Spring 5-2022

Identifying Risk Factors for Anchoring Bias during Emergency Department Transitions of Care

Roni Matin

Identifying Risk Factors for Anchoring Bias during Emergency Department Transitions of Care

By

Roni Matin, M.Sc.

APPROVED:

_____
Susan Fenton, PhD, Chair

_____
Amy Franklin, PhD

_____
Robert Murphy, MD

_____
Amit Mehta, MD

Date approved: _____

Identifying Risk Factors for Anchoring Bias during Emergency Department Transitions of Care

A

Dissertation

Presented to the Faculty of
The University of Texas
Health Science Center at Houston
School of Biomedical Informatics
in Partial Fulfilment of the Requirements for the Degree of

Doctor of Philosophy

By

Roni Matin, M.Sc.

University of Texas Health Science Center at Houston

2022

Dissertation Committee:

Susan Fenton, PhD[1], Advisor
Amy Franklin, PhD[1]
Robert Murphy, MD[1]
Amit Mehta, MD[2]

[1]The School of Biomedical Informatics
[2] McGovern Medical School

Dedication

To my late parents Abdul and Jahanara Matin for their constant love, support and all the sacrifices they made to allow me the opportunities to pursue my dreams and for teaching me the importance of education, compassion, integrity, and hard work.

To my late husband, my beloved David for our time together filled with love, laughter, music and adventure and for always believing in me, supporting me and encouraging me to invest in my education and my development.

I miss you all so, and wish you were here to share in this moment.

Acknowledgements

Firstly, my deepest appreciation goes to my committee, whose advice and support were critical to the completion of this dissertation. I would like to give special thanks to my advisor Dr Susan Fenton, for her continuous steadfast guidance, attentiveness, and encouragement to help me stay focused and determined to complete this work. Similarly, I would like to thank Dr Amy Franklin for always generously giving advice and support throughout this endeavor. Her enthusiasm, brilliant insights and vast knowledge and expertise in human factors, cognitive engineering and research design have not only helped guide this dissertation research but have helped me grow and develop as a researcher. I would also like to thank my committee members Dr Amit Mehta and Dr Robert Murphy for their insightful and essential guidance, particularly on the clinical relevance of this work. I am deeply indebted to Dr Mehta for always making time to help and advise me about medical data and clinical decision making, and for directly facilitating my research with McGovern Medical School.

Through this entire process I have gained an appreciation of the skills, perspective and rigor required to conduct research in the fields of biomedical informatics and healthcare delivery.

I have been fortunate to have had the advice and support of a special group of physicians. Thank you, Dr Nathan Hoot, Dr David Robinson Dr Charles Maddow, Dr Kimberly Chambers and notably Dr Brent King, who enabled and championed my research with the University of Maryland Baltimore Medical School. Their generosity with their time and advice was essential to the success of this work. Also, thanks to those who participated in my studies and made this research possible.

Also, I would like to make an honorable mention for my high school teacher, Mr Walkinshaw, for first instilling in me the love and wonder for science and for being a great teacher.

I must thank my friends, colleagues, and faculty especially, Sarah May and Dr Kimberly Smith for their enduring encouragement and support. Also, thank you to Laila Rasmy Bekhet, Nina Smith and Carlos Aldana, for their friendship and humor. Thank you also to Dr Deevaker Rogith, and Dr Tiffany Champagne-Langabeer for their advice and encouragement.

I would also like to thank the staff at SBMI; Robyn Cavalier-Brown, Erica Unice, Miranda Hedrick, Emin Saglamer, Marcos Hernandez, Camelia Addison, and Joe Menchaca for the help and support they have given during my time at SBMI.

Finally, I would like to thank my children, Joshua, Natalie, and Jacob, for their constant love, support, understanding and encouragement throughout this dissertation. They are my inspiration for embarking on this journey and for everything I do. I hope one day they can reflect on this time and know that they too can accomplish their dreams.

Abstract

Transitions of care have been associated with breakdowns in communication and medical errors. In emergency departments (ED) these handoffs are typically known as sign outs. Sign outs provide continuity of care for ED patients whose diagnosis and care fall across shift changes. They are short interactions where pertinent information and responsibility for the patient is transferred to the physician assuming care for them. However, these exchanges may also be an opportunity for cognitive biases to be transferred or introduced, leading to erroneous decision making. Anchoring bias is known to have a significant impact on clinical decision making. Yet, little is known of the factors that increase the risk of anchoring bias during patient diagnoses that involve sign outs.

This exploratory research aims to understand how the communication of patient information during sign out influences the clinician's use the information and develop the patient's diagnosis and thus identify the factors that contribute to anchoring bias in clinical decision making in the ED.

A mixed method approach was used to identify and evaluate potential risk factors for anchoring. Initially a review of a dataset from a medical incident reporting system was conducted to identify potential contributing factors from known cases of medical error. This was followed by an interview study with emergency medicine (EM) physicians to gain their perspectives on peer influence and communication factors between outgoing

and oncoming clinicians that might affect sign outs and thus potentially impact decision making.

The findings were used to design an experimental evaluation study to assess the impact of potential risk factors identified on diagnostic and treatment planning of EM clinicians. The study was conducted using patient case vignettes as control cases and stimuli cases, which contained these risk factors as test conditions to assess their effect on clinical decision making. The cases were presented in a format simulating sign out communications and the volume of information presented at sign out. Volume of information was represented by the two test conditions of explicitness of the sign out information and the stage in the diagnostic process the case was in at the time of sign out. The study was conducted at two academic hospital ED sites with a total 69 participants. The results indicated that the explicitness of the sign out information had no significant influence on the diagnostic accuracy in stimuli cases or on the confidence of the clinician participants in their diagnosis for the case. However, the stage in the diagnostic process of the case at the point of sign out, did significantly influence both clinicians' diagnostic accuracy and their confidence in the diagnosis. The earlier stage stimuli cases were associated with lower diagnostic accuracy and lower confidence in the diagnosis. The test condition of explicitness did not have a significant effect on a number of outcome measures whereas the test condition of stage of the case did not.

These findings suggest that additional support may be required for during sign out for cases that are in an earlier stage in the diagnostic process at the time of sign out to as they are at higher risk for diagnostic error and for the influence of anchoring bias.

Vita

1989.....BSc. Honours, Immunology,
University of Glasgow, United Kingdom

1996.....MSc., Information Systems Design,
University of Westminster, United Kingdom

2015 to present..... PhD Student, School of Biomedical
Informatics, University of Texas Health
Science Center at Houston, TX

2016-2018..... Graduate Research Assistant, School of
Biomedical Informatics, University of Texas
Health Science Center at Houston, TX

2019-2021..... Graduate Teaching Assistant, School of
Biomedical Informatics, University of Texas
Health Science Center at Houston, TX

Field of Study

Biomedical Informatics

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Most emergency departments (ED) receive patients throughout the day and night, all days of the year. The need to provide continuous medical care, necessitates transitions of care between shifts of emergency medicine professionals. Most Eds involve two to three transitions of care daily between their medical teams, with the off going team providing a handoff to the oncoming team. Transitions of care between ED physicians at shift changes typically involve verbal handoffs, known as sign outs.

The ED is a pressured environment. Patients often present with serious conditions and limited medical history information available. The treating physician follows a process of examinations, tests, and imaging studies to work through a set of differential diagnoses to ultimately determine the patient's diagnosis. In addition, emergency medicine staff typically manage multiple patients in varying states of acuity. (Apker et al., 2007; Horsky et al., 2015) Patients present in the ED often in critical condition, requiring urgent care and attention. In such situations ED physicians must expediently assess the patient and promptly develop a working diagnosis and plan for the next steps of treatment. Their care involves diagnostic testing, monitoring and rapid decision making and is often complicated by a lack of patient information. (Horsky et al., 2015; Royce et al., 2019) During any given shift ED physicians must manage a diverse patient case load, while having to operate within high levels of uncertainty due to the nature of the patients coming to the ED. Emergency medicine providers must cope with demands on their memory, high cognitive load, and fatigue, as the ED can be subject to unpredictable periods of high patient volumes. Then at the time of sign out, the outgoing physician

must provide a brief but informative synopsis of their diagnostic findings and patient-specific details for each patient to their counterpart coming on duty to take over care for the patients.

1.1Transition of Care and ED Sign Outs

Due to the high patient volumes and the complexity of patient cases presented in the ED, sign outs have been recognized as important point of risk for patient safety. Emergency care settings have been linked to errors in diagnosis (V. Arora et al., 2005; Okafor et al., 2016). Analysis of medical malpractice case data indicates that 65% of ED medical malpractice claims involved missed diagnoses. In addition, up to 24% of ED malpractice claims suggested poor patient handoffs as a contributing factor. (Dhingra et al., 2010; Kachalia et al., 2007) Studies from The Joint Commission (TJC) found that 70% of medical errors involved a breakdown in communication and of these errors, 50% occurred during transitions of care. Indeed TJC studies found 80% of serious medical errors were due to faulty transitions in care. (*Sentinel Event | Joint Commission*, n.d.) There have been efforts to improve patient safety within transitions of care. A number of studies have been conducted investigating the processes and communication formats that occur during critical care handoffs. (Abraham et al., 2011; Abraham, Kannampallil, Brenner, et al., 2016; Abraham, Kannampallil, Patel, et al., 2016) While sign outs in the ED are similar to ICU handoffs, in terms of transfer of information and care for the patient, there some significant differences. ICU handoffs involve multi-disciplinary teams and are conducted over longer periods of time. Whereas in the ED, sign outs involve the outgoing physician verbally signing out their patients to the oncoming physician. In the interest of minimizing time away from patient care, most ED sign outs last around 20 minutes. However, some can be longer due to sudden increased patient volumes, such as

in large regional hospitals. Typically, between 10-20 patient cases are transferred during this relatively brief interaction, meaning that the time for each patient is typically no more than a couple of minutes. (Murphy RJ, MD, oral communication, 05/31/19, 13:30; Mehta A., MD, oral communication 02/23/22 13:20) During this time the exiting physician must transfer all pertinent information and responsibility for the patients under their care, to the physician coming on to their shift. Sign outs provide the mechanism to provide continuity of care for those patients whose diagnosis or the urgent treatment needed to stabilize them, extends over a change of shift. This could be either due to the patient being admitted close to the change of shift or to a protracted diagnostic process. Consequently, for some of the patients in the sign out list a confirmed diagnosis my not be known at the time of sign out. (Cheung et al., 2010; Frye et al., 2018; Sullivan et al., 2015)

Sign outs are invariably conducted verbally and the use of a standardized handover protocol, are not routine in all hospitals. (Cheung et al., 2010; Dhingra et al., 2010) Studies indicate significant information loss occurs during verbal transitions of patient care. As such, with the limited time available to transfer all the pivotal patient information in the case means that ED sign outs represent a significant risk for patient safety. (V. Arora et al., 2005; Frye et al., 2018)

There have been calls to develop oral sign out skills curricula as part of medical training (Horwitz et al., 2007). As such, all emergency medicine residency programs are mandated by the Accreditation Council for Graduate Medical Education (ACGME) to ensure that residents are competent in sign outs communication. In addition, there have been efforts to standardize the sign out process. Some Eds have adopted structured handoff protocols based on templates such as SBAR (situation, background, assessment)

and another template often used is IPASS (illness severity, patient summary, action list, situation awareness and contingency planning, synthesis by receiver) (Sullivan et al., 2015) However, studies show that such protocols can make sign outs significantly longer to conduct, which could affect how consistently they are used in practice. (Dhingra et al., 2010; Heilman et al., 2016; Tews et al., 2012). The Safer Sign Outs protocol recommended by The American College of Emergency Physicians (ACEP) may be more appropriate for use in ED transitions of care, having been developed specifically to consider the needs and constraints of ED settings and ED sign outs. Many EDs however, do not require the use of protocols, but with or without protocols, sign outs are typically conducted in the chronological order of the beds on the ED floor. Hence critical patients may not be prioritized for acuity and their information might be discussed somewhere down the order of the sign out. Additionally, while protocols may provide guidelines on the overall structure and good practices, they don't offer advice on the specific content and information details within a sign out. Consequently, in addition to the potential for communication gaps, sign outs also pose a risk for the introduction of cognitive mistakes that may contribute to diagnostic errors. (V. Arora et al., 2005; Dhingra et al., 2010; Okafor et al., 2016)

There have been important advances in efforts to reduce medical errors and they have proved very successful at identifying and reducing system errors. (Singh & Sittig, 2015; Menon et al., 2017) However, tackling cognitive errors has proved challenging, as they are not easy to identify or address. (Singh, 2014; Singh et al., 2017) Studies indicate that in cases of diagnostic error, cognitive factors contributed to errors in the majority of cases. (Graber et al., 2005) More specifically, studies into cases of diagnostic error in the ED, suggested that while system factors were involved in some cases, cognitive factors

were involved a far greater proportion of the cases. (Graber et al., 2005; Okafor et al., 2016)

1.2 Cognitive Biases and Anchoring Effect

Clinical decision making and diagnostic reasoning are extremely complex processes. It involves the integration of many data elements in terms of clinical signs and symptoms related to the patient presentation and tests results. This information then must be reviewed against mental schemas for disease conditions, clinical guidelines and protocols to rule out or rule in different potential diagnoses for the patient. As such clinicians often use a combination of analytical reasoning and unconscious heuristics when diagnosing patients. (Chapman & Sonnenberg, 2003a; Mohan et al., 2017) Heuristics are 'rules of thumb' that are based on pattern recognition and enable fast, almost instantaneous decisions making. (Tversky & Kahneman, 1974) In clinicians, these heuristics may be developed as implicit knowledge gained from the experience of treating countless patients over time. (Croskerry, 2013; Croskerry & Norman, 2008; Jenkins & Youngstrom, 2016; Mohan et al., 2017) These unconscious heuristics, while frugal in terms of mental effort, can have a significant influence in decision making. In most circumstances, heuristics can be useful in terms of decision making mental shortcuts. However, these heuristics when not well calibrated to the conditions may lead to erroneous assumptions known as cognitive bias and these biases can lead to systematic errors. Thus, in clinical settings these automatic cognitive processes have the potential to impact clinical reasoning. (Evans, 2006; Boyle, 2014)

The information constraints, time pressures, and potential for communication breakdowns related to sign outs make them vulnerable to the introduction of cognitive biases. Also, their verbal nature, coupled with their often unstructured format, could

influence the amount and quality of information transferred about each patient. These factors may increase the risk of potentially introducing cognitive biases within the transfer of information from the exiting physician to the physician receiving the sign out. (Cheung et al., 2010; Frye et al., 2018; Long, 2015) One bias known to influence clinical decision making is anchoring bias. In diagnostic decision making, anchoring bias could be thought of as the tendency to lock onto salient features of the case too early in the diagnostic process and then failing to adjust this initial diagnostic impression. (Croskerry, 2003; Mull et al., 2015) In sign out the patient case is transferred often with the diagnostic process already initiated. So, in some circumstances there is a possibility that information passed during the sign out could have an anchoring effect on the diagnostic reasoning of the oncoming physician. Systematic reviews of literature on cognitive biases indicate that anchoring bias is one of the most prevalent biases in medical decision making. (Furnham & Boo, 2011; Saposnik et al., 2016) One study reviewing a case of serious diagnostic error found that many factors had influenced the process and introduced an anchoring bias during the ED sign out. (Campbell et al., 2007) Anchoring has been identified as a potential contributory factor in many cases of medical error, often manifesting in the form of either wrong diagnosis or severely delayed diagnosis.(Keeney & Halalau, 2017; van Geene et al., 2016)

The influence of cognitive biases has been studied in many fields and anchoring bias is recognized as highly influential in decision making. (Clegg et al., 2015; Mussweiler & Englich, 2005; Tversky & Kahneman, 1974, 1986) It has been studied in many diverse fields including sales negotiations (Galinsky, 2001), marketing (Wansink, 1998), and courtrooms (Davis & al, 1984). In healthcare, the anchoring effect is recognized as a prevalent and persistent cognitive bias (Augestad et al., 2016; Garcia-Molina & Chicaiza

Becerra, 2015) that resists de-biasing efforts. Proclivity towards anchoring bias is also not diminished by greater expertise or experience (Kaustia et al., 2008; Ogdie et al., 2012). Cognitive biases are prevalent in many everyday decisions and affect all types of people, including physicians. They may be considered as an aspect of human reasoning.(Croskerry, 2014) Consequently, learning to recognize them and being cognizant of their influence in decision making is important. By being more aware of the influence of anchoring bias, for example, may enable individuals to implement mitigating strategies to prevent errors, particularly in pressured settings like ED sign outs.

1.3 Efforts to Develop De-biasing Solutions

The increased awareness of the impact cognitive biases can have on medical decision making and the potential risk to patient safety, has led to efforts towards de-biasing strategies. (Croskerry et al., 2013; Reilly et al 2013; Morewedge et al., 2015) While many de-biasing solutions and tools have been developed, studies suggest that for most, their success has been mixed. Many solutions appeared very effective at first implementation, but over time their initial efficacy was not maintained (Jenkins & Youngstrom, 2016; Kaustia et al., 2008; Wershofen et al., 2016) De-biasing solutions that have been developed have been varied in nature, from mnemonics and checklists (Chew et al., 2016; Ely et al., 2011), to cognitive forcing strategies. Or they involved reflection and mindfulness (Clegg et al., 2015; Mumma & Wilson, 1995; Ogdie et al., 2012). Some have involved an educational approach with didactic seminars to raise awareness of biases (Jenkins & Youngstrom, 2016), while others have developed very sophisticated serious video game-based solutions. Studies of these experiential gaming technology solutions to mitigate cognitive bias, have demonstrated better results in terms of de-

biasing efficacy, both at the initial trial and subsequent post deployment re-test after the elapse of several months. (Clegg et al., 2015; Mohan et al., 2017)

However, the issue remains that the information needs, reasoning processes and physicians' preferences for sign out content and structure are not well understood. (Arora et al., 2005) Novel ways of studying handoffs have been proposed (Gogan et al., 2013; Mamykina et al., 2016) but understanding the cognitive processes involved diagnostic decision making over handoffs remain a challenging area of study. There have been studies looking at the time allocation given to patients in critical care handoffs. Findings indicated that patients later in the transfer list were associated with less time and that this compression of time allocation increased further relative to how much further down the list the patient was. (Abraham, Kannampallil, Patel, et al., 2016; Jones et al., 2013) This phenomenon could have implications in ED settings where the frequency of transitions is higher and typically occur over a much shorter time period of time. Moreover, studies indicate that unstructured handovers contribute to communication breakdowns. These breakdowns may be in the form of omissions of key information such as active medical problems, medications, test results or consults. (Abraham et al., 2011; V. Arora et al., 2005; Cheung et al., 2010)

In terms of debiasing solutions, few of the primary research studies investigating cognitive biases have approached anchoring bias on an individual basis. Instead many studies refer to the presence of anchoring in their outcomes, collectively along with other common biases, such as availability bias and framing effect. (Ludolph & Schulz, 2018; Richie & Josephson, 2018) Few studies have been designed with targeted empirical methods to detect the presence or the mechanisms of anchoring bias. Indeed, in practice providers may not even have awareness of the presence of cognitive biases in the

decisions they have made or even agree on the presence or absence of individual biases. One study found physicians' judgements were heavily influenced by hindsight bias. During case review, the cases where the outcomes were associated with an implied diagnostic error, were likely to have twice as many biases identified by providers. (Zwaan et al., 2017) This has implications not just for the detection and reduction of medical errors but also on the development of de-biasing solutions.

