagnosis and other information with regard to the etiology of the patient’s primary problem. ****One primary diagnosis can have more than one rationale/Indicator of uncertainty.
Provider Perspectives on Diagnosing Anxiety Disorders in the Veterans Healthcare Administration

T. Fletcher, P. Chen, H. Singh, M. Kunik, T. Giardina
Baylor College of Medicine, Houston, TX

**Background:** Few studies have investigated diagnostic errors in psychiatry, where accurate diagnostic labels greatly impact case identification and access to appropriate treatments. Unspecified anxiety disorder is the most common anxiety diagnosis label in primary care settings but is erroneous in 60-78% of cases. This erroneous label is a barrier to receipt of appropriate evidence-based care for more specific anxiety-related disorders, such as panic disorder, generalized anxiety disorder, and posttraumatic stress disorder. However, little is known about mental health providers’ attitudes, perspectives and practices related to diagnosing anxiety disorders in primary care.

**Method:** We conducted qualitative interviews with 17 mental health providers working in Veterans’ Healthcare Administration primary care mental health integration (PCMHI) clinics, drawn from a national, randomized list of providers. Interviews assessed PCMHI providers’ perceptions about barriers to anxiety diagnostic specificity, including their knowledge and expertise in anxiety differential diagnosis, attitudes toward anxiety diagnosis, and external barriers. Qualitative interviews were analyzed using a framework-analysis approach, which allows inclusion of existing concepts, and emergent themes.

**Results:** Knowledge: Most providers demonstrated adequate knowledge of the diagnostic criterion for specific anxiety disorders. Several stated that they often consulted the Diagnosis and Statistical Manual for Mental Disorders during their assessment process to ensure an accurate diagnosis. Attitudes: Almost all providers said that anxiety symptom assessment measures were important and useful in the diagnostic process, and that accurate and timely diagnoses had significant impacts on patients’ health outcomes. However, while many providers said that having a specific anxiety diagnosis could impact patients’ health outcomes, this did not necessarily translate to making specific anxiety diagnoses. Providers reported that their primary goal was to improve patient functioning, and that they could achieve this through interventions that were transdiagnostic and not targeted to a specific anxiety disorder. External Barriers: PCMHI providers described the conflict between balancing assessment and intervention within the time constraints within of six 30-minute visits they are allotted with each patient in their clinics. Additionally, providers cited the lack of readily available assessment measures for different anxiety disorders as an additional barrier to making specific anxiety diagnoses. These PCMHI clinic barriers were the most cited reasons for not using assessment measures for more specific anxiety diagnosis.

**Conclusion:** Both mental health providers’ attitudes about the importance of diagnostic specificity and external barriers in PCMHI clinics interact to impede accurate anxiety diagnosis. Interventions targeting these two barriers are needed to ensure accurate and timely anxiety diagnosis.

Developing a Data Stream to Detect Diagnostic Errors in the Pediatric ED

J. Grubenhoff¹, D. Lam², J. Leonard², F. Dominguez³, E. Nathaniel², L. Bajaj²
¹University of Colorado School of Medicine, Aurora, CO
²University of Colorado, Aurora, CO
³Children’s Hospital Colorado, Aurora, CO

**Background:** The incidence diagnostic errors (DxEs) among children evaluated in emergency departments (EDs) and urgent cares (UCs) has not been described. DxEs are infrequently submitted to traditional incident reporting systems precluding opportunities to learn from them. Defining the incidence of DxEs in the ED/UC and describing predisposing patient demographic and clinical characteristics will improve the ability to identify vulnerabilities in the diagnostic process amenable to systematic improvement in diagnostic performance and provide context for measuring improvement.

**Methods:** Retrospective chart review of patients initially evaluated (index visit) in a large pediatric hospital ED/UC system (5 satellite ED/UCs; 1 regional pediatric trauma center ED; 162,347 annual visits) followed by hospital admission within 14 days from January 1 to December 31, 2018. An e-trigger identified cases meeting this definition which were then screened to determine if a possible DxE occurred. Cases screened in were reviewed by two reviewers [Cohen’s ? = 0.60]
(95% CI 0.47-0.73) using the Revised SaferDx instrument to identify probable DxEs. Patient demographic and clinical characteristics at the index encounter were abstracted for comparison. Descriptive statistics were used to compare groups.

**Results:** In 2018, from a total visit volume of 162,347 encounters, 926 patients were identified using the e-trigger; 251 (27.1%) were screened in for review. DxEs were identified in 47 (18.7%) for an incidence of 50.1 DxEs/1000 e-trigger-identified encounters per year. The incidence across all encounters for 2018 was 0.3/1000 encounters per year. The Table shows the demographic and clinical characteristics of the cohort. Five patients had no prior migraine history and were initially diagnosed with migraine headache but subsequently diagnosed with serious intracranial pathology (cerebral abscess, cavernous sinus thrombosis, ruptured arachnoid cyst).

**Conclusions:** The incidence of diagnostic errors resulting in subsequent admission within 2 weeks is infrequent in a large pediatric hospital ED/UC network. However, application of an e-trigger followed by screening demonstrated that almost 20% of cases screened in represented a likely diagnostic error after structured chart review. These patients had higher utilization of ICU resources and longer lengths of stay compared to patients screened out suggesting diagnostic delays contribute to higher morbidity.

<table>
<thead>
<tr>
<th>Patient Demographics and Visit Characteristics for Patients Identified Using e-trigger</th>
<th>Screened In DX Error (n=47)</th>
<th>No DX Error (n=201)</th>
<th>Screened Out (n=675)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>6.9 (2.3-13.4)</td>
<td>5.4 &lt; 1.8-11.9</td>
<td>2.6 (0.9-9.0)</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Male</td>
<td>32 (68.1%)</td>
<td>100 (49.8%)</td>
<td>337 (49.9%)</td>
<td>0.05</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>0.55</td>
</tr>
<tr>
<td>Asian</td>
<td>0 (0%)</td>
<td>8 (4.0%)</td>
<td>23 (3.5%)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>4 (8.7%)</td>
<td>15 (7.6%)</td>
<td>76 (11.5%)</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>30 (65.2%)</td>
<td>130 (65.7%)</td>
<td>404 (61.0%)</td>
<td></td>
</tr>
<tr>
<td>Other (Amer Indian, Native Hawaiian, Other, &gt; 1 race)</td>
<td>12 (26.1%)</td>
<td>45 (22.7%)</td>
<td>159 (24.0%)</td>
<td></td>
</tr>
<tr>
<td>Ethnicity, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>0.09</td>
</tr>
<tr>
<td>Latino</td>
<td>24 (51.1%)</td>
<td>70 (35.9%)</td>
<td>231 (35.2%)</td>
<td></td>
</tr>
<tr>
<td>Not Latino</td>
<td>23 (48.9%)</td>
<td>125 (64.1%)</td>
<td>425 (64.8%)</td>
<td></td>
</tr>
<tr>
<td>Preferred Language</td>
<td></td>
<td></td>
<td></td>
<td>0.14</td>
</tr>
<tr>
<td>English</td>
<td>37 (78.7%)</td>
<td>173 (86.1%)</td>
<td>581 (86.1%)</td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td>10 (21.3%)</td>
<td>20 (9.9%)</td>
<td>75 (11.1%)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>0 (0%)</td>
<td>8 (4.0%)</td>
<td>19 (2.8%)</td>
<td></td>
</tr>
<tr>
<td>ESI Acuity</td>
<td></td>
<td></td>
<td></td>
<td>0.13</td>
</tr>
<tr>
<td>L2</td>
<td>3 (6.4%)</td>
<td>37 (18.4%)</td>
<td>113 (16.8%)</td>
<td></td>
</tr>
<tr>
<td>L3-L5</td>
<td>44 (93.6%)</td>
<td>164 (81.6%)</td>
<td>561 (82.2%)</td>
<td></td>
</tr>
<tr>
<td>Index Visit ED Arrival Time</td>
<td></td>
<td></td>
<td></td>
<td>0.58</td>
</tr>
<tr>
<td>6am – 2pm</td>
<td>18 (38.3%)</td>
<td>65 (32.3%)</td>
<td>235 (34.8%)</td>
<td></td>
</tr>
<tr>
<td>2pm – 10pm</td>
<td>24 (51.1%)</td>
<td>96 (47.8%)</td>
<td>330 (48.9%)</td>
<td></td>
</tr>
<tr>
<td>10pm – 6am</td>
<td>5 (10.6%)</td>
<td>40 (19.9%)</td>
<td>110 (16.3%)</td>
<td></td>
</tr>
<tr>
<td>Financial Class</td>
<td></td>
<td></td>
<td></td>
<td>0.28</td>
</tr>
<tr>
<td>Contract</td>
<td>21 (44.7%)</td>
<td>89 (44.3%)</td>
<td>249 (36.9%)</td>
<td></td>
</tr>
<tr>
<td>Medicaid</td>
<td>25 (53.2%)</td>
<td>103 (51.2%)</td>
<td>400 (59.3%)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>1 (2.1%)</td>
<td>9 (4.5%)</td>
<td>26 (3.8%)</td>
<td></td>
</tr>
<tr>
<td>Revisit ICU Stay</td>
<td>12 (25.5%)</td>
<td>48 (23.9%)</td>
<td>110 (16.3%)</td>
<td>0.02</td>
</tr>
<tr>
<td>Number of Admissions in the Last 6 months (median, IQR)</td>
<td>0 (0-0)</td>
<td>0 (0-0)</td>
<td>0 (0-1)</td>
<td>0.01</td>
</tr>
<tr>
<td>Stratified by # of visits (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>27 (57.4%)</td>
<td>106 (52.7%)</td>
<td>291 (43.1%)</td>
<td>0.03</td>
</tr>
<tr>
<td>1</td>
<td>11 (23.4%)</td>
<td>47 (23.4%)</td>
<td>161 (23.9%)</td>
<td></td>
</tr>
<tr>
<td>≥2</td>
<td>9 (19.2%)</td>
<td>48 (23.9%)</td>
<td>223 (33.0%)</td>
<td></td>
</tr>
<tr>
<td>Days between Index Encounter and Admission</td>
<td>2.2 (1.6-4.0)</td>
<td>2.9 (1.8-6.2)</td>
<td>3.3 (1.9-7.0)</td>
<td>0.01</td>
</tr>
<tr>
<td>Admission Length of Stay</td>
<td>4.1 (1.5-7.1)</td>
<td>3.1 (2.1-6.6)</td>
<td>2.3 (1.6-3.9)</td>
<td>&lt;0.001†</td>
</tr>
</tbody>
</table>