In order to prevent for the influence of anchoring bias in sign outs, debiasing efforts need to address the factors that lead to anchoring. To identify these factors, it is necessary to understand which aspects of the information transferred during sign out affect the decision making of the recipient of the sign out. (Horsky et al., 2015) Therefore, to fully understand the mechanism and factors that promote anchoring bias, it is necessary to focus on and identify the features that influence decision making specifically related to anchoring bias. With this in mind, this dissertation proposal seeks to conduct a focused investigation of anchoring bias. This doctoral research project aims to concentrate on the features and factors that play a role in anchoring during diagnosis and ED sign out. The intent is to identify the risk factors that influence the tendency to anchor during diagnosis in the ED. Ultimately, the knowledge gained will be used to determine the measures needed to mitigate for anchoring bias and inform the development of effective debiasing solutions.

Chapter 2: Literature Review

Decision making is a constant part of everyday life for humans and the outcomes generated can have a profound influence not just for the individual, but for any number of others affected by the decision. Suboptimal decision making in the context of health care delivery has been linked to errors in patient care and so understanding decision making processes may contribute to preventing erroneous decision making and medical errors (Balogh et al, 2015 (IoM- Improving Diagnosis in Healthcare)). However, many factors influence decision making and as such the study of decision making processes has been approached from a multitude of perspectives over time.

2.1 Background

Early research in the field, driven from the perspectives of mathematicians and economists, described prescriptive rational decision making models based on statistical models of rational choice and decision optimization, such Bayesian inference models (Von Neumann, 1944). However, as psychologists began researching and empirically testing human decision making, descriptive models of decision making were developed. These models suggested that human decision making utilized heuristics, rather than being based on pure logic and prescriptive processes. When these heuristics result in a deviation from the correct response, they could be considered as biases. This new perspective on human reasoning spawned a paradigm shift, led by pioneering scholars such as Herbert Simon. Simon's concept of bounded rationality posited that when faced with an overwhelmingly complex world, people form a simplified mental model and

behave within its constraints using heuristics as mental shortcuts in decision making. (Simon, 1955) He suggested that humans challenged with finite cognitive resources are unable to exhaustively consider all available options when selecting the optimal decision. Instead, they select the one that meets their level of acceptability. In other words, they use the strategy of satisficing. (Simon, 1955; Polic, 2009)

Building upon this, Gigerenzer and Selten (2001) proposed that bounded rationality involved heuristics as an 'adaptive toolbox' of fast and frugal rules that operate under the limitations of restricted search, knowledge, and time. They suggested these fast and frugal rules operate well that when there is a "match between the structure of the heuristic and the structure of the environment" (Gigerenzer and Selten, 2001, p.9). The authors suggest these 'fast and frugal' rules could often approximate the accuracy of complex statistical models, with considerably less information and computational effort. Their bounded rationality model describes not only the mechanism of reaching the decision by way of heuristics, but also the outcomes and the environments in which these heuristics will be successful or not. (Gigerenzer & Goldstein, 1996) To that end, the 'adaptive toolbox' provides rules such as the recognition heuristic, which leverages the core cognitive capabilities of recognition memory. (Marewski & Gigerenzer, 2012)

Building on from Simon's concepts of bounded rationality, Kahneman and Tversky (1979) proposed their Prospect Theory about facilitating evaluation and choice. This describes how decision making begins by structuring the decision problem in relation to the possible outcomes in terms of gains and losses and the size of probability. (Kahneman and Tversky, 1979) They suggested that the decision-maker's conceptualization of actions, outcomes and dependencies related with a particular choice equated to the decision frame. The decision frame provides the context for any given choice, or in

terms of Simon's works, the context for the different possible models of the world. (Polic, 2009; Simon, 1955; Tversky & Kahneman, 1974) Since the decision context is subject to both the nature of the problem and the qualities and preferences of the decision maker, by altering the actions, dependencies, and outcomes it is possible to alter the decision maker's preferences. Kahneman and Tversky described this as the framing effect within the healthcare setting. They demonstrated that by using different coding words for the problem outcome, namely by expressing in terms of gains or losses, despite the outcomes having the same connected probabilities, the decision maker's assessment of the outcome could be influenced. (Tversky & Kahneman, 1985) Their extensive research into these cognitive systems demonstrated that people resort to heuristics as mental shortcuts for judgement and decision making, particularly in times of uncertainty. They continued to conduct significant and widely acclaimed research into the role of biases in decision making, elaborating on the Dual Process Theory of human reasoning initially proposed by Peter Wason and Johnathan Evans in 1974 ((Evans, 1984; Rottenstreich & Tversky, 1997; Tversky & Kahneman, 1974, 1986; Wason & Evans, 1974)

2.2 Dual Process Theory

The dual process theory of decision making was developed from the theoretical framework proposed by Jonathan Evans, in which he suggested human reasoning involved two distinct types of processes: heuristic processing and analytical processing. (Evans, 1984) Evans suggests that two different cognitive systems are involved with the perception and processing of information to achieve diagnosis of a problem scenario. ((Evans, 2003; Petty & Cacioppo, 1986)

Evans described these two types of processing, heuristics and analytical, as System 1 and System 2 processes, respectively. (Evans, 1984) He suggested that System 1 processes

are ‘old’ in evolutionary terms and correspond to the selection and ‘judgement of relevance about the features of the problem’. (Evans, 1984, p. 451) If the information taken in is deemed ‘irrelevant’ it is not processed further, while the ‘relevant’ information undergoes processing. (Evans, 2003) The System 1 processes are unconscious, fast, and represent intuitive, often instantaneous decision making. (Evans, 2003; Wason & Evans, 1974) System 2 processes allow abstract complex analytical thinking for decision making. System 2 consists of conscious processes that are logical, focused and involve deliberation. This processing engages the central memory system and so are slower and cognitively effortful. While these two systems are considered to potentially occur in parallel, System 2 can override or inhibit the default outcomes that result from System 1. (Kahneman, 2003)

2.3 Other Decision Making Models

Additionally, others focused on aspects like team dynamics and expertise in decision making in complex dynamic systems. (Cook et al., 2004; Salas et al., 2008; Salas & Klein, 2001) Subsequently, alternative models for decision making emerged. One such model that gained attention is The Recognition Primed Decision (RPD) model proposed by Gary Klein and colleagues (1997). They wished to understand how experts made quick effective decisions in complex environments, such as firefighting and medical trauma care. Klein’s RPD model suggested that people contemplate a situation based on their experience in previous similar scenarios while also considering the constraints present in the current situation. Their reasoning involves rehearsing mental models of plausible solution options to select the most adequate response. These options are assessed based on the impact of their consequences by projection into the future for their

potential undesirable effects. The option with the least potential unwanted effects is the one selected and acted upon. (W. M. Klein, 1997)

Rather than viewing any one decision making model as better or more appropriate than another, it is important to recognize this field of research is still evolving. The models of Dual Process Theory, Situation Awareness, and Recognition Primed Decision could instead be considered as complementary ways of viewing aspects of human cognition and decision making. There are even areas of similarity between the models regarding certain concepts but at this time, the exact ways in which these models are related are not well understood. However, aspects of each of them can be considered when investigating decision making. Further research in this field will contribute to our greater understanding of their part in the overarching model of human decision making.

Meanwhile, there was additional research on decision making in other disparate fields such as the military and aviation. Researchers sought to understand decision making with a different emphasis; decision making that occurs within natural environments, which are inherently more dynamic and complex than laboratory settings. (Collyer & Malecki, 1998; M. R. Endsley, 1995; Zsambok & Klein, 2014) This work in naturalistic decision making models viewed decision making within the context of systems and tasks. Consequently, it examined decision events from the perspective of fulfilling the achievement of wider goals within an activity. Much of the more contemporary research in this field of naturalistic decision making has focused on aspects of cognition within operational situations and complexity. Decision making was studied within the context of goals, tasks, fault management and planning. (Gorman et al., 2006; D. A. Norman, 1995; Schraagen & van de Ven, 2008) Advances were made by researchers in the study

of situation awareness (SA), which is the perception and comprehension of information in a given situation to make decisions while understanding the consequences of the decision on the near future state. (M. Endsley, 1995)

2.4 Clinical Decision Making

Clinical reasoning involves judgement associated with providing medical care. Understanding the factors that impact the clinical reasoning process are pivotal to improving diagnosis. (Croskerry, 2003) Diagnostic reasoning is the process by which physicians develop an actionable diagnosis by collecting, interpreting, and integrating the information and medical data about a patient. There has been much debate about the different paths clinicians may use to arrive at their diagnosis. Elstein and colleagues (Elstein, 1978) suggest that clinicians use a hypothetico-deductive approach. This suggests that based on their perceptions of the disease condition, clinicians may formulate a mental shortlist of clinical hypotheses. They then seek further clinical evidence for these different hypotheses, to reach a final diagnosis and treatment plan. (Elstein, 1978, 1999) When making their diagnostic assessment they may compare patterns of the disease they observe against those of disease patterns recalled from their previous experience, to find a match. (G. Norman et al., 2007) This approach is reminiscent of the RPD model. For RPD, Klein suggests there are cues in any situation that enable people to recognize patterns based on intuition formed from their past experiences.(G. A. Klein, 2004) Similarly the SA model of M. Endsley (1995) also describes the perception of situational cues as level 1 SA, when assessing a particular scenario. Clinicians often use a combination of analytical reasoning and unconscious heuristics when diagnosing patients. (Chapman & Sonnenberg, 2003; Croskerry, 2002) Heuristics are intuitive cognitive processes that involve pattern recognition. These ‘rules

of thumb' are influenced by implicit knowledge developed from the experience of treating numerous patients during their practice. These heuristic processes are subconscious, automatic and frugal in terms of cognitive effort. They have a significant influence in decision making and, consequently, may influence clinical reasoning. (Croskerry, 2013; Croskerry & Norman, 2008; Marcum, 2012; Reilly et al., 2013) Clinicians are known to utilize heuristics when developing a differential diagnosis (Croskerry, 2002, 2013) Diagnostic decision making includes considering patient information such as medical history, diagnostic test results and medication lists as part of the diagnostic process. (Croskerry, 2009; Ely et al., 2011) Applying the concepts of heuristics and analytical decision making, Croskerry and colleagues proposed their Universal Model for Diagnostic Reasoning. (Croskerry, 2009)

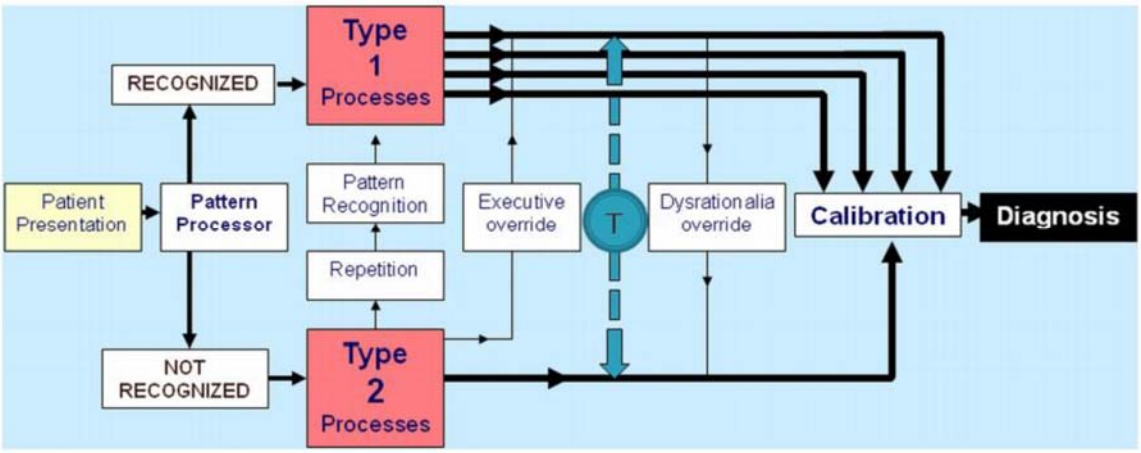


Fig 1: Model for diagnostic reasoning based on dual-process theory adapted from Croskerry P. Cognitive biasing 1: Origins and theory of debiasing. BMJ 2013 (Croskerry et al., 2013)

This model for diagnostic reasoning proposed by Croskerry (Croskerry et al., 2013) is based on Dual Process Theory. Croskerry suggests that humans are primed for pattern recognition so the System 1 processes (labelled Type 1 processes in Figure 1), are the first to activate when the clinician is assessing a patient's presentation of illness. When the cues are not recognized to fit an implicit pattern, the mind switches to the analytical approach of System 2 processes (labelled Type 2 processes in Figure 1). He also describes the ability of the mind of the clinician to be able to quickly move back and forth between the System 1 and System 2 processes to achieve a diagnosis. (Croskerry, 2009)

When the cues are well calibrated to the context of the situation, the System 1 processes lead to correct and appropriate diagnosis. However, when there is poor calibration, cognitive biases occur as 'predictable deviations from rationality' (Croskerry, 2013, p. 25) Researchers suggest that clinical decision making is vulnerable to cognitive biases, as with decision making in other domains. (Croskerry, 2003, 2009b, 2014; Croskerry et al., 2013) In some circumstances, cognitive biases may cause faulty understanding of the situation and lead to erroneous decisions making. (Cheung et al., 2010; Croskerry, 2013, 2014)

Cognitive biases have multiple causes, but it is these heuristics, that are known to play a major contributing role. (Committee on Diagnostic Error in Health Care et al., 2015; Marcum, 2012; Zwaan et al., 2017) Cognitive biases exist in a multitude of forms and affect clinical decision making. (Croskerry, 2013, 2014) Rather than viewing cognitive biases in terms of fault, they should be considered as inevitable and part of "the normal operating characteristic of the brain". (Croskerry, 2014, p. 25) Alternatively, it is important to be cognizant of their influence in decision making. This may enable

individuals to compensate for them, particularly in error-prone settings like the emergency department. (Croskerry, 2003, 2013; Mohan et al., 2017)

Cognitive biases have been studied extensively in domains outside of healthcare. Numerous cognitive biases have been identified in the literature (Blumenthal-Barby & Krieger, 2015; Clegg et al., 2015; Ludolph & Schulz, 2018; Mussweiler & Englich, 2005; Walmsley & Gilbey, 2017)

Some of the more commonly known ones are:

- Availability heuristic – situations where people judge the likelihood of an event by the ease with which instances of it can be brought to mind
- Confirmation Bias – the tendency to perceive more supporting evidence of one's prior belief than actually exists
- Framing effect – described earlier – the evaluation of probabilities and outcomes produce shifts of preference when the same problem is presented in different ways e.g., by using different language
- Anchoring Bias – the tendency to lock on to an initial value or information received early on in the decision making process and judgement.

  ((Ludolph & Schulz, 2018; Richie & Josephson, 2018)

There is increased awareness of cognitive biases, which is leading to other biases being identified and defined in the literature. The literature also includes informative and comprehensive systematic reviews of cognitive biases, as well as, different debiasing techniques and solutions. ((Blumenthal-Barby & Krieger, 2015; Furnham & Boo, 2011; Ludolph & Schulz, 2018; Saposnik et al., 2016)

2.5 Anchoring Bias

The focus of this research has been anchoring bias, which is known to influence clinical decision making. Anchoring is described as 'the disproportionate influence on decision-makers to make judgements that are biased towards an initially presented value' (Furnham & Boo, 2011, p. 35) This initial value acts as an 'anchor' and the final assessment is drawn towards this initial value as a starting point. Anchors can be an externally provided or a self-generated value. (Epley & Gilovich, 2005; Greenstein & Velazquez, 2017; Richie & Josephson, 2018)

The influence of anchoring has been studied by researchers from such disparate fields like sales negotiations (Galinsky, 2001), courtrooms (Davis & al, 1984) marketing (Wansink, 1998) general knowledge quizzes (Strack & Mussweiler, 1997) and lotteries and gambling. (Chapman & Johnson, 1994) Within healthcare anchoring is a very prevalent cognitive bias. Anchoring bias has been identified as a contributing factor in many cases of medical error that was attributed to either an incorrect diagnosis or severely delayed diagnosis. (Keeney & Halalau, 2017; van Geene et al., 2016) The anchoring effect is considered a robust bias (Augestad et al., 2016; Garcia-Molina & Chicaiza Becerra, 2015) that persists despite incentives and debiasing efforts. (Epley & Gilovich, 2005; George et al., 2000) and affects novices and experts. (Kaustia et al., 2008; Ogdie et al., 2012) Flaws in diagnosis have been attributed to the misdirection of reasoning by salient features, that upon review turn out to not be relevant to the patient problem. (Croskerry, 2013; Mamede et al., 2012)

2.6 Medical Errors and Transitions of Care

Since the publication of the Institute of Medicine report drew the spotlight on the prevalence of medical error there has been a growing focus on identifying the types and causes of such error. According to current estimates from the Centers for Disease Control (CDC), medical error is a top leading cause of death in the US. (Bates & Singh, 2018; Makary & Daniel, 2016) Studies looking at autopsy data, indicate rates of 10-15% for errors in diagnosis and some reports indicate these figures may be even higher. (Gandhi et al., 2006; Schiff et al., 2005, 2009) One study analyzing one hundred cases of diagnostic error in internal medicine, found that 65% of cases involved system related factors but 75% involved cognitive factors. (Graber et al., 2005) Analysis of data from an Emergency Department (ED) incident reporting system, which contained voluntary reports from ED physicians of patient cases involving error, indicated similar findings.(Okafor et al., 2016) Results indicated that only 4% of diagnostic errors were related to system factors and the majority of errors involved cognition, with 18% related to cognitive factors alone and 40% due to both system and cognitive factors. (Graber et al., 2005; Okafor et al., 2016) Improvements in patient safety measures over recent years have been successful in the identifying and addressing systems errors. (Singh & Sittig, 2015) However, addressing the causes of diagnostic errors related to cognitive factors is challenging, as these errors are not easily identified and when detected, typically require self-reporting. (Singh, 2014; Singh et al., 2017)

Emergency care settings, in particular, have been linked to errors in diagnosis. Studies that examined ED medical malpractice claims, suggest that 65% of them involved missed diagnoses ((Kachalia et al., 2007; Okafor et al., 2016). In addition, studies from The Joint

Commission found that 70% of medical errors involved a breakdown in communication. Of these errors, transitions of care were involved in 50% of them. (*Sentinel Event | Joint Commission*, May 2019)

2.7 Emergency Medicine

In the ED clinicians work in challenging and pressured conditions. They typically manage multiple patients presenting in various states of acuity, often complicated by high levels of uncertainty due to the lack of available patient information. Because many of the patients in the ED present in a critical condition that requires urgent attention and care, ED physicians must promptly assess the patient and develop a working diagnosis and treatment plan. In many urban trauma centers, on a busy night, the ED physicians may deal with extended periods of high patient volumes and a succession of serious and complex patient cases with and therefore make rapid diagnostic decisions under time pressure and stress. (Franklin et al., 2011; Laxmisan et al., 2007; Marcum, 2012) The work environment of the ED requires providers to constantly switch between their long and short term memory, while managing high cognitive load and fatigue, making optimal decision making challenging. Not surprisingly, ED settings have been associated with medical errors during both routine treatment and transitions of care. (Cheung et al., 2010; Franklin et al., 2011; Okafor et al., 2016)

Most hospital emergency departments operate on a twenty-four-hour basis, throughout all the days of the year. Consequently, the ED staffing teams will undergo two or three shift changes daily and therefore conduct handoffs at these points of care transition. These handoffs allow providing patients continuity of care across shift changes. ED physician transitions of care are known as “sign-outs”. Sign-outs represent the transfer of key patient information, responsibility, and accountability. The high patient volumes,

complexity of patient cases, and limited medical history in the ED may contribute to the potential for medical errors. Hence, these transitions of care pose a risk for patient safety due to information gaps. (Dhingra et al., 2010; Kachalia et al., 2007) The limited time available for sign outs may provide the opportunity for the introduction of cognitive errors that lead to missed, delayed or incorrect diagnosis. (V. Arora et al., 2005; Dhingra et al., 2010; Okafor et al., 2016)

2.8 Transitions of Care and Handoffs

The Joint Commission recognizes transitions of care as a major contributor to sentinel events. (*Sentinel Event | Joint Commission*, May 2019.) There has been interest in understanding the factors associated with patient safety issues and transitions of care. Various researchers have conducted studies investigating the processes and communication that occur in transitions of care in critical care settings, known as handoffs. (Abraham et al., 2011; Abraham, Kannampallil, Brenner, et al., 2016; Jones et al., 2013; Kitch et al., 2008) ED sign outs are similar to critical care handoffs in that they represent a transfer of care, responsibility, and accountability for the patient from the clinician leaving at the end of a shift to their counterpart clinician coming on at the start of a new shift.