Data are presented as n (%) except age. days between visits, and revisit length of stay which are presented as median (IQR). Data are missing for: race (n=17); ethnicity (n=25); ESI acuity level (n=4) *Difference related to large number of infants initially diagnosed with a viral upper respiratory infection or bronchiolitis who progressed to requiring admission for respiratory support for bronchiolitis. †Too few cases to provide stable estimate.
Measuring the Proportion of Cancer Diagnosed an Emergency in the VA Health System

P. Kapadia¹, D. Murphy², A. Zimolzak², G. Lyratzopoulos³, G. Abel⁴, T. Scott², H. Singh²
¹Baylor College of Medicine, Houston, TX
²Michael E. DeBakey VAMC, Houston, TX
³University College London, London, England
⁴University of Exeter, Exeter, England

Background: Emergency diagnosis of cancer is associated with poor outcomes. This phenomenon is generally defined as diagnosis made through presentation requiring urgent medical attention and is driven by both unavoidable (e.g., from tumor aggressiveness) and avoidable (e.g., from suboptimal pre-diagnostic care) presentations. In the UK, emergency cancer presentation has been used to measure cancer-related diagnostic quality of local health economies. As a first step in developing similar measures for the US healthcare system, we created an electronic algorithm ("e-measure") to measure the proportion of lung and colorectal cancer diagnoses that were emergency diagnoses.

Methods: We developed and validated an algorithmically calculated “e-measure” to quantify the proportion of emergency cancer diagnosis (Emergency Diagnoses divided by All Diagnoses). We consulted an expert panel of physicians, administrators, researchers, and patients using the Safer Dx Trigger Tools Framework, an iterative model for the development, validation, and refinement of e-triggers, to develop this e-measure. Guided by this framework and algorithmic definitions in the UK, we operationalized the e-measure in a national repository of electronic health records (EHRs). “Diagnosis” of cancer was operationally defined as the first occurrence of a cancer ICD-10 code in a patient’s EHR and “emergency diagnosis” was operationally defined as a diagnosis occurring in an inpatient setting or in an outpatient setting but following an emergency department visit that occurred within 30 days prior to date of diagnosis. We applied this e-measure to the Veterans Affairs data containing EHRs of approximately >10 million active veterans to measure the amount and proportion of emergent cancer diagnoses.

Results: Our e-measure algorithm found that 607 (435 inpatient and 172 outpatient) out of 7371 new colorectal cancer diagnoses and 1054 (823 inpatient and 231 outpatient) out of 8953 new lung cancer diagnoses in 2018 were potential emergency diagnoses. This represented 8.23% and 11.77% of new cancer diagnoses, respectively.

Conclusion: Our preliminary e-measure algorithm found that a substantial proportion of patients with lung and colorectal cancer have an ER visit in the 30-days before their diagnosis. We will review randomly selected patient records of these potential emergency diagnoses to validate the accuracy of our e-measure and explore patient and cancer characteristics associated with emergency cancer diagnosis. We will additionally validate our e-measure by examining its prognostic implications for treatment and survival in our cohort, accounting for stage at diagnosis. Furthermore, we will continue to refine our e-measure and explore new operational definitions through consultation with domain experts.

Patterns for Diagnosis of Pancreatic Cancer among U.S. Veterans

N. Khalaf, K. Choi, F. Kanwal, H. Singh
Baylor College of Medicine, Houston, TX

Background: Pancreatic cancer is the number three cancer killer in the U.S. and is projected to rise to number two in the next decade. Nearly 85% of patients have incurable (locally advanced/metastatic) disease at diagnosis suggesting need for early cancer detection efforts. As a first step to identify future opportunities for diagnosis improvement, we characterized symptoms, sociodemographic and clinical characteristics and diagnostic intervals in a pancreatic cancer cohort at a large VA center.

Methods: We reviewed electronic health records of 50 consecutive patients diagnosed with pancreatic adenocarcinoma from 1/2015 to 12/2017 with at least 3 years of comprehensive pre-cancer diagnosis medical data. We evaluated clinician notes, referral patterns, and tests, including radiology and pathology files. We used descriptive analyses to characterize the patient population in terms of: (1) known pancreatic cancer risk factors; (2) cancer-associated signs and symptoms and initial presentation; and (3) diagnostic intervals.

Results: Among 50 patients, 98% were men and 50% were White, with a mean age of 68.3 years (SD 9.0) and BMI of 25.9 (SD 5.7) at diagnosis. 70% were =65 years of age, 42% had diabetes mellitus, and 78% were current (36%) or former (42%)
tobacco users. Average survival was 8.2 months. Abdominal pain was the most common symptom and emergency department the most common setting of presentation (Table). In terms of diagnostic intervals, we calculated average time between: onset of first cancer-associated symptom to physician report of symptom (patient delay: 42 days, range 1-311 days); onset of first symptom to histological diagnosis (diagnostic delay: 96 days, range 6-359); histological diagnosis to treatment initiation (treatment delay among the 25 people who received chemotherapy: 46 days, range 15-86 days); and onset of first symptom to treatment initiation (overall delay among the 25 people who received chemotherapy: 134 days, range 48-344 days).

**Conclusion:** Most patients with pancreatic adenocarcinoma had at least 1 known risk factor and a classic warning sign or symptom of pancreatic cancer. The average diagnostic delay was >3 months, suggesting the presence of opportunities for earlier diagnosis. As a next step, we will evaluate diagnostic process breakdowns in a large national cohort.

Table: Symptoms and other characteristics of pancreatic cancer patients at one VA facility (2015-2017)

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of participants</td>
<td>50 (100)</td>
</tr>
<tr>
<td><strong>Initial cancer sign or symptom</strong></td>
<td></td>
</tr>
<tr>
<td>Abdominal or back pain</td>
<td>27 (54)</td>
</tr>
<tr>
<td>Anorexia or weight loss</td>
<td>21 (42)</td>
</tr>
<tr>
<td>Nausea or vomiting</td>
<td>7 (14)</td>
</tr>
<tr>
<td>Change in bowel habits</td>
<td>11 (22)</td>
</tr>
<tr>
<td>Jaundice</td>
<td>11 (22)</td>
</tr>
<tr>
<td>Other</td>
<td>10 (20)</td>
</tr>
<tr>
<td>Incidental finding on imaging</td>
<td>7 (14)</td>
</tr>
<tr>
<td><strong>Setting of first complaint</strong></td>
<td></td>
</tr>
<tr>
<td>Ambulatory clinic visit</td>
<td></td>
</tr>
<tr>
<td>General provider</td>
<td>17 (34)</td>
</tr>
<tr>
<td>Specialty provider</td>
<td>4 (8)</td>
</tr>
<tr>
<td>Emergency department</td>
<td>19 (38)</td>
</tr>
<tr>
<td>Upon presentation to outside hospital</td>
<td>3 (6)</td>
</tr>
<tr>
<td>Incidental finding on imaging</td>
<td>7 (14)</td>
</tr>
</tbody>
</table>

*A patient could have more than one sign/symptom accounted for in this section*

Missed Opportunities to Diagnose Epidural Abscess are Common Among People Who Inject Drugs

A. Miller, P. Polgreen
University of Iowa, Iowa City, IA

**Background:** People who inject drugs (PWID) are at increased risk for developing epidural abscesses. Failure to diagnose and treat an epidural abscess may result in lasting neurologic sequelae. However, symptoms of epidural abscess (e.g., back pain) may be dismissed in PWID as “drug-seeking behavior.” Thus, diagnostic delays may be common among PWID. **Methods:** We conducted a retrospective cohort study using insurance claims data from the IBM Marketscan Research Database for the period 2001-2017. We identified index case visits when patients were first diagnosed with an epidural abscess. We then identified cases of epidural abscess where the patient also had a diagnosis of drug abuse at, or within 6 months of, the index infection. Next, we analyzed all healthcare visits (inpatient, outpatient or emergency department) in the 180 days prior to the epidural abscess diagnosis for related symptoms. A time-series change-point detection analysis was used to identify the point in time where symptoms first began to appear prior to the index infection diagnosis. A simulation analysis was then used to estimate the likelihood and duration of diagnostic delays. Finally, a logistic regression model was used to estimate risk factors for experiencing a diagnostic delay. **Results:** We identified 39,818 patients with an epidural abscess, 1,428 of these were in PWID. Visits for related symptoms increased dramatically before the index diagnosis (see attached figure). Among PWID, we identified 3,024 visits, from 813
patients, with at least one symptom within 30 days prior to the diagnosis. We estimate that 563 visits (95% CI 547-579), in
379 patients (CI 364-394), represented likely diagnostic delays. PWID were more likely to experience a diagnostic delay
compared to non-drug users: 26.5% (CI 25.5%-27.6%) of PWID experienced a delay, compared to 12.0% (CI 11.9%-12.3%)
of non-drug users. On average, delays were longer in PWID: 8.50 days (CI 8.28-8.71) in PWID vs 4.58 days (CI 4.51-4.65) in
non-drug users.