Studies in critical care settings indicate that unequal allocation of time can occur for individual patient cases in a handoff, with patient cases lower down the transfer list receiving approximately 50% less time. This time compression further increases as more patients are added to the list. (Abraham, Kannampallil, Patel, et al., 2016; Jones et al., 2013; T. Kannampallil et al., 2011) This has implications in the ED setting, where transitions are more frequent. There are some notable differences between sign-outs in the ED compared to critical care or general medicine settings. The time over which sign

outs are conducted is significantly shorter and involve transfer of patient information in an abbreviated and condensed form. Most ED sign-outs are around 20 minutes in duration but can be longer due to sudden increases in patient volumes. (Abraham, Kannampallil, Patel, et al., 2016; Frye et al., 2018) Sign-outs are conducted primarily verbally and standardized handoff protocols such as SBAR have not been universally used. (Cheung et al., 2010; Dhingra et al., 2010) Studies indicate that unstructured handoffs can lead to communication breakdowns, in the form of omissions of key information, such as active medical problems, medications, test results, or consults. (Abraham et al., 2012; V. Arora et al., 2005; Cheung et al., 2010; Kachalia et al., 2007) Verbal transfers of patient information during handoffs have been associated significant information loss, making ED sign-outs a significant risk for patient safety. (V. Arora et al., 2005; Frye et al., 2018) Furthermore, one study into verbal handoffs, found that over 30% of residents reported the quality of the handoffs to be suboptimal. Only 26% reported that they were conducted in a quiet place and 37% reported interruptions during receiving handoffs. (Kitch et al., 2008) Studies investigating interruptions indicate that the frequency of interruptions during sign out is very high, with attendings being interrupted every 9 minutes and residents every 14 minutes. These interruptions may impact task completion and communication flow, as well as lead to the loss or omission of information during sign out. (Brixey et al., 2008; Laxmisan et al., 2007) Much of the literature on the structure and content of handoff processes or the communication between clinicians during handoffs has been in critical care, surgical, or general medicine settings. (Abraham et al., 2011, 2014; Abraham, Kannampallil, Brenner, et al., 2016; Anderson et al., 2015; T. Kannampallil et al., 2011; T. G. Kannampallil et al., 2014) The information needs and cognitive process during ED sign-outs are poorly understood and

researched. Moreover, literature on the influence of cognitive biases associated with ED sign-outs is limited and thus needs further study. (V. Arora et al., 2005; V. M. Arora et al., 2013; Cheung et al., 2010)

2.9 Anchoring Bias and Sign Outs

According to the APA Dictionary of Psychology, anchoring bias is the tendency, in forming perceptions or making quantitative judgments under conditions of uncertainty, to give excessive weight to the starting value (or anchor), based on the first received information or one's initial judgment, and not to modify this anchor sufficiently in light of later information. In clinical decision making and, in particular, diagnoses that occur over transitions of care, this could be heavily influenced by the information transferred in the sign-out. (Campbell et al., 2007; Croskerry, 2003; Mull et al., 2015) Reviews of patient cases in ED settings involving sign-outs where diagnostic error or delays have occurred, indicate that several factors may have been involved. (Okafor et al, 2017) It is possible that these factors may potentially introduce an anchoring bias. In such cases, could it result in premature closure of the diagnostic process? Literature exploring anchoring in healthcare have mentioned a number of risk factors that may have contributed to anchoring during the diagnostic process, such as:

- Anchoring on a patient's chronic conditions (Keeney & Halalau, 2017; Montemayor et al., 2018; Parekh et al., 2013)
- A provisional diagnosis provided by the departing physician (Frye et al., 2018; Mull et al., 2015)
- Not questioning the person giving their opinions or assessment of the situation or test results because they were considered an expert due to either seniority or

specialization e.g. an attending physician or radiologist (Campbell et al., 2007; Cree et al., 2007)

• Information of previous possible diagnoses mentioned by the patient (Armstrong & Thurber, 2014)

As many ED sign-outs occur without the use of protocols, critical patients may not be prioritized in terms of chronological order of the presentation. Furthermore, interruptions may result in their information being dispersed within the overall handoff.

The information and time constraints, coupled with the potential for communication breakdowns make sign-outs prone to the introduction of cognitive biases. Due to their unstructured nature and short duration, the amount and quality of information transferred about each patient may be compromised in ED sign-outs. These factors may lead to the introduction of biases from the mental model and picture painted by the exiting physician. (V. Arora et al., 2005; Frye et al., 2018; Long, 2015)

2.10 Debiasing Efforts

With the growing understanding of cognitive biases in terms of their pervasiveness, and the negative consequences in terms of error generation, there is greater research interest in mitigating their impact. Consequently, there has been growing recognition of the need for debiasing strategies and tools. While there have been efforts to build or develop debiasing solutions, studies reviewing these suggest their effectiveness or success has been mixed. Many of these tools or interventions appear very effective at first implementation, but then fail to maintain their efficacy over time. (Jenkins & Youngstrom, 2016; Kaustia et al., 2008; Wershofen et al., 2016)

Debiasing solutions developed have been varied from cognitive forcing strategies, like 'think of the opposite' or modifying behavior through reflection and mindfulness (Clegg et al., 2015; Mumma & Wilson, 1995; Ogdie et al., 2012), to mnemonics and checklists (Chew et al., 2016; Ely et al., 2011) Others have involved raising awareness through education and didactic seminars.(Jenkins & Youngstrom, 2016) Indeed, some have involved the development of sophisticated video game-based solutions. Studies of these experiential gaming technology solutions have demonstrated better results in terms of debiasing efficacy, both at the initial trial and at subsequent re-testing conducted several months post initial implementation. (Clegg et al., 2015; Mohan et al., 2017)

Some of the limitations observed in the debiasing literature are that many studies investigating cognitive biases have study designs examining anchoring bias in conjunction with other common biases, such as framing effect and availability bias. The studies do not separate out the constructs, observations, or measures for anchoring bias independently and objectively, but rather investigate it as part of a more generalized study of biases. (Mamede et al., 2012; Richie & Josephson, 2018) Some of the studies are designed to simply detect the presence of cognitive bias. They do not go further to specifically tease apart the factors that influence individual biases, like anchoring. Systematic reviews of literature on cognitive biases indicate anchoring bias is one of the most prevalent biases in medical decision making. However, the same reviews also show that it is resistant to debiasing efforts such as raising awareness or incentives. (Augestad et al., 2016; Chapman et al., 2002; George et al., 2000; Richie & Josephson, 2018; Rottenstreich & Tversky, 1997)

2.11 Efforts to Standardize Sign-outs

The variability in sign out practices across providers and institutions have led to calls for the adoption of more structured protocols for ED sign-outs. Efforts to standardize sign-outs have resulted in various products being implemented or used. Three main handoff protocol templates that have become popular are SBAR, I-PASS and The American College of Emergency Physicians (ACEP) Safer Sign Out.

SBAR Protocols. Situation-Background-Assessment-Recommendation (SBAR) is a technique originally developed by the military for naval communications. It provides a framework for communication that is easy to remember and has been widely adopted within the healthcare domain. It is useful for framing conversations, especially critical ones, requiring a clinician's immediate attention. It has been used in many settings to facilitate communication between members of a team and fostering shared understanding the situation. (Hern et al., 2016; Riesenberg et al., 2009; Tews et al., 2012)

I-PASS Protocols. I-PASS is a mnemonic that stands for illness severity, patient summary, action list, situation awareness and contingency plans, and synthesis to standardize sign out communications. It is often used as a handoff protocol in critical care settings. It is promoted by the Agency for Healthcare Research and Quality (AHRQ) and Accreditation Council for Graduate Medical Education (ACGME), which regulates physician training and now requires all residents to receive handoff training. I-PASS now serves as the cornerstone for the resident handoff training. (Hern et al., 2016; O'Toole et al., 2020; Starmer et al., 2014)

ACEP's Safer Sign Out Protocols. ACEP's Safer Sign Out is a protocol that may be more appropriate for use in ED transitions of care, having been developed to

consider the needs and constraints of ED settings and ED sign-outs. The Safer Sign Out protocol is based on information, documentation and review combined with effective communication with and across teams, including receiving appropriate feedback (Hazan & Haber, 2018)

SIGNOUTS Mnemonic. Many clinicians themselves also use mnemonics as well as protocols. And one such mnemonic that is used for sign out is conveniently called SIGNOUTS (sick or not, identification data, general hospital course, new events of the day, overall health status, upcoming possibilities with plan, tasks, and questions). (Horwitz et al., 2007) Looking at both I-PASS and the SIGNOUTS mnemonic, there seemed to be common components between them. These are items like identifiers such as patient name, age, room number, etc. There is the activity of assessment which includes examination, the medical history, labs, and therapeutics. There is also a contingency planning component around possible events that may happen or if the patient does not respond as expected. Finally, there is a verification stage which involves a recap of the case, as well as questions and answers between the two physicians.

The introduction of the protocols above has been met with mixed reactions, as in some environments they have made sign outs significantly longer to conduct, adding to the workload of already heavily burdened ED physicians. (Dhingra et al., 2010; Heilman et al., 2016; Tews et al., 2012) In one study that surveyed 175 institutions about the use of standardized protocols, their results indicated there had been a 14% increase in the use of the protocols in the ED over 5 years. However, while the respondents felt the sign outs were more structured, they did not feel any increase in their confidence in their ability to provide sign outs competently. This suggests more training and support is needed to help

physicians conduct efficient structured sign outs. (Hern et al., 2016) Another study found ED physicians perceived handoffs contained ambiguity about the patient's condition and treatment. They also felt faulty communication behaviors and conflicting information were related to poor handoffs and poor handoffs contributed to patient boarding related errors. (Apker et al., 2007; Cheung et al., 2010)

2.12 Summary

In summary, to know how best to structure and improve sign outs, it is important to understand the information needs of the clinicians involved and their cognitive processes in terms of clinical reasoning and treatment planning. It is important then to understand the factors that may influence their reasoning and decision making, as well as whether those factors are conscious and deliberate or implicit and intuitive.

It is for this purpose this research project was conducted with hopes of contributing to the knowledge about ED diagnostic decision making and the influence of cognitive biases like anchoring bias. Ultimately, the knowledge gained will be used to determine the measures needed to mitigate for anchoring bias, develop debiasing solutions and identify areas for future research.

Chapter 3: Methodology

This aim of this dissertation research was to identify and evaluate factors that might influence clinical decision making and potentially lead to anchoring bias. To achieve this the research project was conducted over two phases. The first phase aimed to identify the factors that contribute to the development of anchoring bias when formulating clinical diagnoses. The second phase was to evaluate the effect of the identified risk factors on clinical decision-making amongst emergency medicine physicians in cases of sign outs. This was achieved by way of a study involving fictious patient case vignettes presented in a manner similar to an emergency department sign out.

The first phase represented Specific Aim 1 of this research proposal and was further comprised of two studies; the first study involved a review patient cases within a limited dataset from a medical incident reporting system database (MIRS). The purpose of this was to identify factors that may have been involved in the generation of errors in those cases. The second study involved a series of interviews with emergency medicine physicians to gain insights regarding how sign outs are conducted, physician information needs, and the diagnostic process within the environment of the emergency department and sign outs.

In summary the dissertation research comprised:

Specific Aim I: Identifying Risk Factors for Anchoring Bias in Sign-outs

1) Review of a medical incident reporting system dataset for cases of reported medical error that involved sign out

2) Interviews with emergency medicine physicians

Specific Aim 2: Evaluating Risk Factors for Anchoring Bias in Simulated ED Sign Outs

1) Design and review of Patient Case Vignettes and study instruments
2) Emergency Department Sign Out (EDSO) Study – to evaluate the risk factors with emergency medicine physicians and advanced practice providers as participants.

The overall methodological approach of this dissertation research involved a triangulation approach where two qualitative data collection and analysis studies were conducted to inform the development of a quasi-experimental study aimed at evaluating the effect of the features selected as potential risk factors for anchoring. The substantive nature of the design process involved in developing the study instruments for conducting Specific Aim 2 work necessitated discussion in a separate chapter and so will be addressed in Chapter IV of this dissertation.

This chapter will discuss the methodology of the two Specific Aim 1 studies of this project, which focused on understanding the diagnostic process in the ED and how this is affected when conducted across shift changes. The review the medical incident reporting system data aimed to gain insight on the types of features specified by clinicians as contributing factors in reported cases of medical error in the emergency department. The aim of the interview study was to gain a more detailed understanding of the communication and information needs and the diagnostic process in the ED from the perspective of ED physicians.

3.1 MIRS Analysis Study

The review of the medical incident reporting system data was intended to identify common risk factor themes amongst known cases of diagnostic error in the ED that involved sign outs. A review of a subset of de-identified patient cases within a dataset

extract from the medical incident reporting database for two academic teaching hospitals affiliated with the University of Texas Health Science (UTHealth) McGovern Medical School

The aim of this study was to conduct a retrospective review of patient cases that involved instance of medical error. The study reviewed closed patient cases that were de-identified for any provider or patient personal identifying information.

3.1.1 The MIRS System

The McGovern Medical School Department of Emergency Medicine's Medical Incidence Reporting System (MIRS) database was created and implemented as a department specific medical error reporting system for the two academic teaching hospitals of McGovern Medical School. The MIRS system contained case reports from patients treated at either of hospitals; one being a large urban tertiary referral, level 1 trauma center that averages over 75,000 ED patient visits annually and a second site that is an urban county hospital with an annual volume of 88,000 ED patient visits. The system was developed in house, led by a physician champion, and supported by the Emergency Department Quality Assurance Committee of McGovern Medical School.(Okafor et al., 2016)

The system was intended to provide a database to collect patient cases in which there may have been a medical error or near miss. It was operationalized in December 2011 and is maintained by the Emergency Medicine Quality Assurance program. It contains more than 3,000 cases primarily from voluntary reports submitted by various sources, including attending physicians, EM residents, EM faculty and other associated advanced practice providers. Each case undergoes an iterative review process to determine whether an error occurred, contributing factors, error type, and clinical impact. Both sites where

this MIRS system is deployed have well established and integrated and electronic health records systems (EHRs) that can be accessed at the time of review.

3.1.2 Dataset information

A dataset of de-identified case reviewed reports was provided by Kimberly A. Chambers, MD, who was an Assistant Professor at the Department of Emergency Medicine and the study contact at the time this study was conducted. The dataset contained a subset of de-identified patient case reports within the MIRS system in which there were either:

- diagnostic errors,
- delays in diagnosis
- suboptimal management
- cases that extended over at least one shift change, therefore involved a sign out.

The data in the extract contained no EHR medical record identifiers or patient personal identifiers other than age and gender. The records did contain a MIRS unique identifier and date of incident, but all other distinguishing information was stripped including any identifying information about the clinical staff who were involved in the cases. The records were loaded into Microsoft Excel and processed to combine the three extracts and remove duplicate records.

3.1.3 Protection of Human Subjects

As this dataset contained patient information an Institutional Review Board (IRB) application was submitted by the researcher via the UTHealth Integrated Research Information Software (iRIS) system accompanied by the appropriate the supporting IRB formatted study protocol. The study protocol listed Amy Franklin, PhD as the principal investigator, Roni Matin, MSc, Kimberly A. Chambers, MD and Yashwant

Chathampally, MD, MSc, CMQ, who is Vice Chair of the EM Quality Assurance program, as co-investigators. The study IRB details are as below:

Study Title: Understanding Decision Making Across Shift Changes.

IRB Protocol #: Number HSC-SBMI-19-0817

The study was approved as Exempt status by the IRB board on October 11th 2019.

As part of the protection of human subjects, the protocol specifies that only a subset of de-identified patient cases which were three years or older, panel reviewed and closed at the time of extraction would be included in the dataset for review. This ensured the data did not contain any patient personal health information (PHI) and the cases were beyond the statute of limitations thereby protecting the patients and the institutions involved from the risk of any potential accidental data breach. This study contained no active participants.

3.1.4 Method

Data Processing. The dataset was provided as three separate extractions in MS Excel Spreadsheet files file format over the period of October 26th2019 -January 10th 2020. The files were accessed from UTHealth Kiteworks secure store and file transfer drives via encrypted dual security sign in procedures.

Table 1

MIRS Data Extract File Details

#	FILE NAME	NUMBER OF ROWS	UNIQUE MIRS ID#	UPLOADED BY	COMMENTS
1	MIRS Matin 1to16.xlsx	**263**	**16**	K. Chambers	**Contained full records**
2	Matin 2nd Set of cases 2012.xlsx	**95**	**52**	K. Chambers	**Contained partial data**
3	MIRS Linking 2nd Set.xlsx	**3290**	**487**	K. Chambers	**Contained multiple rows of data per record**
4	MIRS all merged data012521.xlsx	**734**	**68**	R Matin	**Three files merged, reduced for duplicates**
4	MIS Analysis cases for follow up.xlsx		**7**	R. Matin	**Post initial analysis follow up**
6	MIRS all merged data 012521xlsx revised	**734**	**65**	R. Matin	**Ready for analysis**

The data was minimally manipulated following extraction. However, for the purpose of consistency the three MS Excel extract files were processed to combine the data in the three files into one master file using the column headers and MIRS unique ID numbers to match and align the records. The processing took several steps:

- The first file MIRS Matin1to16.xlsx contained the following columns *Location*, *MIRS ID*, *Summary* (narrative details about the case), *Initial Impression* (narrative field with type of case data**)** *Clinician Narrative* (with notes from the various clinicians on the case), *Age*, Unlabelled column (with error type) and *Contributing Factor* (specified by clinician involved in the patient case as to what they considered contributed to the error that occurred)
- The first extract file was extracted in a format where every patient case had a row of data for each instance of contributing factor

- The second data extract file Matin2nd set of cases.xlsx did not appear to have the same format as the first and many patient records only had a single row of data with one contributing factor.
- Dr Chambers was contacted and informed and subsequently a third linking file MIRS Linking 2nd set .xlsx was extracted that contained multiple rows of data for the *Contributing Factor* field.
- The third extract file had 3290 rows of data for what appeared to be 487 unique records.
- The second file was appended with the addition rows of data for the patient records copied from the third file for those cases where the *MIRS ID* corresponded to the *MIRS ID* numbers in the second file.
- 11 patient records in MIRS Linking 2nd set.xlsx file did not have corresponding *MIRS ID* records in the second file Matin 2nd set of cases.xlsx
- This combined file was then appended to the first file of 16 patient case reports.
- The final merged file contained 68 patient cases with 734 records.
- Seven cases were selected for follow up with Dr. Chambers due to anomalies in the data
- After initial review for consistency in the merged file one patient record was found to contain incomplete data for the *Contributing Factor* field and was excluded from analysis
- Following a further review another two patient records were excluded from analysis as their rows were missing case data and so had insufficient data for review
- A final merged file of 65 patient records and 732 rows of data were reviewed.

3.1.5 Data Synthesis and Analysis

The fields in data extract files included columns with headings of *Location*, *MIRS ID*, *Summary* (narrative details about the case), Initial Impression (narrative field with type of case data) *Clinician Narrative* (with notes from the various clinicians on the case), *Age*, Unnamed column (with what appeared to be error type). The data in the two main narrative fields contained information about the patients' course in the ED including any cross-shift sign outs.

3.1.6 Definitions

The cases were reviewed and coded for the type of chief compliant the patient presented with in the emergency department and the contributing factors that were listed by the clinicians who conducted the review of the case. The following are the categories that were assigned to the data:

Cardiovascular: Cases that were in the chest area and involved the heart and/or lungs. Conditions such as chest pain or hypertension were labelled as cardiovascular

Abdominal: Cases that involved the abdomen and had symptoms like abdominal pain, nausea or vomiting or renal issues were labelled as abdominal.

Head: Cases that involve the head and neck were labelled Head. These included cases of headaches, dizziness, altered mental state and vision issues.

Trauma: Cases where there were sign of trauma such as injury from a fall or assault along with injury from an accident or motor vehicle accident were labelled as Trauma

Sepsis: Cases that were recorded as cases of suspected sepsis at the time of presentation in the emergency department or were diagnosed as sepsis were labelled as such sepsis.

Contributing Factor: This field corresponds to the field also labelled Contributing Factor in the dataset and represents an option of a factor contributing to the error that has been selected by the clinician involved on the case during the panel review of the case. A clinician can select multiple contributing factors for each case from a dropdown list.

3.1.7 Case Review Process

The 65 cases were analyzed for the type of cases that presented in the ED in terms of the main body system related to the chief complaints. Cases that involved the chest area, such as chest pain, heart or lung conditions and hypertension were labelled Cardiovascular. Cases that involved the abdominal area such as belly pain, vomiting, renal issues, and cases involving complications from diabetes were labelled Abdominal. Those cases involving the head and neck area, including headaches, dizziness and altered mental status were labelled Head. Cases which were obvious cases of trauma such as injury from a fall or accident, motor vehicle accidents, injury from physical assault were labelled Trauma. Finally, a small number of cases presented as clear cases of sepsis, and these were labelled as such.

The field *Contributing Factor* was of most interest in the data as this referred to the cause attributed by the clinician(s) involved in the patient case. Each clinician involved in a particular patient's care completed this field with their assessment of what factors may have led to an error or suboptimal care. Each clinician could specify more than one type of Contributing Factor which they selected from a drop down list in the MIRS system. Table 2 below shows the 21 options of Contributing Factor. In an effort to rationalize the different codes for this field some options were given the same code. Because the records were de-identified for not just the identity of the clinician but also the number of clinicians involved in each case it was not possible to determine which clinician provided

which entry for the Contributing Factor field. For Faulty Information Processing it was not possible to determine whether this was a self-reported factor or attributed by a team member during the Quality Assurance Committee case review. Hence this option was coded as Faulty Information Verification in the analysis. Similarly, the options Inefficient Processes and Insufficient Resources, which were system issues rather than issues of communication or clinical decision making were coded as Inefficient Resources.

Table 2

Mapping of Contributing Factors Field in Extract Data Files to Codes in Analysis

1	Inadequate handoff/sign out	1	Inadequate handoff/sign out
2	Faulty information processing		
3	Faulty information verification	2	Faulty information verification
4	High workload	3	High workload
5	Faulty information gathering	4	Faulty information gathering
6	Complicated medical history	5	Complicated medical history
7	Atypical presentation	6	Atypical presentation
8	Inefficient processes	7	Inefficient processes
9	Insufficient resources		
10	Supervision Failure	8	Supervision Failure
11	Premature Disposition	9	Premature Disposition
12	Interruptions	10	Interruptions
13	Non-handoff communication error	11	Non-handoff communication error
14	Faulty knowledge	12	Faulty knowledge
15	Limited History	13	Limited History
16	Faulty workload management	14	Faulty workload management
17	Obesity	15	Obesity
18	Patient non-adherence	16	Patient non-adherence
19	Rare condition	17	Rare condition
20	Language barrier	18	Language barrier
21	Other	19	Other

Note: Contributing Factors listed in the data mapping to Contributing Factor codes

The Initial analysis was conducted of the narrative data. Where mentioned in the text, information regarding the patients' conditions, corollary orders, diagnoses, chief complaint and contributing factor(s) were coded and recorded in an MS Excel file.

3.2 Physician Interviews

This component of the project involves an interview-based approach to investigate the interaction and information exchange between physicians during ED sign outs to better understand the factors that affect the transfer of information and characteristics of the communication between clinicians during transitions of care in the ED. An interview format was selected as the method of data collection to ensure the data contained the perspectives and reflections of emergency medicine physicians themselves in relation to their experiences with the transfer of patient information and responsibility as part of the handoff that occurs at sign out. This data is intended to inform understanding of the factors that influence the sign out communication and the development of a shared mental model during the process of handoff between the physicians at change of shift.

3.2.1 Study Participants

The study participants were emergency medicine physicians, who had experience of working in mid-level to tertiary hospital emergency departments and included attending physicians and resident physicians.

Participant recruitment was conducted from McGovern Medical School, via direct communication at the monthly Emergency Department Quality Review departmental meetings following an introduction from Dr. Kimberley Chambers and followed up with an email containing the study adult consent form to those participants who expressed an interest in participating in the study at time of scheduling the interview

Participants who were board certified physicians and who had at least one year of recent experience working in emergency departments were included. Attending physicians, residents (doctors in training) as well mid-level providers were eligible to be participants. However, medical students were not included in this study.

3.2.2 Protection of Human Subjects

An Institutional Review Board (IRB) application was submitted via the UTHealth Integrated Research Information Software (iRIS) system accompanied by the appropriate the supporting IRB formatted documents of a study protocol, adult informed consent forms, and the interview field guide containing the semi-structured questions that would be asked of participants in the interviews. The study protocol listed Amy Franklin, PhD as the principal investigator, Roni Matin, MSc as the co-investigator. The details of the IRB application are as below:

IRB Study Title: Interviews with Physicians to Understand Influence and Shared Mental Models during Handoffs in Emergency Medicine

Study Alias: Emergency Medicine Physician Interviews

IRB Protocol #: HSC-SBMI-20-0654

The study was approved as Exempt status by the IRB board on June 17th, 2020.

3.2.3 Methods

The interviews were conducted using the video conferencing software GoToMeeting. Each participant was scheduled for a one-to-one session via GoToMeeting, in order to observe social distancing CDC guidelines and recommendations related to the COVID19 pandemic.

Participants were scheduled for the interview and were requested to complete the study consent forms emailed to them at the time of scheduling. They were asked to return the signed consent via email back to the researcher prior to commencing the interview. An interview field guide containing a list of semi-structured interview questions was used to direct the course of the interviews.

A constant comparative method was used during the interview to ensure consistency and completeness in the interviews. With this approach the topics and suggestions raised by the initial interviewee that were not already included in the interview field guide were also provided to remaining interviewees to elicit their view and perspectives about the subject. The interviews were conducted as back and forth exchanges where either party was able to request further clarification on the points covered. The interview sessions were approximately 60 minutes in duration, audio recorded and transcribed using the voice transcription software Otter.ai. (https://otter.ai/). The Otter.ai files were converted to Microsoft Word files and edited to remove extraneous time stamp data and participant personal details.

A qualitative data analysis tool, Quirkos (https://www.quirkos.com/) was used for analysis of the interview transcripts. The transcript files were imported into Quirkos software and analyzed to identify common themes related to features that influence clinical diagnosis when there are transitions of patient care.

3.2.4 Data Synthesis and Analysis

Interviews were conducted with nine participants. Of the nine, seven were attending physicians and two were resident physicians. Unfortunately, the software did not record one of the interviews with an attending physician properly and the recording was significantly truncated. Consequently, this recording was not included in the thematic analysis. A total of eight interview transcripts were imported into the project workspace in the Quirkos software and thematic analysis was conducted

3.2.5 Quirkos: A Visual Qualitative Data Analysis Tool

Quirkos is a cloud based qualitative data analysis tool that has a visual graphical user interface (GUI) where thematic coding can be done on the canvas, the tool's GUI workspace, directly from the narrative text contained in data files such as interview transcript files. Coding for themes is done by dragging and dropping sections of text onto visualizations of code labels called Quirks, represented as spheres on the canvas. The Quirks are color coded and text sections connected to a particular Quirk are highlighted in the color for that Quirk. Text excerpts can be assigned to more than one Quirk, i.e., to more than one code. The Quirk properties can be edited to include name, description, choice of color, and other features such as assigning groups and associations

The Quirkos tool provides a very visual approach to conducting data synthesis from text data using grounded theory approach where the data in the narrative evokes the codes. (Glaser et al., 1968) As more items of text are assigned to a code the Quirk grows in size. Quirks can also be moved to be positionally on the canvas to be closer to other codes based on the user's preferences. For example, grouping similar or related quirks to enable visualizing the formation of relationships or similarities and thus the development of themes. The canvas is very interactive and allows real-time editing of the Quirks. The Quirks can be nested to produce axial coding of Quirks, which provides visualization of hierarchical relationships.

3.2.6 Thematic Coding and Grounded Theory

The transcript files were analyzed, and data synthesis conducted using Grounded theory. Because a question guide was used to conduct the interviews it might suggest this would

direct the analysis approach to one of inductive coding of the transcript data. However, the questions in the interview guide aimed to be open ended to allow participants to express their experiences freely, and in their own words. The thematic coding process involved a hybrid method of coding where a combination of the top down approach deductive coding and inductive coding, where codes are generated from the data, was used. A few initial codes were set at the commencement of coding and then a grounded theory approach was used to code the narrative data from the interviews. This involved creating codes as codes that were invoked by the data. As the physician participants were asked open-ended questions they spoke freely and their comments, statements and opinions provided the basis of the codes generated.

3.2.7 Definition of terms

1. Quirk: The term Quirk is the tool's name for a code label and so the term code will be used throughout to refer to the codes derived from thematic coding via the Quirks in the Quirkos software.

2. Question Driven Codes: some of the codes that were generated from the transcripts were directly related to questions from the interview guide that were asked in the interview session. These codes were question driven codes and so the associated code was edited to reflect question driven code in its properties.

3. Participant Driven Codes: this type of inductive coding was done when codes were created by being evoked from the narrative data of any participant. When a new code was generated by this inductive process it involved creation of a new code, which was edited to include its description, color and any appropriate associations, such as being a participant driven code.

3.2.8 Thematic Coding Activities

1. The first participant transcript was imported into the Quirkos system and was reviewed. The transcript data, which was displayed on the right side of the Canvas, was scanned to identify sections of text from the narratives of the participant.

2. Sections of text that related to the questions were highlighted, dragged, and dropped onto the appropriate Quirk code.

3. As the coding process commenced a few deductive codes were generated by the questions on the interview guide. The sections of text directly related to these questions were then associated with these initial codes.

4.The remainder of the text was reviewed in detail and a similar process of highlighting, dragging and dropping was conducted, with new codes being inductively generated as they were evoked by the text data of the participant's narrative.

5.This was conducted until the entire narrative of the participant was reviewed and coded.

6. The next transcript file was then loaded into the system and the entire process repeated. As many codes were already generated, the data was coded by adding more excerpts of text to the existing codes where appropriate. New codes were generated when the text provided a new concept or statement that was novel compared to the existing codes on the canvas.

7. As the coding process progress, axial coding was conducted to identify and generate themes. This was done by either arranging the codes together in groups in terms of location on the canvas or by nesting the codes within a tree schema denoting the hierarchical relationships. This was conducted until thematic saturation was achieved.

3.2.9 Data Analysis

Once all eight transcripts were coded the project was analyzed to produce different types of reports from the coding results. The data collected encompassed findings obtained from different levels of granularity and perspectives through inductive and deductive analyses. The results of the data analysis provided insights into emergency medicine sign- out, the communication process and information needs of EM physicians. The analysis also identified their perspectives on the role of anchoring bias and factors related to it that impact diagnostic reasoning and sign- out.

The analyses included:

- Identifying themes generated from interview questions
- Identifying participant generated themes
- Identifying the concerns and information needs of clinicians
- Identifying clinician perspectives on the risk factors for anchoring bias during sign-outs

This qualitative coding process allowed the development of a visual representation of the codes and relationships of the concepts related to sign out from the perspective of the emergency medicine clinicians. The result provided an in-depth understanding of the information needs, preferences and actions of these clinicians during and immediately following sign- out based on their experiences.

Chapter 4: Findings

4.1 MIRS Dataset Analysis Study

The dataset extracted from the MIRS system resulted from processing of the three data extract files provided by the clinicians managing the dataset. The extracts contained cases that were de-identified both for patient and provider details.

The files contained fields which included a unique MIRS identifier, gender, age, location, date of visit and a narrative field containing a summary of the patient case. They also contained a field with the contributing factors that were considered to be linked to the resulting error. This data were specified by the clinicians involved in the case during case review. The contributing factors were selected from a predetermined set of categories in the MIRS database by each clinician that was involved in the case and they could select multiple such factors for any given case. Analysis of the contributing factors involved primarily counting the prevalence of each category for each case specified by the clinicians rather than inferring it from the information in the cases.

4.1.1 Descriptive Statistics

The three data files resulted in 65 unique patient cases. The patient records showed approximately even distribution between the genders with 55% of records being for male patients. The age range was relatively broad ranging from 20 years old for the youngest patient to the oldest patient being 89 years of age.

Review of the records for these 65 cases yielded the following results:

Table 3

MIRS Descriptive Statistics

TYPE	VALUE
Number of unique patient cases	65
Male	55%
Female	45%
Age range	20-89yrs
Average age	50.5yrs
Median age	55yrs

Table 3 represents some details about the dataset. The three extract files contained 732 rows of data in total, which contained 65 individual patient cases. Of the 65 cases, 36 were male and 29 were female. The patients ranged in age from 20-89 years, with a median age of 55years.

4.1.2 Analysis Findings

Table 4

MIRS Main Case Types

Case Type	Count	% Cases
Cardiovascular	25	38.5%
Abdominal	18	27.7%
Head	10	15.4%
Trauma	10	15.4%
Sepsis	2	3%
Total	**65**	

Table 4 shows that of the 65 patient cases the most common were cardiovascular cases making up 38.5% of cases followed by abdominal cases making up almost 28% of cases. Cases related to the head and trauma cases were both approximately 15% of cases and a further 3% of cases were related to sepsis.

Table 5 lists the contributing factors first in terms of the count of each contributing factor for the 65 patient cases. Where the case had the same contributing factor listed multiple times the count for that contributing factor represents however many times it was listed. Hence there were 624 instances of the contributing factors and the frequency each factor was listed is shown in the first column labelled 'Count of Contr. Factors'. The next column presents the percentage of the total count of contributing factors for each contributing factor. For example, Inadequate Handoff was listed 77 times across all the cases making up 12.3% of the total count for contributing factor. Then Faulty Information Processing was listed 74 times across all cases.

The next column represents the number of cases where a particular contributing factor is mentioned, and the next column represents the corresponding percentage of cases where the particular contributing factor is mentioned. The most frequently present contributing factor across cases was Inadequate Handoff, which was present in 60 of the 65 cases, 92% of cases. The next most frequently present contributing factors across the 65 cases were Faulty Information Processing (44, 67%), High Workload (40, 61%), Faulty Data Gathering (38, 58%), Atypical Presentation (30, 46%), Inefficient Processes/System Issues (30, 46%) and Complicated Medical History (27, 42%)

Additional analysis was conducted to review the other contributing factors present in cases that had Premature Disposition listed as a contributing factor in the error for the case. Of the 65 cases, 24 cases had Premature Disposition listed. Of these 24 cases the five most frequently co-listed contributing factors are presented in Table 6 below.

Of the 24 cases with premature disposition specified as a contributing factor, all but two cases also listed Inadequate Handoff as a contributing factor. Faulty Data Gathering as well as Faulty Information Processing were both also listed in 20 of the 24 cases where

Premature Disposition was listed. The next most prevalent contributing factors in the 24 Premature Disposition cases were High Workload, Atypical Presentation and Complicated Medical History, which were also present in 16, 15 and 10 cases respectively.

Table 5

MIRS Contributing Factors

Contributing Factor	Count of Contributing Factors	% of all Contributing Factors	# of cases with Contributing Factor	% of cases in which Contributing Factor is present
Inadequate handoff/sign out	77	12.3%	60	92%
Faulty information processing	74	11.9%	44	67%
High workload	58	9.3%	40	61%
Faulty data gathering	53	8.5%	38	58%
Complicated medical history	36	5.8%	27	42%
Atypical presentation	36	5.8%	30	46%
Inefficient processes/system issues	36	5.8%	30	46%
Supervision Failure	34	5.4%	26	40%
Premature Disposition	28	4.5%	24	37%
Interruptions	24	3.8%	22	34%
Non-handoff communication error	23	4.28%	24	37%
Faulty knowledge	22	4.10%	19	29%
Limited History	19	3.54%	12	18%
Faulty workload management	15	2.4%	12	18%
Obesity	10	1.6%	7	11%
Rare condition	5	0.8%	4	6%
Language barrier	5	0.8%	5	8%
Patient non-adherence	4	0.6%	4	6%
Other	50	9.6%	4	6%
N	**624**		**65**	

Table 6

MIRS Premature Disposition

Category	**Count**	**%**
LBJ	14	58%
MHH	10	42%
Male	16	67%
Female	8	33%
Co-reported Contributing Factors		
Faulty data gathering	20	83%
Faulty information processing/verification	20	83%
Inadequate handoff	22	92%
High workload	16	67%
Atypical presentation	15	62.5%
Complicated medical history	10	42%
Other	23	96%

4.2. Emergency Medicine Physician Interviews

The themes identified represents the perspectives of these clinicians and is based on their own observations as well as their responses to the questions asked during the interview. Thematic analysis of the 8 interview transcripts was conducted. A combination of grounded theory and deductive coding was used to code the narrative data in the transcripts. Axial coding of the core codes resulted in identifying main themes, some of which were associated mainly with responses to questions in the interview guide, hence referred to as question generated themes. The themes derived from codes invoked by participants' own thoughts and statements, thus were labelled as participant generated themes.

4.2.1 Descriptive Statistics

Eight physicians were interviewed following the semi-structured interview field guide (see appendix A). The eight interviewees were practicing emergency medicine physicians working as either attending physicians or resident trainees and were affiliated with either UT Physicians or McGovern Medical School. Table 5 describes demographic details about the participants, who were interviewed about their perceptions and experience of sign outs.

Table 7

EM Physician Interviews Participant Information

#	**Position**	**Years of experience**	**Gender**	**Affiliation**
1	Attending	21years– residency + 19yrs	Male	UT Physicians
2	Attending	10years – residency + 7yrs	Male	McGovern Medical school
3	Attending	24years – residency + 21yrs	Female	McGovern Medical school
4	Attending	16years – residency + 13yrs	Female	McGovern Medical school
5	Attending	5years – residency + 2yrs	Male	UT Physicians
6	Attending	13years – residency + 10yrs	Female	McGovern Medical school
7	Resident	3 years	Female	UT Physicians
8	Resident	3years	Male	UT Physicians

Of the eight physicians interviewed, six were attending physicians and two were emergency medicine residents. Participants were required to have at least one year's emergency medicine experience and all participants met this criterion. The two residents were PGY3 status, which means they were in the third year of their residency. All the participants had experience of working in the emergency departments of large urban academic hospitals affiliated with McGovern Medical School.