**Conclusion:** Many opportunities to diagnoses epidural abscess are missed or delayed. PWID are at greater risk for
diagnostic delays and tend to experience longer delays than patients who do not inject drugs.

---

**Retrospective Analysis of Inpatient Brain MRI Usage Reveals Opportunities for Improving Utility**

A. Pisupati¹, J. Beatty-Chadha¹, J. Thomas²

¹Penn State College of Medicine, Hershey, PA
²Penn State Health, Radiology, Hershey, PA

**Background:** The comparatively high resolution of magnetic resonance imaging (MRI) as a modality makes this tech-
nique vital for the diagnosis of a wide array of neurologic diseases. Brain MRIs are not immune to diagnostic error,
however. Moreover, there are numerous logistical and monetary constraints that can limit how much brain MRIs can be
utilized in the inpatient setting. Given these limitations, it is imperative that use of brain MRIs is clinically appropriate and
necessary.

**Methods:** In this retrospective analysis, use of brain MRIs by different clinician services at a single institution is
quantified with respect to whether or not each MRI diagnosis is consistent with the primary diagnosis at discharge.
Secondary analysis of radiologist recommendation follow-up for additional studies by the primary clinical team is
performed. Finally, diagnostic inpatient brain MRI utilization is assessed using the American College of Radiology (ACR)
appropriateness criteria.

**Results:** Preliminary analysis of diagnostic non-contrast brain MRIs over a three-week period suggest that almost half
(17/35) of MRI diagnoses are only partially consistent or inconsistent with the diagnosis at discharge. This would suggest
that a significant number of MRIs that are ordered as an inpatient may not be necessary. Secondary analysis suggests
that diagnostic inconsistency is more likely in the setting of patients with multifactorial discharge diagnoses or if the MRI
order is made in conjunction with or in anticipation of a specialist consult.

**Conclusions:** These findings suggest that MRI utilization in an inpatient setting requires greater scrutiny given the
significant number of studies that may not actually alter disease management. Diagnostic errors could be reduced simply
by reducing the number of unnecessary imaging studies that are ordered, allowing for greater resource allocation to
studies that may be more worthwhile. Possible solutions to reducing unnecessary MRI orders and prevent overuse
Trainee Diagnosis in the Age of Clinical Pathways

I. Rasooly¹, M. Congdon¹, C. Clancy², N. Muthu¹, D. Balmer¹, C. Bonafide¹
¹Children’s Hospital of Philadelphia, Philadelphia, PA
²University of Pennsylvania, Philadelphia, PA

Introduction: Clinical pathways, which translate evidence into algorithms for standardized diagnosis and management, are ubiquitous in US hospitals. There is concern that widespread use of clinical pathways has altered trainee knowledge acquisition and clinical reasoning skill development. Interactions between pathways and the development of diagnostic reasoning skills among trainees may have long-term unintended effects on delivery of safe, effective care.

Objectives: Characterize ways residents employ inpatient clinical pathways in diagnostic decision making, to enhance their educational value.

Methods: We conducted semi-structured interviews with pediatric residents purposefully sampled at a large, academic hospital. Interviews were recorded and transcribed. We created codes deductively from human factors theory and inductively from patterns in the data. Data was analyzed using directed content analysis.

Results: Nine interviews lasting 30-60 minutes were conducted from May to Dec 2019. Residents primarily situated their role in the diagnostic process as primarily “communication of the diagnosis,” “treatment,” and “outcomes,” rather than engaging in diagnostic reasoning for patients on clinical pathways. All described using clinical pathways primarily as a reference for specific management guidance. For instance, one resident stated, “I’ll go [to the pathway] because I really just want to find the antibiotic recommendations.” While all used pathway-associated orders, few looked at pathway diagnostic algorithms. Trainees described rare occasions in which they reconsidered diagnoses but expressed that these occurrences prompted them to restart the diagnostic process from scratch, rather than utilizing the diagnostic guidance embedded within the pathway. Residents relied on bedside assessment to identify when patients are “not quite fitting the mold” of the current pathway diagnosis and many expressed that diagnostic reassessment was not built into pathways. This was explained by one resident as, “I have to admit that if a kid isn’t improving, I don’t necessarily look at [the] pathway. I kind of then use my clinical judgment about what comes next.”

Conclusions: Understanding the workflow and cognitive processes of trainees is essential in developing strategies to address educational implications of new technologies. Trainees turn to pathways for details of clinical management but are not engaged with pathway diagnostic guidance or, at times, in making a diagnosis at all. To promote development of diagnostic skills, it may be helpful to embed prompts or tools for diagnostic reassessment within the order sets that trainees commonly look at rather than relying on passive diagnostic support in the form of infrequently referenced algorithms.
Prevalence and Perceptions of Differential Diagnosis Support Systems Amongst Students, Residents, and Faculty Physicians

S. Sathianathan, A. DeWaters, N. Duca
Penn State College of Medicine, Hershey, PA

Background: Differential diagnosis support systems (DDSS), such as Isabel, Diagnosaurus, DynamedPlus, are historically unpopular among physicians. However, at early stages of training, learners are searching for methods to improve their clinical reasoning skills. DDSS may provide students with the immediate feedback they need to develop clinical reasoning skills and expose heuristics in critical thinking. The purpose of this study is to identify the prevalence and perceptions of DDSS use amongst students, residents, and physicians.

Methods: Between July 2019 and March 2020, we surveyed 22 adult primary care physicians and residents (family medicine and internal medicine) and 45 medical students in primary care clerkships. The survey assessed prevalence of DDSS use, utility based on desired outcomes, and perceptions about using DDSS and teaching clinical reasoning in medical education. Responses were analyzed via multinomial test for equality of proportions. P-values from post-hoc comparisons were adjusted for multiple comparison error.

Results: There was no significant difference between the proportions of students, residents, and physicians who used DDSS \((p = 0.23)\). On average, students used DDSS more frequently (between “sometimes” and “most times”) than physicians (“rarely” to “never”). Dynamed Plus was used among 49% of respondents, significantly more than other DDSS \((p < 0.05)\). There was no significant difference between physicians, residents, and students when they used DDSS to help “broaden” or “narrow” differentials, or to “avoid biases.” However, students and residents used DDSS significantly more than physicians to “determine further testing,” “to learn something new,” and “when not sure what to do” \((p < 0.01)\). Nonusers of DDSS felt that DDSS “hindered clinical reasoning” more than users. Otherwise, users and nonusers similarly felt that DDSS were “[less] reliable than an expert physician’s clinical reasoning” and agreed that they were “user dependent.”

Conclusion: Our results suggest that students and residents use DDSS more frequently than physicians for specific educational and clinical purposes. Students, residents, and physicians use DDSS to broaden differentials and avoid biases in clinical reasoning. Learners are more likely than physicians to use DDSS in practice to help combat uncertainty, increase their fund of knowledge, and to determine next steps. DDSS may mitigate early sources of clinical reasoning errors among learners, but both users and non-users recognize the limitations of DDSS in supporting clinical reasoning. Further research is needed to evaluate the benefit of these tools for education and clinical care amongst learners.

Development of a Machine Learning Enhanced Trigger to Detect Diagnostic Error

V. Vaghani\(^1\), P. Kapadia\(^2\), A. Zimolzak\(^3\), H. Singh\(^4\), D. Subramanian\(^5\)
\(^1\)Baylor College of Medicine & Michael E. DeBakey VA Medical Center, Houston, TX
\(^2\)Baylor College of Medicine, Houston, TX
\(^3\)Michael E. DeBakey VA Medical Center, Houston, TX
\(^4\)Rice University, Houston, TX

Background: Missed opportunities for diagnosis (MODs) are an important type of medical error but are difficult to identify in large-scale EHR data. Prior work has screened for MODs using rules-based electronic triggers applied to EHRs followed by manual chart review. We explored development of a machine learning (ML)-enhanced e-trigger.

Methods: Using the Safer Dx Trigger Tools and the Symptom-Disease Pair Analysis of Diagnostic Error frameworks, we previously developed a rules-based e-trigger for potential stroked-related MODs. We used diagnostic codes to identify patient records with “treat-and-release” emergency department visits based on return hospitalizations for stroke. We applied this e-trigger to the Veterans Affairs corporate data warehouse and reviewed trigger-positive records, creating a secondary database of manually abstracted variables (e.g. symptomatic flags, risk factors, exam findings, and management) and annotation for definite, potential, or absent MOD. In this project, we trained and tested logistic regression
and random forest classifiers of MODs using 37 selected variables, including risk factors, exam findings, and demographic data, on this secondary database from prior work.