Table 8

EM Physician Interviews Descriptive Statistics

TYPE	**VALUE**
Participants	8
Attendings	75%
Residents	25%
Male	50%
Female	50%
Years of Experience of Attendings	5-24
Average years of Experience of Attendings	15.7
Years of experience of Residents	3

Table 8 presents some descriptive data about the participant group. There was an exactly balanced split in terms of gender with 50% being male and 50% female. The group was comprised more heavily of senior physicians, with 75% being attendings and only 25% physicians were emergency medicine residents. Consequently, as expected, the attendings had a much higher average years of experience (15.7yrs) compared to the two PGY3 residents, who had 3 years of experience in emergency medicine.

4.2.2 Sign Out Themes

A total of 60 concepts or codes were generated from the thematic analysis. The codes were invoked from the text in the interview transcripts. Axial coding of the initial open codes resulted in 17 main themes, which are listed below. The main themes about sign out (SO) were grouped into higher level themes that represented the following six categories of Information Capture, Information Processing, Aspects of Sign Out, Reliability of the Information, Teams and Training, Workflows and Processes

Table 9

Sign Out Themes, Definitions, Examples and Categories

THEMES		DESCRIPTION	EXAMPLE	QUESTION/ PARTICIPANT GENERATED
Information Capture				
1.	Function/ Overview of SO	What purpose or function do sign outs provide	'The sign out, I think, the intention is that they are trying to relay as much information that is reliable and accurate as possible.'	Question generated – Questions 1& 4
2.	SO Preferences	What does the physician want in SO info transfer	'Start providing the context. So, the story is always first. I like to know the context of the patient, so, before I hear the chief complaint'	Question generated – Question 3
3.	Desired Content Detail	How much detail the receiving physician wants info in SO to be provided	'It really is important to have all that information, every presentation with age, gender and at the least in brief their conditions.'	Question generated – Question 8
4.	Focus Receiving Info	What are you focusing on when you get the patient info in SO	'Definitely, like any test results, or response to medications, consults and recommendations, definitely that.'	Question generated – Question 7
Information Processing				
5.	Recording/Reviewing	What do they do to record the information provided in SO	'see what the lab is, keep writing, that kind of stuff. So, the actual mechanism is for me, is on paper.'	Question generated – Question 5
6.	Verification	What are the factors related to verification of patient information	'I circle important points, and there are open boxes next to the things that I need to verify.'	Participant generated

Aspects of Sign Out				
7.	Factors Impacting SO	The factors the clinician feels affect or are challenges of sign outs	'that's affected by the time of day, because there are some times of day that are reliably more busy than others.'	Question generated – Question 10
8.	Problem Cases	The types of patient cases are problematic to diagnose	'Rare, rare diagnoses that are bifurcating threatening can be very difficult to diagnose.'	Participant generated
9.	Standardization	How does the clinician feel about handoff tools	'We've tried to standardize our sign-up process, because there are data items that have been missed'	Question generated – Question 12
Reliability of Information				
10.	Confidence in SO info	What affects the confidence the receiver has in the information provided in SO	'We take them at their word but we also weigh the evidence to make sure it lines up with it.'	Participant generated
11.	Source	What is the source of the information about the patient	'The information to be radiology is different because it's, it's more data.'	Question generated – Question 6
12.	Anchoring Bias	What factors does the clinician think contributes to instances of anchoring	'...is there a potential chance of getting an anchoring bias from say like a differential diagnosis that you've been given during the signup. I think that there definitely is'	Question generated – Question 14
Teams and Training				
13.	Team Dynamics	The factors that affect or are affected by team interaction/ communication	'I think that any, any group of people who have spent time working together will be better at relaying information to each other.'	Participant generated
14.	Training Environment	Aspects related to the department being an academic environment	'So, our interns and the early in the first part of their first year, we spent a lot of time trying to help them	Participant generated

		where residents receive training in emergency medicine and sign outs	organize their presentations into a cohesive structure'	
Workflows and Processes				
15.	Diagnostic Process	Diagnosis is a process and can be dynamic	'we live, and we work in a very dynamic setting that's always changing '	Participant generated
16.	Eyes on the Patient	Does the oncoming physician see patients signed out to them	'And then and patients that are complicated or early in their course, I will often go in and speak with them.'	Question generated – Question 13
17.	Improvements	What changes would be desired to improve sign outs	'so it might be nice to have like a dedicated space that have some sound insulation'	Question generated – Question 15

Note: The axial code from the open codes led to 17 main themes around the information and communication related to sign outs (SO).

The themes identified were 'Function/Overview of SO', 'SO Preferences', 'Desired Content Detail', 'Focus Receiving Info', 'Information Processing', 'Recording/Reviewing', 'Verification', 'Aspects of Sign Out', 'Factors Impacting SO', 'Problem Cases', 'Standardization', 'Reliability of Information', 'Confidence in SO info', 'Source, Anchoring Bias', 'Teams and Training', 'Team Dynamics', 'Training Environment', 'Workflows and Processes', 'Diagnostic Process', 'Eyes on the Patient' and 'Improvements'.

The codes identified are listed in the table below. The table includes the code name, which is also the same as the label for the Quirk item used on the canvas of the thematic coding software. Also included is a description of the code and which main themes the code is associated with.

Table 10

Codes for Constructs Related to Sign Outs

	Quirk/Code Name	**Description**	**Related Themes(s)**
1.	Patient Safety point	Clinicians view SO is a point for patient safety	FUNCITIONS OF SO
2.	Transfer patient info	Purpose of SO is to transfer patient info	FUNCTIONS OF SO
3.	Diagnosis desired	Clinician wants to be provided a working diagnosis in SO	SO PREFERENCES
4.	Patient info	Want to get patient specific information e.g. presentation and medical history	SO PREFERENCES
5.	My own info needs	Personal preferences in info content and detail	SO PREFERENCES CONFIDENCE IN SO INFO
6.	Workup	What gets done to patient in EM e.g. diagnoses labs, tests, meds. Running tests and imaging to determine the differential diagnosis	SO PREFERENCES DIAGNOSTIC PROCESS CONFIDENCE IN SO INFO
7.	Plan	What the plan for diagnosis and treatment for the patient is	SO PREFERENCES
8.	Patient story	SO information presented like a patient story	SO PREFERENCES CONFIDENCE IN SO INFO
9.	Acuity of patient	How serious the patient's condition is	FOCUS WHEN RECEIVING SO
10.	Stage in the process	Tests, labs and results for the patient up to the point of SO	FOCUS WHEN RECEIVING SO
11.	Medical History	The medical history of the patient e.g. the chronic conditions or past medical conditions	FOCUS WHEN RECEIVING SO
12.	Missing/needed	Workup that is pending	FOCUS WHEN RECEIVING SO
13.	Concurrent review	the clinician considers SO patient info as presented with data	VERIFICATION
14.	Review Later	Clinician packages up SO patient info to consider later	VERIFICATION
15.	Notes	Make notes during sign out for follow up and review – typically on paper	VERIFICATION
16.	Electronic	Uses electronic mechanism to record – direct to EHR or tablet etc.	VERFICATION

17.	Residents	Sources type providing info in SO	SOURCE
18.	Attendings	Source providing info in SO	SOURCE
19.	Nurses	Source providing info during SO	SOURCE
20.	Md-levels	Source providing info in SO	SOURCE
21.	Consultants	Source providing info in SO – e.g. radiologist, neurology	SOURCE
22.	Patient/Family	Source providing info	SOURCE VERIFICATION
23.	Return patients	Patient seen recently in the EM	PROBLEM CASES VERIFICATION
24.	Frequent Flyers	Patients who repeatedly come to EM	PROBLEM CASES FACTORS/CHALLENGES AFFECTING SO
25.	Rare	Patients with unusual or rare cases	PROBLEM CASES DIAGNOSTIC PROCESS
26.	Complicated cases	Patients with more co-morbidities/complex medical history or more complicated presentations	PROBLEM CASES FACTORS/CHALLENGES AFFECTING SO
27.	Trust in individual	Trust in sign out information is related to the person giving sign out	CONFIDENCE IN SO INFO TEAM DYNAMICS
28.	Nonverbals	Tonal elements not related to spoken words that convey communication	CONFIDENCE IN SO INFO
29.	Match with patient story	If the test and imaging results received in sign out are consistent with overall patient story	CONFIDENCE IN SO INFO
30.	All patients	See all patients received at sign out; new with no issues, with issues, with disposition already agreed	EYES ON PATIENT IMPROVEMENTS
31.	Patients with concerns	See only patients about whom physician has concerns e.g. diagnosis seems inconsistent with work up or is very sick	CONFIDENCE IN SO INFO EYES ON PATIENT
32.	Active patients	Patients without a disposition transferred in SO and needed active management	EYES ON PATIENT
33.	Questions during SO	verify with the SO provider at the time of SO	CONFIDENCE IN SO INFO VERIFICATION
34.	Patient record	Go through EHR notes or patient record	VERIFICATION
35.	Labs & tests	Results of laboratory and other diagnostic tests	VERIFICATION

36.	Orders	Exiting teams' orders for the patient	VERIFICATION DO NOT VERIFY
37.	Discharged Pts	Oncoming clinician does not verify or follow up disposition of patients	DO NOT VERIFY
38.	Case Specific	Amount or type of info depends on the patient's case e.g. how serious the patient is	DESIRED CONTENT
39.	Pertinent info	Only information that is relevant and pertinent to the situation	DESIRED CONTENT
40.	More detail	How much information do they like to receive in a sign out	DESIRED CONTENT
41.	More Attention	When the clinician gives more attention to in the sign out	CONFIDENCE IN SO INFO
42.	Less Attention	When the clinician pays less attention to the information in the sign out	CONFIDENCE IN SO INFO
43.	Structured SO info	Prefer info provided in a structured format	STANDARDIZATION TRAINING ENVIRONMENT
44.	Disorganized SO	When the information is given in a disorganized and unstructured way during the sign out	CONFIDENCE IN SO INFO FACTORS/ CHALLENGES AFFECTING SO
45.	Information loss	When important information transferred in SO can get lost	STANDARDIZATION FACTORS/CHALLENGES AFFECTING SO
46.	Better Communication	What does not get communicate in SO but that should have been	IMPROVEMENTS TEAM DYNAMICS
47.	Incorrect Diagnosis	Diagnosis turns out to be different from diagnosis at SO	DIAGNOSTIC PROCESS
48.	Change path	Oncoming physician takes steps or considers taking the case in a different direction to that of the exiting team	FACTORS/CHALLENGES AFFECTING SO DIAGNOSTIC PROCESS
49.	High Workload	When department is busy or there are a lot of complex cases	FACTORS/CHALLENGES AFFECTING SO
50.	Interruptions	Interruptions during the sign out process e.g. from nurses needing sign off on orders or calls	FACTORS/CHALLENGES AFFECTING SO
51.	Compression	Patients later in the sign out get less time	FACTORS/CHALLENGES AFFECTING SO
52.	Data Resident	A resident physician who has been appointed to track and present health data during the SO	STANDARDIZATION

53.	Patient order	The order that patients are signed out	PROBLEM CASES STANDARDIZATION TRAINING ENVIRONMENT
54.	Workarounds	The approaches adopted to support sign out specially to mitigate for interruptions	IMPROVEMENTS
55.	Anchoring risk	What contributes to anchoring and instances of anchoring	ANCHORING BIAS CONFIDENCE IN SO INFO PROBLEM CASES
56.	Avoiding bias	Participants actions to avoid or consider bias in their decision making	ANCHORING BIAS FUNCITIONS OF SO
57.	COVID	The effect of COVID infection on the emergency department and on sign outs	ANCHORING BIAS FACTORS/CHALLENGES AFFECTING SO
58.	Training residents	Things to help residents e.g. give oncoming residents pointers to help them	CONFIDENCE IN SO INFO TRAINING ENVIRONMENT
59.	Diplomacy/discretion	Trying not to criticize or embarrass the provider if their assessment may be flawed	TEAM DYNAMICS
60.	Team Familiarity	Familiarity with team members influences SO	CONFIDENCE IN SO INFO TEAM DYNAMICS TRAINING ENVIRONMENT

The codes developed from the narrative were from the narrative data in the transcripts. They included constructs about how and what the physicians liked to receive in terms of the information such as ‘Work up’, ‘Patient story’ and ‘Plan’. The codes ‘Stage in the process’, ‘Medical History’, ‘Acuity of the patient’ and ‘Missing/Needed’ were constructs related to how they used the information translating it into the aspects they focused on about the patient’s course in the ED.

The constructs ‘Review Later, ‘Notes’ and ‘Electronic’ ‘Question during SO’ were related to what the participants do to verify the information they received in sign out e.g., ‘Notes’ is about making notes on paper.

There were also codes around the theme of 'Confidence in the SO', which were about what influenced the confidence the participant felt in the information being transferred at sign out. The constructs 'Patients with concerns', 'Disorganized SO', 'Match with patient story' were about factors that contributed to confidence in the reliability of the information. There were also constructs related to the challenges of sign outs in the ED and were represented by codes such as 'High Workload', 'Compression' and 'Interruptions'.

There were also codes associated with the core theme of 'Standardization' that related to efforts to standardize sign out such as 'Data Resident', which is a practice where a resident is designated to check and present test results and other data from the EHR during sign out. There was also 'Structured SO info' and 'Patient order' which was a code that denoted the way sign out is conducted following the order of patients' bed numbers.

The theme of 'Source' contained codes that represented the different sources for the information transferred and discussed at sign out and during verification e.g., 'Residents', 'Attendings', 'Nurses', 'Mid-levels', 'Consultants', 'Patient/Family'.

'Better communication', 'Team familiarity' and 'Diplomacy/discretion' were codes for concepts related to 'Team Dynamics', a theme describing participants views on the factors related to how the physicians communicate and build rapport in the ED and the effect this has on sign outs. Then there were the codes such as 'Training Residents' that were related to the theme of 'Training Environment'. This theme also included the codes of 'Patient order' and 'Structured SO info' highlighting the efforts to instil standardized practices in the residents who are doctors in training. Some codes like 'Structured SO info' were associated with multiple themes.

There were also a number of codes such as ‘Better Communication’, ‘Workarounds’ that were related to areas for improvement that participants felt might help with the challenges they experienced during sign out in the ED.

4.3. Overlap Analysis: Anchoring Bias Results

Analysis to identify the codes and themes that were related to anchoring bias was conducted. The Quirkos software enables a report that maps the themes and codes related to a specified code called an Overlap Analysis. An overlap analysis report for the theme ‘Anchoring Bias’ yielded the following results shown in figure 2.

.

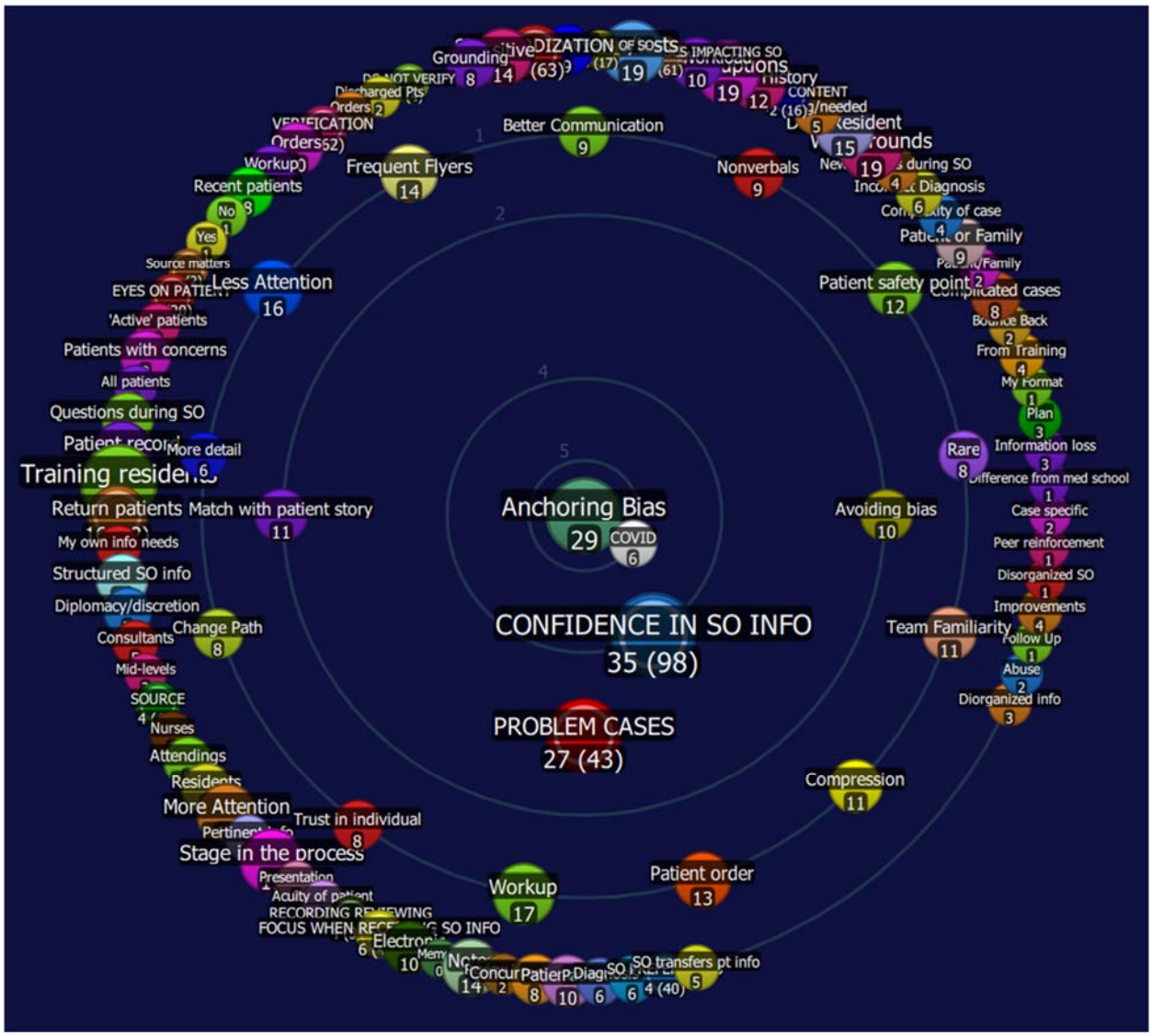


Figure 2: Anchoring Bias Overlap Analysis

This overlap analysis (figure 2) is an illustration of the theme of 'Anchoring Bias' and the themes and codes that are related to it. 'Anchoring Bias' is in the center of the image. The proximity of the other codes and themes indicates how closely they are related. The code 'COVID' is most closely related followed by 'Confidence in SO info' and then 'Problem cases'. The concentric circles then have other codes that are related but less strongly.

Table 11

Themes and Code Related to Anchoring Bias

Proximity/ Association	**Code/Theme**	**Definition**
1st level	COVID	The effect of COVID infection on the emergency department and on sign outs
2nd level	Confidence in SO	What affects the confidence the receiver has in the information provided in SO
	Problem cases	The types of patient cases that are problematic to diagnose
3rd level	Match with patient story	If the test and imaging results received in sign out are consistent with overall patient story
	Avoiding bias	Participants actions to avoid or consider bias in their decision making
4th level	Frequent flyers	Patients who repeatedly come to the emergency department
	Trust in individual	Trust in sign out information is related to the person giving sign out
	Team familiarity	Familiarity with team members influences SO
	Change in path	Oncoming physician takes steps or considers taking the case in a different direction to that of the exiting team
	Rare	Patients with unusual or rare cases
	Compression	Patients later in the sign out get less time
	Patient order	The order that patients are signed out
	Non-verbals	Tonal elements not related to spoken words that convey communication
	Workup	What gets done to patient in ED e.g. diagnoses, labs, tests, meds. Running tests
	Better communication	What does not get communicate in SO but that should have been
	Patient safety	Clinicians view SO is a point for patient safety
	Less Attention	When the clinician pays less attention to the information in the sign out
	More Attention	When the clinician gives more attention to in the sign out

The results of the overlap analysis (table 12) show that the code ‘COVID’ was closely associated with anchoring bias. The construct ‘COVID’ referred to the presence or suspected presence of a COVID-19, an infection, caused by SARS-CoV2 virus, in a patient case. Suspicion of ‘COVID’ positivity was felt to act like an anchoring bias for diagnostic reasoning in recent times due to the high number of patients coming to the ED with symptoms that resemble those of the respiratory system infection observed during the COVID-19 pandemic that began in late 2019. The next most common related construct was the theme of ‘Confidence in SO’, which consisted of codes like ‘Trust in individual’, ‘Nonverbals’, ‘Patients with concerns’ and ‘Questions during SO’. Also, at this second level of correlation was the theme of ‘Problem cases’. The third level of related codes were ‘Match with patient story’ and ‘Avoiding bias’. The code of ‘Avoiding bias’ represented the thoughts and actions that participants mentioned about their attempts to avoid or mitigate for anchoring bias during their practice in the ED.

4.4 Themes across Physician Responses

Figure 3: Data Related Codes for sign out data and information for Attendings and Residents

The codes related to the data of sign outs are shown in figure 4. The majority of the physicians felt that sign outs provided an opportunity for a patient safety point. In terms of their preferences, half of interviewees said they wanted details about the patient and that a working diagnosis was desired as part of the information transfer. Most said they wanted to know about the workup that had been conducted for the patient up to the point of sign out. All the participants said they preferred the patient story rather than just getting the information as data items.

For the desired level of content there was a preference for more details but only pertinent data by the attendings. Receiving the data in a structured format was desired by half the attendings and both residents. In terms of the focus for recording the information received all participants focused on the stage in the diagnostic process of the patient and the majority focused on the patient's medical history. Two attendings and both residents said that they focused on what was missing or pending i.e. lab results or imaging for the patient in the information that transferred during sign out. Fifty percent of the participants expressed a desire to conduct the sign outpatient list starting with the sickest first rather than bed order as is usually practiced in the ED.

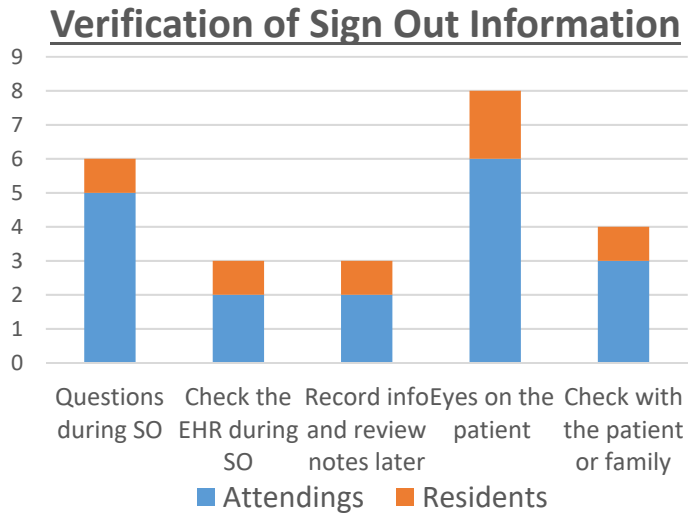


Figure 4: Verification of Sign Out Information – Codes related how to attendings, and residents verify information received at sign out

Figure 5 charts the various methods by which the participants stated they verified the information received at sign out. The majority said they asked questions during sign out, while some checked information directly in the EHR. Two attendings and one resident said they recorded the information with the intent to review their notes at a later time. All the participants stated they made some effort to go to see patients, while half said they also checked with the patient or

their family/caregivers.

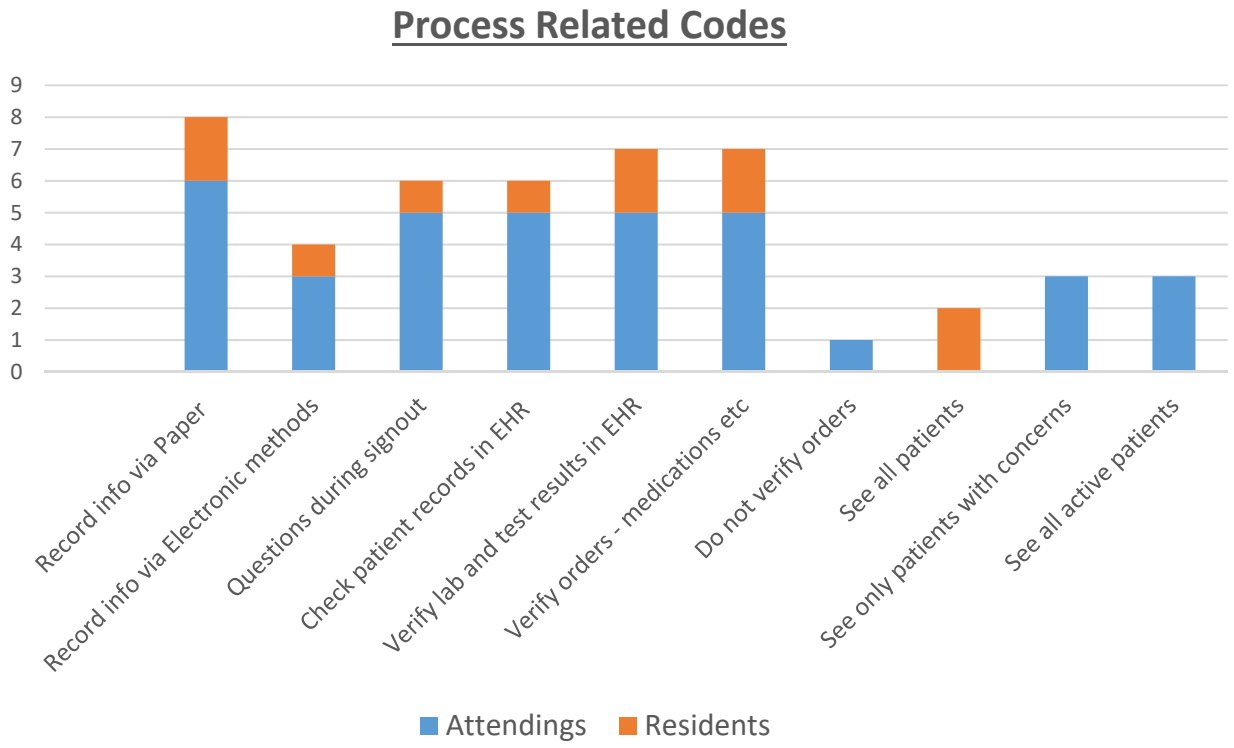


Figure 5: Process Related Codes – Codes related to the process associated with recording and verifying sign out information

This chart shows the different activities conducted by the physicians by role. When recording information provided during sign out all the interviewees stated that they used paper to make notes and only half of attendings and residents said they recorded information using electronic means, such as iPads. In terms of verifying information most said they asked questions during sign out and also said that they checked the patient record in the EHR directly. When they did verify information most focused on checking orders and medications as well as test results, such as lab test results. In terms of verifying information by seeing the patient, there were different approaches depending on the role with only residents saying they would see all the patients transferred during sign out.

4.5 Schematic of Themes and Codes

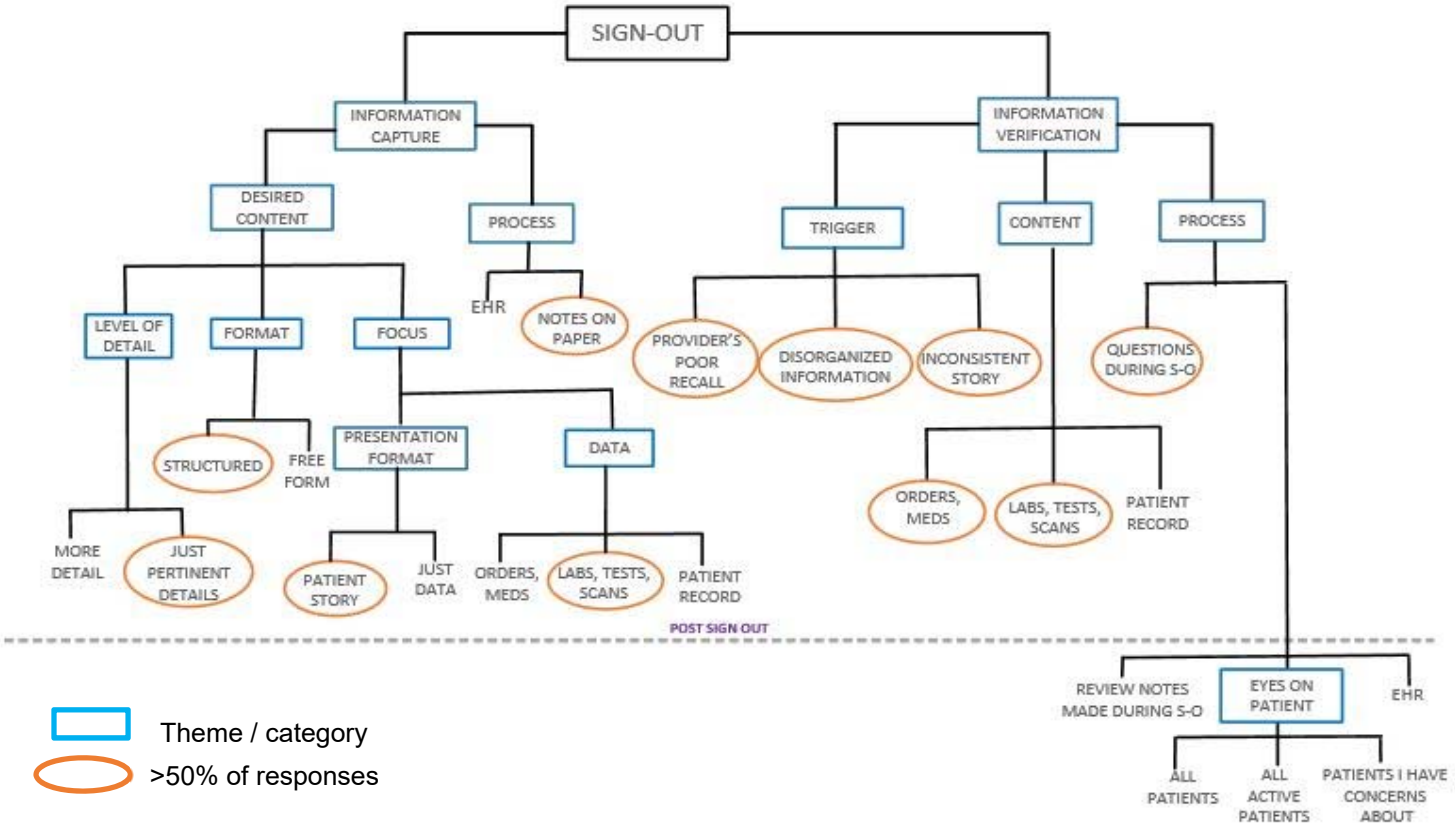


Figure 6: Schematic of Themes and Codes Related to Sign Out

Figure 6 shows a schematic of the themes and codes related to sign out represents the key findings. Participants prefered sign out information provided in a structured way, prioritizing just the pertinent details presented in the format of a patient story. The participants took notes on paper and focused on noting data like tests, scans and lab results. The factors that would trigger a desire to verify the information provided would be if the outgoing provder had poor recall of the patient details, the story the outgoing provider presented contained inconsistencies or was disorganized in its presentation. They would verify information via questions back and forth during sign out and would check on things like orders, medications, or test results. Most would see the patient but this would be after the sign out was completed.

Chapter 5: Aim 1 Conclusions, Discussions, and Recommendations

The first phase of this project aimed to identify potential risk factors that might contribute to anchoring during diagnostic decision making and treatment planning related to sign outs in the emergency department (ED). A combination of qualitative methods was used collect data related to emergency medicine patient cases and the experiences of emergency medicine clinicians. This data was then analyzed to identify features or factors that could be considered as candidates for potential anchoring risk factors. The first study involved the analysis of data from known instances of patient cases containing diagnostic error that were documented in the Medical Incident Reporting System (MIRS) database. The second study involved a thematic analysis of a series of interviews with emergency medicine physicians to gain insights into their information needs and preferences. The interviews were around their perspective of the challenges involved with sign outs based on their experiences working in the ED.

5.1 MIRS Dataset

The 65 unique patient cases within the MIRS dataset were reviewed and analyzed. The analysis aimed to identify common features within the cases that could be considered as risk factors for errors. The intention was to then use these identified features in the design of an experimental study to assess their impact on clinician decision making.

5.1.1 Discussion of Findings

Analysis of the data shows that the dataset was reasonably balanced in terms of gender with 55% of the patients being male and 45% female. No cases with infants or young children were seen in this MIRS dataset. With the age range being 20-89 years across all the cases and a median age of 55 years, the dataset appears to be relatively representative of the general ED patient population. (McCaig & Nawar, 2006) This suggested that diagnostic or management errors in the MIRS dataset were not driven by the gender or age of the patients and this sample was within the typical gender and age distributions of emergency department patient populations.

The analysis of the contributing factors found that over 90% of the cases were associated with an Inadequate Handoff. The next most frequent contributing factors both in terms of overall prevalence and across the cases were Faulty Information Processing, High Workload and Faulty Data Gathering. Given that cases being examined were selected on the basis of containing a sign out and instance of error, factors such as Inadequate Handoff, Faulty Data Gathering and Faulty Information Processing might be somewhat expected. Indeed, breakdowns in communication during transitions of care are known to be a significant cause of medical error. (Abraham et al., 2011; V. Arora et al., 2005; Cheung et al., 2010) The findings show that other factors about the cases were also considered contributory to error, such as a Complicated Medical History and Atypical Presentation, both being the next most prevalent factors. Although much less frequently specified, Rare Medical Condition was also mentioned as a contributory factor.

The findings also suggest that a host of system or environment related issues were common to many cases of error. For example, high workload, interruptions, and

inefficient processes/system issues, all were specified both frequently within each case and across a high number of cases. The aim of this work was to identify features that could be incorporated into an experimental design. Emergency departments are known to be busy, dynamic, complex, and stressful environments and these results confirm this. Although these findings are interesting in themselves, these suggested features cannot be easily incorporated or reliably replicated in an experimental design. Similarly, there were some factors related to individuals that would not be easy to include or reliably replicate in an experimental design. These include patient features such as providing a Limited History or Language Barrier. There are also some provider behavior related features such as Supervision Failure, Faulty Knowledge, Faulty Workload Management and Non-handoff Communication Errors that are very interesting from the perspective of a root cause analysis of ED errors but do not necessarily provide features that are useful for the design of an experiment looking to identify risk factors for anchoring.

The most common chief complaint types were cardiovascular cases making up almost 40% of cases, followed by abdominal cases (18%). Head and trauma cases made up around 15% of cases each. These figures were slightly different to the types of cases reported in the literature on national statistics for US emergency departments patient visits. The national emergency department statistics indicate the most common types of medical conditions that patients present with are abdominal pain, chest pain, fever and trauma in that order. (McCaig & Nawar, 2006) However, the MIRS data suggests that cardiovascular cases were the most common chief complaints, followed by abdominal cases. Next are head and trauma cases in similar numbers. As the experiment design for specific Aim 2 involved developing patient case vignettes, these findings about chief complaint types were useful for informing the design of the test cases.

Finally, the category of Premature Disposition was considered as a potential indicator of premature closure by the clinician decision maker. The cases where Premature Disposition was specified were analyzed further. The results showed that the cases were roughly evenly distributed across the two clinical sites but occurred twice as much in male patients than female. The most common co-listed contributing factors were like the rest of the MIRS dataset. Inadequate handoff was present in almost all the cases, followed by faulty data gathering and faulty information processing and high workload. Results indicated that atypical presentation and complicated medical history were present in the cases where the patients had been given a premature disposition.

The contributing factors were grouped by three categories: provider factors, system factors and patient/case factors. This allowed easier review of the factors when considering the significance of factors in the design of the experiment.

Provider factors:

- Inadequate Handoff/Sign Out – 92%
- Faulty Information Processing – 67%
- Faulty Information Gathering – 58%
- Supervision Failure – 40%
- Premature Disposition – 37%
- Non-handoff Communication Error – 37%
- Faulty Knowledge – 29%
- Faulty workload Management – 18%

System Factors:

- High Workload – 61%
- Inefficient Processes & System Issues – 46%
- Interruptions – 34%

Patient/Case factors:

- Atypical Presentation – 46%
- Complicated Medical History – 42%
- Limited History- 18%
- Obesity – 11%
- Language Barrier – 8%
- Rare Condition – 6%

5.1.2 Limitations

While the analysis of the MIRS dataset did suggest some features that could be used to inform the design on an experiment, there are some limitations to this study that should be mentioned. Firstly, the MIRS dataset was limited in sample size with only 65 unique patient cases resulting from the data extracts provided, making inferences from the analysis limited in their generalizability. Also, the cases were heavily de-identified, so it was not possible to determine other types of information might have been useful, such as how many or which types of clinicians worked on any given case or which assessments were provided by them at case review. In addition, there was wide variability across the cases in terms of the level of information included in the narrative field about the case. This field often contained information such as treatment details, patient course, patient response to treatment, timings, outcomes and other details about the case. Some of this informative data was collected and coding attempted but as the level of detail and type of information was not consistently present across the cases in the dataset, the analysis for these aspects and features was not possible.

It should be mentioned also that the contributing factors field was populated by data input selected from a set of pre-determined categories by the clinicians involved in the case at the time of case review. The timeframe for when the case was reviewed in relation to the management of the case was not available but was assumed to be some time after the actual patient encounter. Hence the data related to the case review may be subject to recall and hindsight bias issues that apply to all post hoc review of cases. Additionally, because the contributing factor was a selection from drop down with a list pre-determined options, it may not have met all the information needs of the clinicians conducting the case review. This is supported by the finding that the option of other i.e. not any of the options provided, made up almost 10% the contributing factors selected. This was so for 6% of the cases. Nonetheless, this study focused primarily on the contributing factors and case specific data that was available and the findings were still useful for informing the design of the experiment for Specific Aim 2.

5.1.3 Conclusions

In conclusion, the chief complaint of the case such as cardiovascular, abdominal pain, head and trauma cases could be an important feature to include in design of the cases in any experiment to assess the effect of risk factor. By including the chief complaints identified from the MIRS cases, the experimental design is more likely to reflect both the frequency and the potential ambiguity in these types of cases that is experienced in the ED. Similarly, factors such as atypical presentation and complicated medical history should also be considered in the design. The design of the cases should also reflect the age and gender distributions seen in general emergency department populations. In other words, the cases in the experiment should reflect an even balance across gender and an

age distribution that does not include patients at the extremes of age ranges such as infants, young children, and extremely geriatric patients.