**Results:** Of 398 trigger-positive records, 181 were excluded due to erroneous/irrelevant billing codes in prior study. Of the remaining 217 records, 43 had definite MOD, 81 had potential MOD, and 93 had no MOD. Neurology consult and stroke symptom flags were the strongest predictors of MOD in the ML models trained from this data. The logistic regression model had a precision (analogous to positive predictive value) of 0.83 and recall (analogous to sensitivity) of 0.82 for correctly classifying MOD, while the random forest model had a precision of 0.82 and recall of 0.83. Most misclassifications across both models occurred between potential and absent MODs due to the classifiers’ inability to utilize features abstracted from the neurological exam and capture heuristic decisions unrecorded in the database.

**Conclusion:** While sample size was small, we explored potential for ML to enhance current rules-based e-triggers to identify MODs. Currently, ML relies in part on manually abstracted variables, but future work will develop classifiers that rely more on structured data with minimal manual abstraction to evaluate if ML can help retrieve large numbers of records with MODs.

**Evaluation of the Diagnostic Process in Patients Admitted with Dyspnea to Hospital Medicine Department**

E. Yousef  
Cleveland Clinic, Cleveland, OH

**Background:** The National Academy of Medicine (NAM) describes the diagnosis process as a complex, patient-centered and collaborative activity. A breakdown in any of its steps can lead to diagnosis error—defined as; failure to establish an accurate and timely explanation of the patient’s health problem or communicate that explanation to the patient. Such errors can result in unnecessary workups, delayed, wrong or missed diagnosis, readmission, and sometimes, mortality. In this study, we evaluated the diagnosis process in patients admitted with dyspnea, a common cause of hospitalization that is often difficult to uncover due to multi-organ involvement—to identify any errors associated with diagnostic process.

**Methods:** In this retrospective chart review, we identified all patients who had been admitted either from ED or were transferred from Medicine ICU to Depart of Hosp Med due to dyspnea between January and March. We used a custom template to comprehensively evaluate the different steps of the diagnosis process as described by the NAM—engagement, information gathering, information integration, information interpretation, establishing an explanation, and communicating the explanation. We then categorized the diagnostic process outcome as; error with harm (patient death), error with no harm, and no error no harm. Causes of identified failures in each step were further analyzed using modified fish
bone model for root cause analysis. Finally, the number of readmissions with dyspnea within 4 weeks of hospital discharge was tabulated.

**Results:** We validated 48 patient charts-38 admitted from the ED and 10 transferred from the MICU. (1-A), The diagnostic process outcomes. (1-B), Where errors occurred in diagnostic process. (2-A) Root cause analysis of the identified failures. (2-B) Readmission rates within 4 weeks for dyspnea. Failures in the diagnosis process were similar between the patients who were admitted from the ED or the MICU. There were overlap of multiple steps failures in the same chart. Other potential failures were difficult to clarify due to lack of documentation related to these steps. This included failure to communicate the explanation (unclear if the diagnosis was verbally communicated) and failure of engagement (no documentation of how soon the patient sought medical care)

**Conclusion:** Failures in the diagnosis process occurred at every step but were mostly evident at the information gathering. Educating clinicians about the diagnosis process and the importance accurately obtaining history, performing a comprehensive physical examination and ordering appropriate work up within the context of the patient’s health problem may reduce errors in the diagnosis process.

![Figure 1: A,B](image-url)

**Prediction of Acute Appendicitis among Patients with Undifferentiated Abdominal Pain at Emergency Department**

X. Zhang¹, D. Su¹, P. Mahajan¹, Q. Li¹, K. He¹, T. Zhang², P. Veliz¹, Y. Chen³,

¹University of Michigan, Ann Arbor, MI
²Sichuan University, China
³Huazhong University of Science and Technology, China

**Background:** Early screening and accurately identifying Acute Appendicitis (AA) among patients with undifferentiated symptoms associated with appendicitis during their emergency visit will improve patient safety and health care quality. The aim of the study was to develop models to predict AA among patients with undifferentiated symptoms at emergency visits using both structured data and free-text data from a national survey.
Methods: We performed a secondary data analysis on the 2005-2017 United States National Hospital Ambulatory Medical Care Survey data to estimate the association between ED patients with the diagnosis of AA, and the demographic and clinical factors present at ED visits during a patient’s ED stay. We used multivariable logistic regression and random forest tree models incorporating natural language processing (NLP) to predict AA diagnosis among patients with undifferentiated symptoms.

Results: Among the 40,441 ED patients with assigned ICD codes of AA and appendicitis-related symptoms between 2005 and 2017, 655 adults (2.3%) and 256 children (2.2%) had AA. For the logistic regression model identifying AA diagnosis among adult ED patients, the c-statistics was 0.72 (95% CI: 0.69-0.75) for structured variables only, 0.72 (95% CI: 0.69-0.75) for unstructured variables only, and 0.78 (95% CI: 0.76-0.80) when including both structured and unstructured variables. For the logistic regression model identifying AA diagnosis among pediatric ED patients, the c-statistics was 0.84 (95% CI: 0.79-0.89) for including structured variables only, 0.78 (95% CI: 0.72-0.84) for unstructured variables, and 0.87 (95% CI: 0.83-0.91) when including both structured and unstructured variables. The random forest tree showed similar c-statistics to the corresponding logistic regression models.

Conclusions: We developed predictive models that can predict the AA diagnosis for adult and pediatric ED patients, and the predictive accuracy was improved with the inclusion of natural language processing elements and approaches.

Clinical Vignettes

Clinical Triads: Tried and Untrue

J. Alexander, C. Jackson
UTHSC-Internal Medicine Residency, Memphis, TN

Case Information: A 52-year-old male with alcohol use disorder, remote left cerebellar CVA, vitamin B12 and folate deficiency presented to the ED with nausea and more than 40 episodes of non-bloody, non-bilious vomiting for 4 days. He quit drinking four days prior to arrival and had been unable to keep any food down during that period. He was hospitalized six weeks ago for pneumonia. He reported having a severe headache, photophobia, and new onset gait instability with ataxia. On initial evaluation, patient was mildly confused with a BMI of 18. His neurological exam revealed left lower facial half droop, and right lateral gaze restriction. Laboratory studies notable for an elevated anion gap metabolic acidosis with an elevated osmolar gap, lactic acidosis, and ketonuria. UDS and EtOH were negative. CT head, MRI brain, c-spine, and CTA were unremarkable for any acute changes. Patient received one intravenous fomepizole, intravenous thiamine, and D5 ½ normal saline for treatment of alcoholic ketoacidosis. While his anion gap metabolic acidosis and left-sided facial droop improved, his right lateral gaze restriction and ataxia persisted. By hospital day two, his facial droop, ataxia, headache, photophobia, lateral gaze restriction had completely resolved. Ethylene glycol/methanol levels were negative. Due to sustained and rapid resolution of his symptoms, negative imaging, and complete and sustained neurologic recovery after only continued parenteral thiamine the diagnosis of Wernicke encephalopathy (WE) was made.

Discussion: WE is an acute neuropsychiatric emergency cause by thiamine deficiency. WE is missed in approximately 75-85% of cases,1 and classic triad of gait ataxia, oculocutaneous deficits, and confusion are reported in only 16%-33% of patients. Treatment within the first 48-72 hours of onset can lead to rapid and sustained symptom resolution, while treatment delay may cause permanent disability and Korsakoff syndrome. Diagnosis is primarily clinical by identifying any two of the following 4 criteria: 1. evidence of dietary deficiency 2. ataxia 3. confusion 4. oculocutaneous signs. Imaging may help confirm the diagnosis but is only 53% sensitive. Parenteral thiamine should be considered in any patient at risk for or meeting the diagnostic criteria for WE, and MRI/lab testing levels should be not delay therapy. Because the symptoms are nonspecific and variable, it is important to rule out other conditions that may have similar presentations to avoid anchoring and premature closure, including stroke, overdose, meningitis intracranial hemorrhage, etc.
Improper Management of Ulcerative Colitis for Jail Inmate

E. Cohen¹, S. S. Sarvepalli¹, T. E. Russeth¹, L. Selitsky², D. L. Levine¹
¹Wayne State University School of Medicine, Detroit, MI
²Johns Hopkins University, Baltimore, MD

Learning Objectives: 1. Recognize challenges in transition of care for jail inmates. 2. Emphasize the need for universal health screenings of chronic conditions upon jail admission.

Case Information: A 33-year-old male with previously well-controlled ulcerative colitis (UC), presented to the emergency room complaining of 3 weeks of abdominal pain. The patient was brought in from a county jail where his previously prescribed adalimumab prescription was not continued over the course of his 3-month incarceration. The patient stated that he did not receive a health screening upon jail admission. He reported severe gastrointestinal distress and 13-15 bloody bowel movements a day. During hospitalization, the patient was treated with steroids and adalimumab. Prior to discharge, the medical team confirmed continuation of adalimumab with the jail, however, in jail, the patient was only given acetaminophen for pain. The patient’s condition further deteriorated, and he was readmitted four days later with worsening symptoms, low hemoglobin, and anemia. He was treated with blood transfusion and methylprednisolone and discharged with the same medical regimen as recommended on the first admission. On return to jail, he received the recommended treatment.