5.2 EM Physician Interviews

The aim of this interview study was to gain insights on the perceptions and experiences of emergency medicine physicians in terms of the information they transferred and received during sign out.

5.2.1 Discussion of findings

The results identified a number of themes around their preferences and focus for the information they received, the factors that affected their confidence in the reliability of the information, the steps they took to verify information and what they believed might contribute to anchoring bias.

5.2.2 Function of SO

As an opener to the interview, participants were asked what purpose they felt sign outs fulfilled in emergency departments. The responses were described by question generated theme of 'Function of SO', which included the idea that sign out allowed the transfer of patient information and responsibility e.g.

> "*I think it's definitely designed to pass along the case and status of the patient, but also to prevent any missteps along the way*." And
>
> "*And so, as at the end of the time, you look to see who's assigned to every patient in the department. If they're still some of the off-going people left, there has to be a very good reason why your new team is not assigned to those patients*."

However, along with the transition of patient care and information, participants felt that sign out also provided a patient safety point in the delivery of care. They felt that sign out allowed an opportunity to re-assess the case and treatment plan in order to prevent or address potential errors, e.g.

> *"the new team coming on comes with a different perspective and more energy and a fresher mind. And so, there can be things that differ."*

Studies have indicated that handoffs such as sign outs are viewed as a risk for medical error (Abraham et al., 2011; *Sentinel Event | Joint Commission*, 2007.) However, these findings suggest that clinicians are cognizant of the inherent challenges of emergency medicine cases and transitions of care. It suggests they seek to mitigate for them and recognize sign outs as means to catch things missed in the case or adjust the plan or correct the diagnostic path. Studies show that physicians do utilize cognitive interventions such as active reflective practice and metacognition to review their cases in the moment to ensure weighing up the evidence appropriately during decision making. (Graber et al., 2014) The awareness that in addition to a transition of care, the role of handoffs can be expanded to include patient safety opportunities is growing. (Gogan et al., 2013)

5.2.3 Information Capture

A number of themes identified were related to the overarching theme of information capture and were about participants' preferences. For example, the theme of 'SO Preferences' related to what information the participant wished to receive during a sign out and there seemed to be a preference amongst attendings and residents (see figure 4) to get a story about the patient rather than just factual data points e.g.

"The story, and then the ideas that prompt to me. I'll look to try to reconcile them with the labs or the studies, that I hear. But the main driver of my thought process is the, by far, is the story."

Additionally, some participants did also wish to receive information about the clinical reasoning of the provider, suggesting the desire to develop a shared mental model about the patient case e.g.

"Like, the top diagnoses they're considering, the top of emergent diagnoses that they're considering. Because then that would lead to other workflow".

Another theme around information capture was 'Desired Content Detail' that was about participant's preferences in the level of detail and granularity of the information provided. There was an overall preference for pertinent information only over extensive overly detailed information e.g.

"Oftentimes, junior learners will, when they first start presenting, they often present too much information, but not necessarily too much. It's just too much of the information that we don't want to know. Right!"

This view of the amount of information transferred may reflect the limited time available to conduct sign outs and the volume of patients that may require sign out at busy times. This is possibly due to the need to balance the conciseness or relevance of the information with its completeness. Alternatively, it may suggest that there can be mismatches in preference between the provider and the receiver of the sign out. This mismatch could be for the level of detail included. Or it could even be a mismatch between experienced clinicians, who prefer only pertinent information (see figure 4) and

junior staff, who are still learning how to sort and prioritize the information they transfer during handoffs. (V. Arora et al., 2005; Cheung et al., 2010)

Another theme related to information capture was 'Focus Receiving Info' which covered what aspects of the sign out information that clinicians focused on at the time of receiving the information. This theme highlighted that the physicians did focus on details like the patient medical history and laboratory test values e.g.

> "*But a lot of times, it's looking and also making sure that the labs in the finals and history of the patient and the results are fitting with that with that mental model, that provider's saying."*

Additionally, both attendings and residents (see figure 4) stated that they focused on not just the details being provided but also on what tests and results were still outstanding e.g.

> *"I try to understand what has already been accomplished, or completed in the, in the case so far, I think that to me that that's more important. When imaging or what labs have not been completed"* and
>
> *"so I think with that was the focus for me is the what's missing, just to figure out the next steps. But a lot of it is within the context of the acuity of the patient."*

This suggests that physicians focus on the components of the sign out to support the development of their mental model about the patient in terms of patient acuity and the stage in the diagnostic process.

5.2.4 Information Processing

There were two main themes identified around the activities that the clinicians performed as part of processing the information received. These themes were

'Recording/Reviewing' and 'Verification'. Participants stated that they took notes during sign out to make sure they recorded important information to allow them to review later if need. 'Recording/Reviewing' contained many different codes such as 'Notes' and 'Electronic', which represented the mechanisms they used for recording information e.g.

> *"Anything that I know that I would want to know, that they pass on to me, I will jot down as a note, if, I think I won't remember it off the top of my head."* And
>
> *"So, I use a hybrid of paper and the I for sign out when I am taking sign out when I am the new team. I write down the patient's name and a little blurb about them and then like what is pending or what the disposition is."*

The majority said they used made notes on paper (see figure 6). This may be due to the way sign out is conducted in huddled groups in the ED making it harder for everyone to have access to the I, or it may be due to individual preferences for artefacts that aid personal information needs.

The other main theme was 'Verification', which was about the factors that prompted clinicians to check or verify the accuracy of the information received e.g.

> *"I circle important points, and there are open boxes next to the things that I need to verify"* and
>
> *"Sometimes we'll try to verify or confirm reality separately. We look at the images ourselves and say. Yes, it does look like XYZ."*

The majority of attendings and residents said they verified information like the laboratory tests and results as well as verifying orders and medications for the patient. (see figure 6). There was also a code 'Do not Verify' that captured the things that did not prompt verification and a few stated that they did not typically take steps to verify sign out information e.g.

"*No, no, I take their word for it, if they said they did it, or it's in place."*

And one clinician specifically stated they did not verify orders (see figure 6).

5.2.5 Aspects of Sign Out

Three main themes identified that were related to general aspects of sign out were 'Factors Impacting SO', 'Problem Cases' and 'Standardization'. The first 'Factors Impacting/Challenges of SO', was more about aspects of either sign out or the challenges of the fast-paced nature of the ED and included codes like 'Complexity of Case' e.g.

> "*If you've got very complicated, very sick people, and the number of those very comprehensive people goes up, sign out times are going to take longer."*

And the code of 'High Workload' e.g.

> *"Number of patients, is also one of the biggest factors with you, get fatigued as we go through sign-outs."*

One of the consequences of high workload and the dynamic nature of the ED is the risk of patient cases receiving less time at sign out e.g.

> *"I'd be very interested to see how sign our communication differs from the first by patients. The last, you know, five patients in a 50 patient sign out. You know, that's just those last five patients. We know this, we all know this. Those last five patients get a short-change sign-up, it's just the truth, right?"*

This may in part be due to unexpected increased patient volumes at sign out resulting in compression of time allocated to patients e.g.

"I think we try to be consistent with all the patients on the list, but as the department changes during sign up, and then those last few, if the place blows up during sign up, then that's when they risk getting more compressed."

This is a known phenomenon that has been observed in critical care setting handoffs also. (Abraham, Kannampallil, Patel, et al., 2016; T. Kannampallil et al., 2011)

Then there are other challenges that clinicians face that interfere with sign out, such as 'Interruptions' e.g.

"yes, every five minutes to sign an EKG, to answer a phone call about a patient, or from pharmacy, or from the lab, or from the Transfer Center, or from a thousand other people who have reason to speak".

Despite being disruptive these interruptions are viewed as a necessary part of emergency medicine e.g.

"People say, Oh, don't interrupt them doing sign-out. Well, we need the interruptions, because they're usually very important to do so. Where's the solution there? I don't know."

These aspects of the ED and sign outs, such as interruptions and surges in patient volumes are contributory factors for not just for high workload and clinician stress, but also possibly for medical errors in the ED also. (Maragh-Bass et al., 2017)

The theme of 'Problem Cases' incorporated constructs such as 'Return patients', patients who return to the ED possibly because something got overlooked or they are very sick e.g.

"I let them know when patients are what we call a bounce back, that they were seen, that this is the preliminary diagnosis, they were discharged, and they are back again".

There are also 'Frequent flyers', patients who repeatedly come into the ED e.g.

"I'll also usually point out to the oncoming resident that, you know, yeah, he's here all the time."

Repeat patients are known to place a burden on emergency department resources and staff. (Cook et al., 2004) Then there are 'Complicated cases', where patients have complex medical histories e.g.

''Some patients are very complicated. Have significant complicated, confounding issues. I would say the sicker the patient is, the more information. That, if they are sicker, we want to know more."

Problem cases also includes the construct 'Rare' for cases that are not typical or commonly seen conditions e.g.

"*Rare, rare diagnoses that are bifurcating threatening can be very difficult to diagnose."*

The last main theme in this category was that of 'Standardization', which is about the efforts and protocols to standardize sign out e.g.

"However, they're drafted, they emphasize granularity of data, quantity of data, and relevance of data."

The use of handoff protocols is not mandated in hospitals affiliated with McGovern Medical School and therefore are not used consistently. The theme of 'Standardization' included constructs like 'Data resident', which is a resident assigned the task of checking and providing lab and test results during sign out e.g.

> *"Because I was a part of developing the sign-out process and the idea of the data resident. I always say, 'Who is the data resident?' and that task is assigned before we start."*

While there were a few negative feelings about the use of handoff protocols, on the whole they were considered to be a positive tool in sign outs e.g.

> *"We've tried to standardize our sign-out process, because there are data items that have been missed, either things that the off going provider feels have resulted, but they may be mistaken, or action items that have been performed. That have not been performed."*

The construct of 'Structured SO Info' referred to the concept of presenting sign out information in a consistent structured format, which it was felt reduced 'Information Loss' e.g.

> "*it really does help organize them. It decreases lost information or not communicated information. Yeah, and so it helps organize the whole handoff, and it totally, it totally works,"*

5.2.6 Reliability of Information

The main themes of 'Confidence in SO info', 'Source' and 'Anchoring Bias were under the category of Reliability of Information, which primarily related to trust in the

information and those providing it. The theme of 'Anchoring Bias' will be discussed later. 'Source' was a theme that was related to the various people that can provide input to the sign out information and included attendings, residents, advance practice providers, nurses, consultants and even family of the patient. Some participants also stated that the source of the information mattered in how they considered it,

"I think the more familiar you are with the source, the more reliable you feel."

'Confidence in SO info' was a major theme that covered the factors that affected how reliable the participants felt the information received at sign out was. It contained constructs such as 'Match with patient story' e.g.

''We look at the images ourselves and say, Yes, it does look like XYZ. Sometimes we don't believe it because we know more about the patient story''

as well as codes such as 'Trust in individual', which is about assessing the reliability of the person providing the sign out information e.g.

"We trust colleagues have done the due diligence to evaluate and investing with trust what they're reporting to us in their findings."

This theme also contained the construct 'Non-verbals' which referred to actions or gestures participants made to convey their intent without using spoken words e.g.

"*the language gets a little shorter and more pointed and you may see someone actually just stare at their computer, as opposed to looking at the person that's speaking. Hopefully there's not too much eye rolling, but that does occur occasionally."*

How confident the receiving physician felt about the information strongly influenced whether they then went on to check out and verify the information provided e.g.

> *"we kind of, we take everything, and we see if it has good face validity. And then we see, you know, there's corroborating information in the labs, and the task or, if there's things that disagree with the plan, or the testing plan".*

The ways in which they would do this is by either asking questions during sign out, checking the patient record in the I, go to see the patient or check with the patient's family (see figure 5). In addition to the whether the data presented was reasonable and matched the patient story, the manner in which it was presented also influenced how much trust the receiving physician placed in the information e.g.

> *"If the story is disorganized, and they're just throwing random details at me, I'm much less confident in the sign out and in their thought process"*

5.2.7 Teams and Training

There were two main themes associated with the teams and training. 'Team Dynamics' was associated with the factors that affect teams and team communications. While 'Training Environment' was associated with the teaching requirements of academic hospitals and the need to provide training and development for residents in the emergency department. Many felt that familiarity within teams and with providers of sign out was beneficial in terms of better communications and the development of shared mental models e.g.

> *"I think there's a bias to trust providers, that you have a longstanding history with"* and *"some of these faculty were actually my teachers, as well. And so, they*

are the ones who helped me develop a script, and so we have spent 15 years speaking, the same language."

There was also awareness of the importance of supporting the development of physicians' training in emergency medicine, such as with the code 'Training Residents' e.g.

"You know, processing the information, and organize it, especially for young learners." And *"We do it as a group, and so usually it is the residents who are giving the sign out, with faculty interjecting with either pertinent information and or clarifications on what the resident is saying."*

Indeed, there was recognition of the history that develops as physicians progress through their training and professional development. This familiarity was felt to influence not just their communication but also their working practices e.g.

"If I know the provider that's telling me about a patient, I know their capabilities and their level of diligence there, you know their thoroughness. I'm much more comfortable with the plan".

5.2.8 Workflow and Process

The last main category of themes had to do with the workflows and processes in the emergency department and included the three core themes of 'Diagnostic Process', 'Eyes on the Patient' and 'Improvements'.

The theme of 'Diagnostic Process' was related to the understanding that developing a diagnosis is a multistep process that included a 'Workup', a code representing the process of labs and other diagnostic tests e.g.

"they had an initial, you know, differential diagnosis they whittled it down with their testing to a presumed diagnosis, and they've admitted the patient to the

hospital. Let's say what's the percentage of time that a person's diagnosis changes? Maybe like 10% of the time".

There is also recognition of the dynamic nature of the diagnostic process and the need to be adaptive to this e.g.

"And part of emergency medicine and medicine in general is the ability to be flexible and weighing a clinical position or clinical course changes, you have to allow yourself to cognitively go that direction."

Patients frequently present in the ED with limited information about them available for those treating them. This, coupled with the need to stabilize seriously ill patients promptly, means EM physicians are often required to practice in situations with a high degree of uncertainty, to a level that is unlike any other field of medicine. This can be challenging for physicians as medical training is so exacting and requires them to be highly knowledgeable and skilled. –

"So, it's very important to communicate your level of uncertainty or concern to your next team, and that does take humility. And it takes kind of constantly saying that you don't know things, and that's hard, especially, I mean, honestly, especially for, for physicians, they're not, we're not taught that, saying, we don't know is OK. … We don't really train people for uncertainty in our system."

Studies show that there is much focus on diagnostic accuracy and confidence in understanding the diagnostic process during medical training. However due to many factors, including the fragmented nature of healthcare, practicing physicians rarely receive feedback about their diagnostic record. Indeed, there was mismatch observed in terms of diagnostic accuracy, case complexity and their confidence in their diagnoses. (V. Arora et al., 2005; Meyer et al., 2013)

The factors that influenced whether and when physicians saw the patients in their care was described by the theme of 'Eyes on the Patient' and included the constructs 'Patients with concern' 'Active patients' and 'All patients'. 'Patients with concern' referred to those patients where the physician felt there were inconsistencies in the case, or the patient was a risk. 'Active patients' was related to the term of 'Active' which is a part of the Active/Stable/Watcher constructs that EM physicians often use to categorize their patients, based on how serious the patient's condition might be. 'Active patients' was used to describe those patients transferred during the sign out that did not have a disposition and needed active management. The code 'All patients' referred to all the patients transferred to a given physician during sign out, regardless of their status. The participants expressed different behaviors in terms of the patients they decided to examine at the bedside immediately after sign out, based on their motivation to verify sign out information e.g.

> "*For everyone else that I am taking responsibility for, I go see them, just because there have been times when things are different than what you were told, and it's so much better to catch those earlier in your shift than later"*.

Only the residents stated they saw 'All patients' while some attendings saw only 'Patients with concern' and other saw 'Active patients' e.g.

> *"I very rarely go to see the patients that has been signed out tonight. I go to see new patients"*.

These differences by role (see figure 6) could be due confidence in accurately assessing a case that comes from experience or due to the different tasks and responsibilities of residents and attendings e.g.

"*I would go to the bedside… you know, be prescriptive about what I want the resident to check when they go in to speak to the patient."*

An important theme of this category was that of 'Improvements', which captured the thoughts and views of the physicians regarding some of the challenges experienced in the ED and ways in which these issues might be addressed. Many of the challenges that the physicians expressed were around the sudden unexpected surges in patient volumes around the time of shift changes. This made it difficult to gather enough information on newly received ED patients e.g.

"if you get six new patients during sign out on each side, then that one resident can't. There's no way they can handle that. And so, in that case, sign-out becomes flexible".

The other big challenge in the ED was to do with interruptions that they experienced during sign out e.g.

"I can say I cannot concentrate on what they are saying, and what is being transferred here because of all of these interruptions".

These interruptions may occur more frequently due to the way sign outs are typically conducted e.g.

"our sign-ups occur in, in the center of the patient care area, with no barriers between us and anyone else. And so, you have free access from patients coming up to the counter, from nursing staff, from consultants, and you are out in the open. And so, that can lead to more interruptions".

Consequently, some suggestions for improvement were focused on ways to address these interruptions e.g.

> *"so, it might be nice to have like a dedicated space that have some sound insulation. And so that people were very cognizant of the fact that if they're in the sign out, the team is in that room."* And
>
> *"Maybe some sound barrier or something. And truly not have an interrupted space unless there was some legitimate emergency."*

The other main type of improvement that was suggested was around changes to the sign out processes to incorporate more patient facing sign out practices, e.g.

> *"the one system thing that I would change with our sign out, is that we don't actually do it in front of the patient, So, we don't actually incorporate the human that we are signing out the care for."*

Some of the physicians had previously mentioned that one approach to verifying the information they received during sign out might be to go and see the patient. The suggestion of conducting sign outs while rounding by the patient bedside would potentially allow the checking of details and gathering more information directly with the patient. Data gathering and verification were themes identified as important to sign out within this interview study and were also specified as major contributing factors in the MIRS dataset study. While conducting sign outs at the bedside may bring many advantages, it would also require significant changes to processes and protocols in the ED e.g.

"*So, those are kind of advantages, I think, of walking around and actually looking at patients as you do sign out. I suggested that when I came here, and it wasn't, it was not, um, excepted as we would change our practice".*

Incorporating this approach would also impact the communication and training guidelines particularly in terms of what would be appropriate to discuss in front patients e.g.

"when you actually look at the patient. And it's also just, it kind of inspires people to be more professional because they are in front of the patients".

Would these new methods impact the effectiveness of the sign out and ensure that all the information needs of the physicians on both sides were being met? It is clear that to implement such changes would require further investigation to understand their full impact on the information transfer processes, clinical reasoning processes and workload of ED physician teams. (Hern et al., 2016)

5.2.9 Anchoring Bias

The final main theme is that of 'Anchoring Bias', which is related to the factors that the clinicians felt led to reaching the wrong diagnosis based on anchoring on the wrong information received during the sign out. The participants recognized that there did exist a risk for anchoring on wrong information presented during a sign out, which could lead to erroneous decision making e.g.

"is there a potential chance of getting an anchoring bias from say like a differential diagnosis that you've been given during the signup. I think that there definitely is".

They expressed an awareness of the effect that anchoring early on in the diagnostic process can have e.g.

"so, it's easy to sort of pigeonhole that patient into a non-acute diagnosis".

In addition, they acknowledged the impact of not just the information they received but also how cues in the environment could influence their decision making e.g.