Discussion: Upon entry to jail, our patient went three months without receiving his previously prescribed adalimumab, which likely exacerbated his UC leading to hospitalization. This lapse in appropriate care may be traced back to the lack of a health screening upon jail admission, as reported by our patient. If a proper screening was performed, his adalimumab could have been continued, and the patient’s subsequent two hospitalizations likely could have been prevented. These occurrences are common in the healthcare of incarcerated patients. While there has been significant attention given to psychiatric screenings, no studies or policies are evident for screenings of chronic medical conditions upon admission. The lack of comprehensive medical screening poses a challenge in proper and timely diagnosis of chronic conditions.

Percutaneous Endoscopic Gastrostomy: Are We Doing the Right Thing?

F. Dell’Otto¹, R. Lo Piccolo¹, A. Morabito²
¹Meyer Children’s Hospital, University of Florence, Florence, Italy
²University of Florence, Florence, Italy

Learning Objectives: Percutaneous Endoscopic Gastrostomy (PEG) is an accepted and increasingly utilised means of obtaining access to the gastrointestinal tract. Complications such as bowel injury have been described. Failure to diagnose PEG dislocation can lead to serious consequences for the patients.

Case Information: A 13 years old patient affected by cystic fibrosis, esocrine pancreatic insufficiency and liver disease underwent “pull-PEG” because of feeding difficulties, weight loss and nervous anorexia. Soon after the procedure biliary fluid leakage from around the PEG insertion was noted. The leakage remains through the years and was associated with diarrhea and abdominal pain. The patient underwent a contrast study, reported as normal, through the PEG, that revealed an opacified intestinal loop without identification of the stomach (Fig.1). No further action were taken at the time. Four years later and after six-monthly PEG changes, a gastroscopy was performed to better understand the nature of the clinical complaints. The PEG was not seen inside the stomach. An upper GI series revealed that the PEG was sitting inside a jejunal loop. The patient underwent laparoscopic closure of the jejunoscopy with resolution of symptoms.

Discussion: PEG carries risks and complications. We believe that Laparoscopic Assisted - PEG should be the gold standard for all patients to avoid complications.
Myasthenia Gravis and Septic Joint: Navigating the Minefield of Antibiotic Induced Myasthenic Crisis

H. Lee, A. Ajaz
Wayne State University School of Medicine, Detroit, MI

Objectives: Literature on Myasthenia gravis crisis precipitated by certain antibiotics is sparse. Our understanding of this topic depends on reported adverse effects or successful treatments. We believe that successful antibiotic treatment of septic arthritis in a patient with a history of myasthenia gravis has not been reported previously.

Case Information: A 48 year old female with history of myasthenia gravis, status post thymectomy, SLE presented with complaints of right wrist pain that started 2 days prior to admission. The pain was sudden, achy, and rated at 10/10, awaking her from sleep. The joint was swollen, warm, and tender to touch, limiting her range of motion. The overlying skin lacked erythema. She denied recent or history of trauma to the area. She endorsed subjective fevers and recorded home measurements of 100.5°F. Other complaints included cough for 2 weeks, productive of yellowish-white sputum, without hemoptysis. Patient denied shortness of breath, chest pain, night sweats, sick contacts, or recent travel. With concerns for septic arthritis, orthopedics was consulted and I&D performed, with a purulent sample and blood sent for cultures. With concern for staph infection, the patient was started on linezolid. Vancomycin was avoided given the history of myasthenia gravis and concern for exacerbation or crisis. The joint fluid culture grew pan sensitive Streptococcus pneumonia, which also grew in the blood cultures, believed to be secondary to dissemination of wrist joint infection. Orthopedics performed washout and closure of the joint. Patient completed 6 days of ceftriaxone treatment. Repeat blood cultures showed clearance of strep infection.

Discussion: Reports of myasthenic crisis, or avoidance thereof, consequent to antibiotics administration is sparse. Of the preexisting, one reported permanent peripheral neuropathy after fluoroquinolone administration. Another case describes successful treatment of bronchopulmonary infection with cefoperazone/sulbactam. The movement to better understand this phenomenon continues to grow.
In Contrast to What You See Here...

E. Leven1, A. Linker2
1Icahn School of Medicine at Mount Sinai, New York City, NY
2Mount Sinai Hospital, New York City, NY

**Learning Objectives:** 1. Critically evaluate alternative explanations to imaging findings that are inconsistent with clinical history. 2. Analyze diagnostic challenges posed by the introduction of new therapies with incompletely recognized effects.

**Case Information:** A 66 year-old woman with complex past medical history was hospitalized with acute respiratory and renal failure secondary to COVID-19. While on therapeutic anticoagulation she developed hypotension and acute anemia with hemoglobin 4.5 g/dL, without clinical evidence of bleeding. CT angiogram of the abdomen and pelvis with intravenous contrast showed “limited evaluation for gastrointestinal bleeding due to presence of enteric contrast” without active extravasation of blood. The patient had not received enteric contrast, raising concern initially for acute intestinal hemorrhage. On further evaluation by the team, radiology and interventional radiology, the “contrast material” was identified as being most consistent with sodium zirconium cyclosilicate, an oral potassium-binder that was new to our institution. It had been given for hyperkalemia in renal failure.

**Discussion:** In this case, the radiopaque material in the gut lumen was atypical for the appearance of blood but there was no obvious alternate explanation for the patient’s acute anemia. The explanation of enteric contrast did not fit the patient’s history, whereas the concern for intra-abdominal bleeding was high. The team and radiologist initially anchored on the hypothesis of acute bleeding as it fit best with the observed lab values and consulted with interventional radiology. This was a “near miss,” as the patient could have undergone invasive procedures to stop bleeding, exposing her to unnecessary risk and potential harm. Adoption of new treatment modalities and care fragmentation can create situations in which clinicians are at risk of misinterpreting data that affects management decisions. This team overcame these challenges by correlating with the patient’s clinical history, an important step that can be missed in acute settings.

Diagnostic Difficulties in Treating Atypical Case of a Patient Being a Doctor

T. Miyagami1, T. Harada2
1Juntendo University Hospital, Japan
2Showa University, Japan

**Learning Objectives:** It is difficult to diagnose patients who are also doctors. It is important to pay attention to overconfidence bias. Performing “zero-based thinking,” such as reflecting on the clinical course to debias the bias, can significantly reduce diagnostic errors.

**Case Information:** A healthy 30-year-old man presented to the emergency department complaining of fever, shivering, and abdominal pain after vomiting. Two days earlier, his family were diagnosed with infectious gastroenteritis. He also admitted having appetite loss, dysuria, and pain at perineum. On arrival, his body temperature 38.4?, other vital sign were almost within normal range. On physical examination, there was tenderness in the mid-lower abdomen without muscular defense or rebound tenderness. Laboratory tests showed a white blood cell count of 15.3×103 cells/μL (neutrophils: 94%) and C-reactive protein level of 5.3 mg/dL. His urinalysis was completely normal. The patient presumed that it was less likely for him to have acute appendicitis, given his recent exposure to infected contacts and his urinary symptoms. The physicians agreed with the patient’s differential diagnosis. However, his pain gradually shifted from the upper to the lower half of his abdomen within one hour of observation. Contrast-enhanced computed tomography (CT) diagnosed an acute appendicitis. The CT scan showed a small volume of pelvic ascites at the rectovesical pouch, possibly from intra-abdominal inflammation due to appendicitis, which may have irritated his urinary system and led to the urinary symptoms.
Discussion: In this case, we experienced diagnostic difficulty not only because of the confusing symptoms, but because the patient was a physician as well as our colleague. In such a challenging situation, re-visiting the patient history and physical examination results from the beginning is preferred, i.e. performing zero-based thinking to debias the bias can significantly reduce diagnostic errors.

Inappropriate Use of Resources During Global Pandemic

M. Ngassa, N. Al-Kourainy, D. Levine
Wayne State University School of Medicine, Detroit, MI

Learning Objectives: As the effects of COVID-19 continue to significantly impact healthcare globally, ensuring high value care and appropriate use of resources has never been more essential. Clinicians must remain vigilant to avoid known cognitive contributions to diagnostic error that may lead to unnecessary testing with the potential for patient harm. This case illustrates multiple forms of diagnostic error resulting in inappropriate use of duplicate imaging in the diagnosis of peripartum cardiomyopathy.

Case Description: A 30-year-old woman with a history of hypertension and obstructive sleep apnea developed dyspnea two days after uncomplicated cesarean delivery. Work-up included Computed tomographic (CT) pulmonary angiography (CT-PE) was unremarkable. Transthoracic echocardiography demonstrated ejection fraction of 45-50%; she was diagnosed with peripartum cardiomyopathy. Patient was discharged after five days on medications for heart failure including furosemide; she was unable to obtain her medications. Two days post discharge she complained “her lungs were filling up with water”. Upon readmission, she was tachycardic and hypertensive. Despite the prior negative CT-PE, she underwent repeat CT-PE, which was negative for PE. The patient was treated with diuretics with rapid improvement in symptomatology.

Discussion: This case represents errors in both clinical reasoning and diagnosis leading to inappropriate use of repeat radiation-intense imaging during a global pandemic. As per Graber, et al [Arch Intern Med. 2005;165:1493-1499], the following errors may be applicable to this case: availability error due to not considering other causes for her symptoms as the patient did not having her diuretics and faulty triggering as PE was considered without a change in symptomatology despite an alternative diagnosis of peripartum cardiomyopathy being more likely. These errors led to unnecessary repeat testing that ultimately did not change medical management and had the potential to lead to patient harm from unnecessary radiation exposure and contrast-induced nephropathy.

A Missed Case of Invasive Aspergillus Niger Found on Autopsy in a Presumed Immunocompetent Patient

M. Pham, B. Lublin, G. Block
University of Colorado Internal Medicine, Aurora, CO

Learning Objectives:
– Identify chronic alcohol use and malnutrition as strong risk factors for immunosuppression.
– Diagnose invasive bronchopulmonary aspergillosis in patients without a history of transplant, malignancy, or immunocompromised status.
– Highlight the value of autopsy to both medical education and cases with diagnostic uncertainty.

An 81-year-old man with a history of alcohol use disorder presented with two weeks of non-productive cough and fatigue without improvement despite finishing a course of azithromycin for pneumonia. He also reported ten-pound unintentional weight loss over several months. Blood pressure was 88/48. He was cachectic. Serum creatinine was 8.0 mg/dL, WBC 26.3 x/µL with 8% eosinophils. Chest radiograph showed a large right upper lobe peripheral consolidation. Initial shock improved with antibiotics and fluid resuscitation but on day three developed worsening shock requiring vasopressors and respiratory failure which precluded safe bronchoscopy without intubation. Computed tomography of the chest revealed a persistent right-upper lobe consolidation with central cavitation. Dialysis was indicated for renal failure;
however he declined continued aggressive interventions. Vasopressors were withdrawn and he died peacefully on day five. Autopsy revealed invasive Aspergillus Niger pulmonary infection with disseminated fungal elements in his kidneys and adrenal glands. Invasive bronchopulmonary aspergillosis (IBPA) is an often fatal, underrecognized cause of disease in critical illness. Chronic aspergillus infection classically presents with progressive respiratory and constitutional symptoms in addition to upper-lobe cavitary lung lesions. While commonly considered in patients with history of transplant or malignancy, there have been a few case reports in previously immunocompetent patients. An autopsy series revealed aspergillus in 2-4% of ICU cases. IBPA has a high mortality rate in critically-ill patients, and early recognition is key to treatment. This case highlights the importance of avoiding anchoring on bacterial sources of infection in a critically-ill patient.

Knee Dislocation: A Vascular Emergency

C. Pilcher
Evergreen Health, Kirkland, WA

Background: Patient falls, dislocating her knee. She is taken to a rural hospital where her knee is relocated after consulting an orthopedic surgeon at a distant hospital. Transfer is arranged (to a third hospital) for vascular studies. She arrives in the ED at shift change 5 hours after injury. Seen by oncoming ED physician 6 hours post-injury. Motor/sensory function variable. Distal pulses absent by Doppler. Vascular ultrasound is ordered. ED learns that no vascular surgeon is available, so transfer to Level I trauma center is begun. Departs the ED 9 1/2 hours post-injury. Foot is pale with no motor/sensory function; Doppler US is abnormal. Vascular repair completed at trauma center 13 hours post injury but amputation is eventually necessary. A lawsuit is filed. PLAINTIFF: Delays and miscommunication were rampant while my leg was dying. I needed immediate vascular surgery. You accepted a transfer without knowing if the care I needed was available. Because my circulation was not restored within 6 hours, I lost my leg. DEFENSE: Each defendant pointed fingers at the other, debated the transfer, whether the knee should ever have been reduced before surgery and the availability of the vascular surgeon. We all did the best we could under the circumstances. Even if we did everything perfect, it wouldn’t have mattered. RESULT: Settlement for undisclosed amount. TAKEAWAYS: * Recognize that a total knee dislocation is a vascular emergency. * Avoid focusing on how awful the joint looks. The danger lies in what is not seen. * The status of distal pulses or timing of reduction matters little. * Admission for monitoring distal circulation is mandatory. * Assure that all consultants and receiving hospitals understand the patient’s injury and needs. * Assure the accuracy of backup call schedules, especially if coverage is not 24/ * Shift change is dangerous if patient’s needs are under-appreciated. * A limb can survive for only about 6 hours with compromised circulation.
Acids, Bases, and Agitation: A Case of Acute Altered Mental Status

J. Raphelson, A. Jagannath
University of California San Diego Internal Medicine, San Diego, CA

Objectives:
- Recognize clinical signs of anticholinergic toxicity.
- Recognize Kussmaul breathing, a physical exam finding highly concerning for severe metabolic acidosis.

Case Information: A 70-year-old man with cannabinoid hyperemesis presented with acute altered mental status and worsened chronic vomiting. Temperature-101.2°F, pulse-112, respirations-28, BP-157/103. He was agitated, speaking nonsensically, with dry skin and mucous membranes, deep and rapid breathing with forceful expirations, and suprapubic distention. Laboratory tests revealed a bicarbonate 19mEq/L, anion gap 20, lactate 8.3mmol/L, venous blood gas pH 7.67/PCO2 16.1/PO2 97, and normal salicylate level. Chest x-ray unremarkable. Foley catheter produced 600 mL urine. After 4L IV fluids, breathing, mentation, lactic acid, and pH normalized. His wife later revealed patient started oxybutynin four days prior for overactive bladder, confirming a diagnosis of anticholinergic toxicity.

Discussion: Diagnosing the cause of altered mental status is challenging without history. In our patient, the acute onset of symptoms, anhidrosis and urinary retention, and anion gap acidosis with respiratory alkalosis made our team highly suspicious for a toxidrome[2]. However, it was not until we learned of his oxybutynin use that diagnosis became definitive. Acid-base status plays a significant role in drug toxicity. Our patient had both an anion gap metabolic acidosis from vomiting and a respiratory alkalosis. His delta-delta ratio was 1.6 which is highly suggestive of a lactate-driven acidosis due to lactate ion buffering intracellularly[1]. Acidosis potentiates effects of oxybutynin in particular by decreasing its affinity for serum proteins, resulting in higher concentrations of active drug[3,5]. The primary respiratory alkalosis likely resulted from Kussmaul breathing to correct for a prolonged period of acidemia[4]. The presentation of acute altered mental status with a mixed acid-base disorder should prompt consideration for ingestion-related toxidrome. Basic physical exam and laboratory findings greatly aid in making this difficult diagnosis.

References
ED. Urinalysis was abnormal, and he was treated with nitrofurantoin. Symptoms returned 10 days later. He was given an additional week of nitrofurantoin from the physician-father of a friend and ultimately visited the clinic. Urinalysis after completing nitrofurantoin remained abnormal. He was started on trimethoprim/sulfamethoxazole. Ten days later, he had an adverse reaction. He was treated with fosfomycin with symptom resolution.

**Discussion:** Physicians are often required to initiate empiric therapy for infection. Knowledge of local resistance patterns is necessary for optimal choice of therapy. E. coli resistance to ciprofloxacin in the patient’s area was 20-40% making it a poor empiric choice. As it turns out, he grew an ESBL E. coli which is uniformly ciprofloxacin resistant. Follow-up on culture results is crucial. Failure to follow-up on tests from the ED is not uncommon. In our patient, this resulted in increased costs and delayed improvement. The patient did not have a PCP and did not realize the rationale for follow-up. Arranging a visit with a new PCP when someone is ill is a daunting task. Use of teach-back communication and identifying barriers to follow-up are necessary to ensure good clinical outcomes.

---

**IgG4-Related Disease: A Great Mimicker**

J. Welkie, A. Hyczko, N. Olsen  
Penn State College of Medicine, Hershey, PA

**Learning Objectives:** 1. Describe common and uncommon causes of SVC syndrome. 2. Recognize the variability in clinical presentation, organ involvement, and histopathology of IgG4-Related Disease and how this may lead to a delayed diagnosis. 3. Understand the benefits of inter-institutional collaboration when diagnosing rare diseases.

**Case Information:** A 26-year-old male with a two-year history of intermittent progressive difficulty breathing presented to an outside hospital with worsening purple discoloration and swelling of his chest, upper extremities, and face. Outside lab tests were unable to be reviewed but were described as unremarkable. Imaging revealed a mediastinal mass encasing the superior vena cava (SVC). The initial pathology from the outside hospital was interpreted to be consistent with a ganglioneuroma. The patient was then transferred to our hospital for surgical removal of the mass. The pathologist reviewed the initial biopsy, noting necrosis, fibrosis, and inflammation, and suggested further evaluation was needed for accurate diagnosis. Following surgical resection, an inter-institutional consultation service was used to evaluate the pathology of the mass, which was found to be consistent with IgG4-Related Disease (IgG4-RD).

**Discussion:** SVC syndrome occurs when compression or invasion of this large vessel by a pathologic process impedes venous return to the heart, leading to swelling of the head, neck and arms, dyspnea, and cyanosis. More common etiologies of SVC syndrome include malignancy, fibrosing mediastinitis, and thrombosis. IgG4-RD, a rare cause of SVC syndrome, is an immune-mediated fibroinflammatory disorder caused by an infiltration of CD4+ cytotoxic T cells, resulting in enlargement and fibrosis of various organs. Diagnosis is complicated by a high degree of variability in clinical presentation and histopathologic characteristics and, as in this case, may be delayed. Given the histopathologic complexity of IgG4-RD, inter-institutional collaboration, as highlighted in this case, may be beneficial in accurately diagnosing this disease.
Deception: Septic Pelvic Thrombophlebitis Mimicking Symptoms of Postoperative Infection

R. White¹, C. White²
¹Wayne State University School of Medicine, Detroit, MI
²Henry Ford Macomb, Clinton Township, MI

**Learning Objectives:**
1) Identify septic pelvic thrombophlebitis as an alternative cause for increased white blood cell count in patients who are in a postoperative recovery period
2) Recognize how the nature of thrombus growth may necessitate repeating imaging studies for diagnosis.

**Case Information:** A 42 y/o female, 11 days status post laparoscopic hysterectomy, presented to the hospital with complaints of diffuse abdominal pain for 5 days. An initial workup revealed a WBC count of 16.1 K/uL so the patient was admitted. Given her postoperative state, an infectious disease consult was ordered along with a CT scan. Consult tested the patient for influenza A/B antibody and RSV IgM. These were both unrevealing. Her CT scan was likewise unremarkable. She was empirically treated with IV Clindamycin and Gentamicin for 5 days for suspected postoperative infection. Yet, over the course of the IV treatment, the patient clinically deteriorated. Further workup tracked a continually increasing WBC count, PT, and PTT time. Her physician transferred her to the ICU in anticipation of repeat surgery and ordered a repeat CT. It revealed a septic right ovarian thrombosis not previously visualized. She was started on Lovenox 60mg BID as anticoagulation therapy. Her labs returned to within normal limits and she was discharged.

**Discussion:** Septic pelvic thrombophlebitis is a rare, but potentially fatal complication associated with a postpartum and postoperative state. Diagnosis is made traditionally via laparotomy, but the popularity of CT imaging continues to increase. Patients typically present with pelvic pain, fever, and right sided abdominal mass. Since symptoms are nonspecific, clinical suspicion should be high in at risk populations, emphasizing the importance of a thorough physical exam and imaging prior to more invasive action.

Reframing for the Rickettsia Diagnosis

M. Yang, J. Nguyen, C. Nguyen
Baylor College of Medicine, Houston, TX

**Learning Objectives:** Overcome anchoring bias from physician handoff through cognitive reframing. Incorporate new data in the context of endemic diseases.

**Case Information:** A 25-year-old woman presented with one week of fever, myalgia, headache, nausea, and vomiting. During that time, she developed hypotension and was started on low-dose norepinephrine and empiric “broad-spectrum” antibiotics (vancomycin and cefepime) for suspected meningitis. Initial work-up was only remarkable for mild anemia and thrombocytopenia of 21 cells/µL. She was then transferred overnight to a tertiary referral facility due to concern for thrombotic thrombocytopenic purpura (TTP) and need for intensive care and plasmapheresis. Upon admission to the medical intensive care unit (MICU), patient was noted to have a diffuse blanchable maculopapular rash in the extremities and abdomen. TTP was ultimately ruled out as there was no laboratory evidence of hemolysis. Given the patient’s symptoms and knowledge of endemic infections often seen in Texas, the possibility of Rickettsial diseases was considered. The patient was ultimately started on doxycycline. The patient improved and antibody test for Rickettsia typhi returned positive one week later.

**Discussion:** The diagnosis of the patient was delayed due to various factors. There was anchoring bias with the handoff from the outside hospital given to the MICU, despite the lack of objective evidence for TTP. The pressure to coordinate with the hematology service and to start plasmapheresis led to the premature closure of the decision-making process. Furthermore, the term “broad-spectrum antibiotics” seems to provide a false sense of security in management of critically-ill patients. Reframing the patient as a pet owner with a newly identified diffuse maculopapular rash, severe flu-like symptoms, and thrombocytopenia could have led to the correct diagnosis earlier. This case highlights the importance of incorporating new subjective and objective data into the history and considering alternative diagnoses to avoid anchoring bias.
Immersed in the Evidence - A Training Program for Laboratory Pathology Professionals

R. Hardie¹, J. Li¹, C. Campbell², W. Punyalack³, A. Georgiou¹, Q. Lam⁴, A. R. Horvath⁵
¹Macquarie University, Sydney, AU
²NSW Health Pathology, AU
³RCPAQAP, St. Leonards, AU
⁴Austin Pathology
⁵NSW Health Pathology, AU

Purpose/Problem: Standardized alert procedures are crucial, especially for pathology test results requiring immediate action due to risk of major patient harm or imminent death. There is currently a lack of outcomes-based evidence supporting existing high-risk result thresholds for pathology testing and alerting. We aimed to engage pathology laboratory professionals in the research process in the evidence-based laboratory medicine field. We developed a training program, designed so that teaching and support is provided to participants to enable them to complete systematic literature reviews. This process will supply evidence to support harmonized high-risk result thresholds and improve diagnostic reliability and accuracy. The results from this study are being translated by The Royal College of Pathologists of Australasia and Australasian Association for Clinical Biochemists and Laboratory Medicine High Risk Results Working Party (RCPA-AACB HRR WP) to produce a best practice guidance document supporting the implementation and uptake of recommended alert threshold values.

Description of Program: The program opened with a face-to-face workshop. Participants from across Australia and New Zealand convened, where they were immersed in problem-based and interactive learning. Participants were each assigned an analyte and learning team. This was followed by 12 weeks of distance learning (online modules/video-conference workshops) and self-directed learning modules. In each module, participants gained the skills to perform each step in the systematic review process according to PRISMA guidelines. Deliverables included database search strategies, PRISMA flow diagrams, EndNote libraries, and a summary systematic literature review.

Outcomes: Systematic literature reviews were produced by laboratory professionals, in partnership with Macquarie University researchers. A summary document of all analytes was produced in partnership with the WP, to be used in their best practice guidance document for recommended HRR. Discussion: This program has resulted in several key innovations: 1) A framework for long-term involvement of participants in a translational research project from literature search, to dissemination, to policy; 2) capacity building for both researchers engaged in course development and delivery, and laboratory participants in developing research and knowledge synthesis skills and; 3) a model for relationship building between academic institutions and professional bodies in the areas of clinical chemistry and pathology.

Significance of Findings: Empowering health professionals in the process of research allows for truly translational work from conception to practice. This project provides not only vital evidence for the creation of new harmonized alert thresholds, but also demonstrates genuine collaboration in the effort to successfully generate much-needed evidence for effective policy.

Learning and Speaking a New Language: A Clinical Problem-Solving Elective for Medical Students

V. Phadke, J. Spicer, M. Lora
Emory University, Atlanta, GA

Purpose/Problem: Clinical reasoning is a core skill for clinicians but is often not taught explicitly in medical school. The transition to clinical rotations is an ideal moment to develop foundational knowledge in reasoning concepts and foster habits that may mitigate diagnostic error.
Description: We created a four-week elective in clinical problem-solving for third-year medical students. The course aims included: 1) Familiarizing students with core concepts in clinical reasoning, 2) Developing learning strategies that emphasize knowledge organization, and 3) Providing forums for deliberate practice. We used flipped classrooms to teach reasoning concepts, using online resources/readings to supplement didactics. Students worked in small groups to solve weekly cases, submitting daily iterative problem representations and differential diagnoses after each aliquot. In-class time involved interactive sessions with unknown cases presented to expert discussants, followed by group reflection on pivot points or potential sources for error. We distributed an end-of-elective survey to measure impact on learner satisfaction, knowledge, and behaviors.

Outcomes: Forty-one students enrolled in the course, and 100% completed the survey. Students spent an average of 3 hours/week completing course readings, 4 hours/week reading independently about diseases/schema, and 6 hours/week on small group work. Most (95%) said they would be “extremely likely” to recommend the elective to others. Based on retrospective pre-post self-assessment of knowledge, students reported improvements in familiarity with core reasoning concepts - on a five-point Likert scale (“not at all familiar”=1, “extremely familiar”=5), the mean score for problem representation increased from 2.71 to 4.87, for illness scripts from 3.08 to 4.68, for diagnostic schema from 2.95 to 4.66, and for Bayesian reasoning from 2.16 to 3.82. Students reported increases in their intended frequency of reasoning behaviors - on a five-point Likert scale (“never”=1, “always”=5), the mean score for developing a problem representation before constructing a differential diagnosis increased from 2.18 to 4.58, for reading horizontally to refine illness scripts from 2.11 to 4.18, and for utilizing cognitive forcing strategies to avoid bias/error from 1.79 to 4.26.

Discussion: Our elective successfully developed students’ fluency in clinical reasoning concepts and demonstrated their utility in the diagnostic process. Incorporating a flipped classroom approach and opportunities for deliberate practice enabled us to translate abstract concepts into practical knowledge and skills.

Significance: Few published curricula for clinical problem-solving exist for undergraduate medical education. Our elective could serve as a model for developing modules or longitudinal courses in clinical reasoning to develop these skills in early learners.

Development of a Machine Learning Algorithm for Feedback on Clinical Reasoning Documentation

V. Schaye1, B. Guzman2, D. Kudlowitz1, J. Rafel1, M. Marin1, I. Reinstein1, L. Miller1, J. Chun1, P. Cocks1, Y. Aphinyanaphongs1

1NYU Grossman School of Medicine, New York City, NY
2Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY
3Stanford University School of Medicine, Stanford, CA

Purpose/Problem: Clinical reasoning (CR) is a core component of medical training, yet residents often receive little feedback on their CR documentation. Here we describe the process of developing a machine learning (ML) algorithm for feedback on CR documentation to increase the frequency and quality of feedback in this domain.

Description of Program: To create this algorithm, note quality first had to be rated by “gold standard” human rating. We selected the IDEA assessment tool—a note rating instrument across four domains (I=Interpretive summary, D=Differential diagnosis, E=Explanation of reasoning, A=Alternative diagnoses) of CR documentation that uses a 3-point Likert scale without descriptive anchors. To develop descriptive anchors we conducted an iterative process reviewing notes from the EHR written by medicine residents and validated the revised tool using components of Messick’s framework—content validity, response process, and internal structure. Using the Hofstee standard method, cutoffs for high quality clinical reasoning for IDEA score and DEA score was set. We then created a dataset of expert-rated notes to create the ML algorithm. First, a natural language processing software was applied to the set of notes that enabled recognition and automatic encoding of clinical information as a diagnosis or disease (D’s), a sign or symptom (E or A), or semantic
qualifier (e.g. most likely). Input variables to the machine learning algorithm included counts of D’s, E/A’s, and semantic qualifiers. ML output focused on DEA quality and was binarized to low or high quality CR.

Outcomes: The revised IDEA and DEA scores ranged from 1-10 and 0-6, respectively. An IDEA score of = 6.5 and a DEA score of = 3 was deemed high quality. 252 notes were rated to create the dataset and 20% were rated by 3 raters with high intraclass correlation 0.84 (95% CI 0.74-0.90). 120 of these notes comprised the testing set for ML model development. The logistic regression model was the best performing model with an AUC 0.87 and a positive predictive value (PPV) of 0.65. 48 (40%) of the notes were high quality notes.

Discussion/Significance of Findings: Next steps are to validate the ML algorithm by comparing its output to human rating. This validation process will help determine whether further training of the model is required. Subsequently we will pilot using this ML algorithm for feedback on residents’ CR documentation which we hypothesize will increase the quality of CR documentation.

Practice Improvement

Virtual Rounding - Improving Diagnosis Through Telemedicine Services on Inpatient Teaching Teams
B. Mathews, A. Doshi
HealthPartners, Bloomington, MN

Statement of Problem: During the COVID-19 pandemic, residents and residency programs find themselves in a unique situation. Balancing the educational needs of a training program with the safety of trainees is a challenging task, specifically when taking care of patients who are COVID-19 positive or patients under investigation (PUI). One increasingly available tool that can help protect trainees while continuing to prioritize patient care and medical education is the use of telemedicine for virtual rounding. For our residents through the University of Minnesota rotating at Regions Hospital in Saint Paul, MN, we have used video visits to continue our mandate as both healthcare and education professionals.

Description of the Intervention or Program: Virtual care can mitigate exposure risk, minimize use of personal protective equipment (PPE) and improve communications with patients and their families. To guide our teaching teams on the right situations for telemedicine, we needed to select those patients who would be most appropriate for a virtual visit. Further, we implemented a simple decision tree (Figure 1). First, the team needs to decide whether the patient needs an immediate in-person assessment; for instance, critically ill patients or those who need end-of-life care discussions, telemedicine would not be an appropriate modality. Next, the decision needs to be made whether a patient requires an in-person exam at that time.

Findings to Date: These are very preliminary data and additional domains are being analyzed and will be updated. So far, by mid-June, 2020 telemedicine was used in 14% of the patient encounters on resident teams and in 60% of COVID-19 patients. Telemedicine usage was highest by patients ages 25-44, and least in 65+. Interpreters were regularly used with non-English speaking patients (45% iPad interpreters, 30% in-person interpreters). Virtual follow up with families of English-speaking patients (57%) and non-English families (8%).

Lessons Learned: Prior to the COVID-19 pandemic, telemedicine adoption was limited due to lack of awareness, barriers in training, understanding, and narrow beliefs regarding the innovation. Telemedicine should be a developing part of the education and training for healthcare learners. Ensuring the growth of telemedicine during COVID-19 and beyond does not worsen disparities in care is also key. This pandemic has underscored the need for providing telemedicine services that will likely long outlast this crisis and to support our healthcare learners in being effective on our care teams.
Breaking Bread to Improve Diagnosis: Implementation of a “Cases and Conundrums” Conference Among Early Career Faculty

J. Matulis, S. Barakat
Mayo Clinic, Rochester, MN

Statement of Problem: Time constraints, EHR work-load and increasing patient complexity are well known challenges to the diagnostic process in outpatient internal medicine practice. These same challenges, shared among colleagues, also limit a clinician’s ability to readily seek the counsel of peers in diagnostic decision making.

Description of the Program: We started an outpatient case conference for early career internal medicine clinicians with an aim of creating a safe space to seek opinions on diagnostic decision making and management. The rationale for this program was that crowd sourcing of diagnostic decision making can facilitate more timely and accurate diagnosis, resulting in improved patient care. Additionally, building a community around diagnosis among like-minded clinicians may improve participant’s sense of professional community, feeling of support, and joy in work.

Findings to Date: 16 early career clinicians (physicians, nurse practitioners and physician’s assistants) agreed to participate in the conference series. In baseline data collection, 15 of 16 participants agreed that their work was meaningful to them, but the median sense of belonging within the division was 6 out of 10 and the median rating of enthusiasm...
for working in the division was 7 out of 10 (using a 0-10 point Likert scale). The majority of respondents indicated one motivation for participating in the conference series was the opportunity to interact with and get to better know their peers. Conference data: 3 separate conferences were conducted and 9 diagnostic conundrums were presented prior to suspension of in-person gatherings due to Covid-19. Management decisions were altered after conference feed-back in 4 of the 9 cases presented. For each case, 1 or 2 clinical pearls were shared with the group. The conference will resume this summer when restrictions on in-person meetings are lifted.

**Lessons Learned:** A safe space for peers to meet and review cases can improve diagnostic decision making and sense of community among early career clinicians. Keys to success include administrative support, achieving the right group size and composition to allow for open dialogue, desirable food options based on group input, and a semi-formal structure for presentations which ultimately summarize and share learning for the entire group.

---

**Developing a Culture of Learning from Diagnostic Errors among Frontline Clinicians**

D. Upadhyay¹, D. Torretti², C. Schuerch², Y. Mehta², A. Bansal², S. Korukonda¹, P. Lamphear¹, H. Singh³

¹Geisinger, Danville, PA
²Geisinger Medical Center, Danville, PA
³Baylor College of Medicine and Michael E. DeBakey VA Medical Center, Houston, TX

**Problem:** Clinicians are uncomfortable reporting diagnostic errors or sharing them widely often due to fear of legal and professional repercussions and an unsupportive institutional safety culture.

**Intervention/Program:** At Geisinger, the Committee to Improve Clinical Diagnosis (CICD) and the Safer Dx Learning Lab are supporting clinical excellence through a program that seeks to reduce diagnostic errors while promoting a culture of safety and learning. We encouraged all clinicians to share diagnostic errors using a dedicated reporting mechanism which included a hotline (phone), EHR-based secure messaging system (email), and paging (text). A system-wide clinician engagement initiative was launched in July 2019. Cases were discussed at M&M and other department meetings. Directors and chief residents championed the cause and sought cases from their teams.

**Findings:** From February 2015 to May 2020, 101 cases were shared by clinicians which account for 20% of cases in our database. More than half of these (n=58; 57.4%) were reported just in the previous 10 months, soon after the initiative’s launch. Most cases were reported by hospitalists (n=17; 16.8%), residents (n=13; 12.9%) and intensivists (n=9; 8.9%). Pediatricians, emergency medicine physicians, rheumatologists and pathologists reported 8(7.9%) cases each. Of the 101 cases, 76 (75.2%) were determined to have missed opportunities in making a timely and accurate diagnosis. Most opportunities were found in clinical encounters (n=58; 76.3%) which includes problems with history taking, physical examination, assessment and considering relevant differential diagnosis. Other opportunities were related to interpretation of imaging and lab tests (n=14; 18.4%) and inappropriate follow-up of abnormal test results (n=11; 14.5%). Emergency medicine (n=27; 35.5%), hospital medicine (n=21, 27.6%) and primary care (n=14, 18.4%) were where most of the reported opportunities occurred. Other departments included radiology (n=7; 9.2%), critical care (n=5; 6.6%), surgery (n=4; 5.2%), pathology (n=3; 3.9%), orthopedics (n=3; 3.9%) and 13 other departments (each n =<2). Several opportunities were reported: cancer (n=12; 15.7%), septic joints (n=6; 8%) and strokes (n=6; 8%) in addition to sub-dural hematomas, sepsis, appendicitis, cardiac tamponade and diabetic ketoacidosis (all n =<3).

**Lessons Learned:** Diagnostic errors reported by a variety of clinicians included different conditions in multiple departments. Preliminary findings indicate that frontline clinicians may be more likely to report breakdowns in the diagnostic process if they are constantly engaged in a non-punitive and constructive manner. Developing a network of champions appears essential for establishing a safe environment to discuss and learn from diagnostic errors in an open and transparent manner.