"can be anchoring off of your prior resident's diagnosis, it can be anchoring off the details of you're getting the complaint about, I can be anchoring off, you know, who that patient is".

A factor that could be influential for anchoring bias is the inherent ambiguity in many cases e.g.

"anyone who complains of chest pain and shortness of breath has a fairly broad differential".

While participants had cited a lack of familiarity with the patient case by the sign out provider as a flag that triggered verification of the information, it seems that the non-verbal cues can also be influential for anchoring also. The absence of reasons to doubt the validity of information provided might cause the decision maker to give undue weight to incorrect elements, e.g.

"someone presents a case in a very confident way, and you think all of the pertinent data is available and you walk in and you get some other piece of information from the patient."

It is not just the information or the person providing the information that is considered to have an influence. Patient qualities may also play a part in forming a mindset or view that has an undue influence on how the sign out information is considered e.g.

"anchoring bias is there for every patient, for every individual encounter, and then again, for those frequent flyers in particular".

Familiarity with repeat patients may cause physicians to anchor on the previous history of the patient rather that assessing the current sign out information with an unbiased perspective. The clinicians in this study appeared to be aware of the need to mitigate for this and to remain vigilant of their own assumptions e.g.

"And then if there's a difference in what's happening today because the hard thing about frequent flyers is that they often come in for things that do not require urgent, urgent medical attention. But occasionally they do."

This supports the notions that in addition to cognitive forcing strategies like checklists, physicians' own metacognition can be a powerful di-biasing solution to addressing medical errors caused by cognitive biases like anchoring bias. (Graber et al., 2012)

5.3 Limitations

A convenience sample of emergency medicine physicians was used and the criteria to be included were that the participant be a qualified doctor and have a minimum of one year's emergency medicine experience. While the physician participants were very generous with their time and forthcoming with their responses, there were a number of limitations with this study.

The small sample size and the fact that all the participants were from one hospital system means that the findings may not be reflective of the general EM physician population and may not be generalizable to other hospitals or to non-academic institutions. However, as a combination of attending physicians and residents were interviewed, the results can be

considered to be representative of those both roles for academic emergency medicine departments.

The method used in this study was a thematic review of the narrative data from interviews that were conducted with an interview guide. The subjective nature of this type of qualitative analysis that relies on the researcher to select the data for inclusion for analysis, is always vulnerable to selection biases. Similarly, the questions drawn up in the interview guide would have the effect of steering the responses from the participants. Hence the nature of the questions themselves can influence the data gathered. To mitigate for these issues the interview guide was reviewed with subject matter experts and the questions were open ended to allow participants to freely express their opinions and thoughts based on their own experiences. A combination of inductive and deductive coding was performed to ensure that the views of the participants were reflected in the data.

As the thematic coding was conducted by a single researcher, inter-rater reliability cannot be provided. However, the IRB has been amended to enable a second researcher to review the transcripts and code the data. This analysis by a secondary reviewer will be conducted shortly after the completion of this dissertation work. In the meantime, the results of this study were sufficient to be used to inform the development and design of a quantitative method to study the effects of risk factors identified on clinical decision making. In addition, it may inform further research into ED sign out protocols and the introduction of de-biasing awareness into medical education and training.

5.4 Conclusions

There has been much work done to understand the processes and structure of handoffs both in critical care and emergency environments. (Abraham, Kannampallil, Patel, et al.,

2016; T. G. Kannampallil et al., 2016; Mamykina et al., 2016) However, there is little published on the way the information transferred during handoffs in the ED impacts the clinical reasoning and diagnostic decision making of clinicians providing and receiving those sign outs. The purpose of this phase of work was to better understand the nature of clinician decision making within the context of emergency department sign outs in order to obtain factors that could be incorporated into an experimental study to empirically measure their effect on diagnostic decision making.

The approach taken in this first phase of this dissertation work was to identify factors that within the context of the environment and activities of the management of patient cases in the ED. This involved review of known cases of medical error from the MIRS database. The results yielded a set of high-level categories or features that could be considered potential factors that may contribute to erroneous decision making in the ED. These categories included the chief complaint in the case with cardiovascular and abdominal cases being most common followed by head and trauma cases. Also, cases that were atypical in presentation or involved complex medical histories were most commonly associated with errors.

These findings were then complemented by findings from an interview study investigating how a clinician might consider and use information transferred during sign out in their clinical decision making. This aim of the interview study was to understand, from clinicians' own perspective, how they consumed the information, their preferences in terms of the structure and content of the information, what affected their trust in the information provided and the actions they did related to their processing of the information to support their decision making. The findings yielded a list of features and conditions that were potential candidates for anchoring risk factors.

5.5 Recommendations from Specific Aim 1

The findings from the Phase 1 studies were translated into design elements for the experiment to simulate sign out communications for the second phase of this of this research project. The findings included when considering the design of patient cases for the experiment were:

- Patient Story

This involved providing the sign out with the key information presented with the context of patient. This included pertinent elements of the patient's history of presenting illness, their medical history and key details about their workup such as relevant lab and test results.

- Chief Complaint

Participants stated in the interview study and the analysis of the data from the MIRS database indicated that type of complaint the patient comes with to the ED is important in that it sets the path for the diagnostic process. As such, the chief complaint should be given important consideration in the design of any patient cases for the experiment.

- Complexity of the Case

Both the analysis of the MIRS dataset and the physician interview study identified that cases with complex co-morbidities pose a challenge for EM physicians in terms of diagnosis. The cases in the experiment should include cases with a representative array of complexity.

- Concise Versus Complete Nature of Sign Out

Sign outs are typically conducted under time pressure to minimize the time physicians are away from patient care. As such the information transferred is presented as a short

packet of patient information from which the recipient has to understand the situation sufficiently enough to decide how to proceed with managing the patient. Consequently, the information in a sign out must strike a balance between being concise and completeness. The experiment design should ensure this is met and that the language EM physicians use is reflected in the cases.

- Volume of Information Presented about the Patient Case

The amount of information presented in the sign out is important and will vary depending on the complexity of the case as well as the stage of care that the patient is in. In addition, how explicitly the information is conveyed, will affect the volume of information. The study design should consider the volume of information included in the cases, in terms of stage in the diagnostic process and the explicitness of how the information is conveyed. The volume of information and the details included has signal to noise implications. (Cheung et al., 2010)

- Structured Sign Out

Disorganized handoffs have been identified as a risk factor for errors in the ED. Studies indicate that in many institutions clinicians resort to their own practices resulting in variation in the structure of the sign out information. (Cheung et al., 2010) The physicians interviewed indicated that the structure of the sign out does affect their perceptions about the information presented.

All these aspects were considered when selecting features that were incorporated into the development of fictitious patient vignettes to simulate the transfer of cases within a sign out scenario. The experiment was intended to test the effect of the factors built into patient cases to determine their effect on the decision making of participants receiving the

information in the experimental setting. These factors served as design considerations of the experimental study for the second phase of this project, Specific Aim 2.

Chapter 6: ED Sign Out Study Methods

6.1 Introduction

This emergency department survey study (EDSO) was an experimental study designed to evaluate features identified as potential risk factors for anchoring bias that might influence diagnostic decision making in the ED. The experiment was conducted via a survey platform that contained patient cases presented in the format resembling the narrative style of information communicated during a sign out in the ED. The patient cases were designed with features identified as potential risks factors for increasing the tendency to anchor on incorrect or inappropriate information during the clinical reasoning process.

6.2 Setting and Subjects

This experiment was conducted with emergency medicine trained clinicians practicing in the emergency departments of two academic institutions across three hospitals. The first site was McGovern Medical School, which included two hospitals, Memorial Hermann Hospital and LBJ General Hospital. The second site was University of Maryland Baltimore (UMB) Department of Emergency Medicine. While both are academic teaching facilities, the main difference between the two emergency medicine departments were that UMB Emergency Departments require the use of a handoff protocol for their ED sign outs, while UTH McGovern Medical School does not formally use handoff protocols.

The participants of this study were any emergency medicine clinicians with a minimum of one year's experience in emergency medicine in their respective departments at the

time of completing the study. This included attending physicians, resident physicians, fellows, advance practice providers (APPs), which included nurse practitioners (NPs) and physician's assistants (Pas). These study participants all had experience of providing and receiving sign outs and in diagnostic decision making for ED patients, which were qualities necessary to perform the experimental study. Medical students were not included in the study because, while they do sometimes participate in patient care and even sign outs, their level of knowledge and experience may not have been sufficient to complete the experiment.

Participant recruitment was conducted at the start of the study via an email which included a flier, sent by the EM physician SMEs in the project for both sites. Another reminder was sent three weeks into the study at both sites. A further request was made in person at the McGovern Department of Emergency Medicine weekly resident didactic meeting and a final email reminder was sent the week the study was set to close. The participants were able to complete the study in their own time via a link sent within the emails.

6.3 Protection for Human Subjects

As this study involved the participation of clinicians an Institutional Review Board (IRB) application was submitted by the researcher via the UTHealth Integrated Research Information Software (iRIS) system, accompanied by the appropriate the supporting IRB formatted study protocol. The study protocol listed Roni Matin, MSc as the principal investigator, Amy Franklin, PhD, Amit Mehta, MD, CMQ**,** Brent King, MD, MMM, FACEP, FAAP, FAAEM and Robert Murphy, MD as co-investigators. The study IRB details are as below:

Study Title: Simulated Sign Out Experiment with Emergency Medicine Physicians to Determine the Influence of Communication Factors on Clinical Decision Making.

UTH IRB Protocol Number: HSC-SBMI-20-0203

The study was approved as exempt status and fully approved by the IRB board on October 29th, 2021.

An additional IRB application was submitted by Dr Brent King for the UMB IRB Board.

UMB IRB Protocol Number: HP-00098665

The study was fully approved by the IRB Board on November 15th, 2021.

6.4 Study Design and Experiment Constructs

The experiment to test the potential risk factors identified was designed to replicate emergency medicine sign out, which involves the signing out provider transferring a set of patient information in a narrative format to the oncoming clinician at the time of shift change in the ED. The narrative handoff consists of patient information summarized to present key information and details. Sign outs are intended to provide the receiving clinician with an understanding of the patient's medical condition and status at the time of sign out. The experiment contained a combination of control cases and stimuli cases. The control cases were designed to reflect the typical cases that come to the ED. Similarly, the stimuli cases would also be reflective of ED cases, but they would contain aspects of the test conditions inserted into them.

The target participants of the experiment were trained emergency medicine clinicians. Participants were required to review the fictitious patient cases and provide their assessment and answers to some question about the cases. The mixture of control and stimuli cases created was reflective of the type and frequency of patient cases that typically present in the ED. An iterative approach was used for the design process with input to the design from multiple sources including a group of subject matter experts

(SMEs), who were experienced practicing emergency medicine physicians affiliated with McGovern Medical School. The survey was anticipated to require 30-60 minutes of the participants' time and participants were offered a $30 Amazon gift card for participating.

6.4.1 Control cases

The term control case refers to the fictitious patient case vignettes that were created to present the types of standard routine patients typically seen the ED. The information for the cases was selected from the medical literature and the cases were set up as documented in the source literature. The clinical details, patient course and outcomes remained the same as described in the source literature. The control case set did not vary in clinical information, details and outcomes. The control cases represent the control measures in the experiment providing the baseline against which the effect of cases containing test conditions can be compared in the data analysis. All the control cases in the experiment were presented to all participants.

6.4.2 Stimuli Cases

The term stimuli case refers to the fictitious patient case vignettes that were created to present the test conditions. The patient cases were selected from medical literature as with the control cases and then enhanced to include the stimuli conditions. The details of the stimuli cases contained the clinical details and outcomes taken from the source material overall but then also contained the risk factors that had been identified previously from Phase 1 of this research. These risk factors identified were known as the test conditions of the experiment. There test conditions that were built into the stimuli cases. Each participant received a randomized permutation of the test conditions from the set of test condition permutations.

6.5 Within-Subjects Design

As this study required emergency medicine trained professionals as participants, the sample population and time for participant recruitment was anticipated to be limited. Hence an across participant study where participants are randomly assigned to different control and intervention study arms would be not be possible. In addition, the various case vignettes may contain different levels of cognitive load thus affecting the successful completion of related tasks. Therefore, careful balancing across participants would be necessary as part of the study design

Consequently, a within-subjects experimental design was selected. This approach requires that every participant is presented with the control and the stimuli cases. This approach enables collection of increased data points from the testing of both the control cases and stimuli cases simultaneously and thereby increases the statistical power of the data collected. The aggregated data from participants' responses was analyzed to compare the responses for control cases against the stimuli cases.

6.5.1 Design Considerations

The iterative design process involved taking forward the findings derived from the Phase 1 MIRS and EM Physician Interview studies and complementing them with advice and suggestions from the SMEs.

6.5.2 Features Important to Design of the Stimuli Cases.

The review of Aim 1 and discussion with SMEs produced several features for consideration in the design of the experiment as well as the identification of potential risk factors as test conditions for the experiment.

i) Representativeness

Care was taken to ensure the cases were not manipulated to be overly atypical in presentation or have deliberately misleading cases built in. There was no intent to include intentionally misleading cases or to have very rare medical conditions for participants in the experiment. Indeed, following on from the findings of the MIRS data, a set of patient cases that were representative of the types of patients that come to the ED in terms of age, gender, medical conditions and other population factors was developed. In addition, the patient cases developed were iteratively discussed with the SMEs to ensure all the details of the cases were appropriate.

ii) Number of Cases

Most sign outs involve transfers of around 20-30 patients in a sign out but can be as high as 40-50 for a busy general hospital. Of the 20 or so patients transferred, an on-coming clinician may take on between eight to fifteen patients under their care. The total cases number for the experiment therefore was chosen to be eleven which was comprised of seven control cases and four stimuli cases.

iii) Complexity

Both the previous studies highlighted that cases that are complex or have complex co-morbidities can be difficult to diagnose or be associated with error. Equally, many patients present at the ED with minor or straightforward complaints, which are dealt with

expediently. Care was taken to include a variety of cases with different levels of complexity to represent the real-world scenarios in the ED.

iv) Patient Story

The factor from the EM Physician interviews that was specified by all the participants as a preference in the way the sign out information was presented was the patient story. They also wanted to be provided patient details. The EDSO patient case vignettes were created to include a brief patient story giving context to the events leading up to the fictitious patient's presentation in the ED, including some patient specific information. It is important to note that this story may be influential in setting a path of thinking for the recipient, against which evidential information such as diagnostic lab and imaging results are weighed up to either support or refute the initial impression.

v) Concise Versus Complete

Sign outs are very limited in time and so the provider must balance including important pertinent information with brevity. To achieve this EM physicians typically communicate their sign out using a coded like language that include abbreviations and commonly used phrases e.g., BMP for when referring to the basic metabolic panel of blood test. The case vignette narratives were developed in the style of the succinct coded language that is typically used by clinicians during sign out.

vi) Volume of Information

There were a number of comments and themes that pointed to the volume of information that is presented at sign out as being important. This was considered in the context of two features of sign out information; the stage in the process that the patient is in at the

time of sign out and whether the information included during the sign out was explicitly conveyed and explained.

vii) Chief Complaint

From the interview findings and throughout many conversations with the EM SMEs, the importance of the chief complaint the patient is presenting with in the ED, was stressed repeatedly. As such, the chief complaint was considered an important factor to incorporate in the design of the experiment. The cases in the experiment were developed to reflect not only the frequency of chief complaints in the ED but also some of the complexity and ambiguity that are associated to the different types of chief complaint.

6.5.3 Features Not Considered for Test Conditions

The findings of Aim 1 studies did identify that the content and structure of handoff information were important to sign outs and clinicians' preferences. After much consideration and discussion with SMEs, it was decided not to use the theme of structure as a category of risk factors for the test conditions.

i) Structured and Disorganized Sign Outs

Disorganized handoffs were mentioned by the physicians as an important factor for how credible they felt the information being transferred was. This suggested that disorganized information presentation during handoffs may possibly lead to decreased anchoring on information presented in that sign out. However, in the interview study physicians stated that disorganized sign out information prompted a desire to verify the information with further information gathering in the form of asking questions or checking the EHR or directly with the patient. The experimental design did not allow for these types of interactions and so for this reason this feature was not taken forward into the design as a

possible test condition. Inadvertently introducing disorganized information in the experimental cases might introduce a potential confounding effect. As such care was taken to ensure the sign out narratives were consistently structured for content and style.

ii) Order of Sign Out Information

Typically sign outs are conducted in bed order although some of the physicians in the interview had suggested that prioritizing the sign out in terms of the sickest patients first would be their preference. In addition to changing patient order, the order of the information in the sign out could be varied e.g. for workup have imaging results first or the order in which the patient's co-morbidities are listed in the medical history. (Meyer et al., 2013) Varying the order of information both within a case and as well as across all cases of a sign out would not be achieved easily within the constraints of an experimental design. Trying to do so would overly complicate the design of any experiment testing these conditions. While it was recognized that the Primacy/Recency Effect of information may play a role in contributing to anchoring bias, the order of information was not taken forward as a test condition for the purposes of this study.

6.5.4 Summary: Risk Factors for Anchoring – Test Conditions

In conclusion the key components that were selected to the be part of the test conditions in the experiment and thus become the test conditions in the stimuli patient vignettes were related to the volume of information. The volume of information in a case can be affected by the stage of the diagnostic process that case is at and by the level of details that are mentioned in an explicit way. Hence the two test conditions were:

Stage in the process: this concept pertains to the diagnostic process and the information related to this rather than the patient course in the ED, which is more akin the care for

that patient. This concept is focused more on the decision making process and the associated information needs of the EM clinician and not on the quality and outcomes of healthcare delivery to the patient.

Explicitness versus Implicitness: this construct is based on the concept of how explicitly information is unpacked and conveyed during the transfer of information in handoffs. Implicit communication was assumed to be associated with less volume of information and explicit was assumed to have more.

Chief Complaint: A special consideration was made for chief complaint, which was cited on so many occasions as a pivotal factor to the decision making process. While the chief complaint itself was not a test condition, it was considered to have the potential to heavily influence the responses of participants to the patient case vignettes constructed for the experiment. As such the design involved blocking for chief complaint.

So, in summary the test conditions for the experimental model were the stage of the case and explicitness of the information. The cases in the study also included blocking design for chief complaint to enable accounting for its influence on decision making.

6.6 Definitions of Test Condition Constructs

6.6.1 Stages in the Diagnostic Process.

Stage in the process refers to the stage in the physicians' diagnostic process rather than the patient course through the emergency department. While the stage within the workup that the patient is in at the time of sign out does have some bearing on the stage of said patient's course in the emergency department, for the purpose of this study the stage in the process is solely related to the diagnostic reasoning process conducted by emergency physicians. (Cheung et al 2010; Gibson et al 2010)

The different stages refer to whether a case is in the middle of the diagnostic process or at a later stage in the diagnostic process and relates to the amount of data about the case. In other words, the number of pieces of data that have come back and the impact that may have on the ambiguity or certainty about the case and its diagnosis.

i) **Early**

Early cases are ones where the initial assessment of the patient has been conducted and may include a primary and secondary examination and taking the patient's medical history

- Low confidence in the diagnosis – diagnosis not clear- requiring more tests
- Patient may present with a specific complaint, but differential diagnosis is broad
- Information gathering stage – more information is necessary to create a management plan
- Some critical tests and information may be back – vitals and early labs may be back, but more is required to develop the differential.

ii) **Middle**

This is during the intervention stage where some tests have taken place or procedures have been started or some medication provided

- Diagnosis is uncertain but requires more confirmatory evidence for management
- Confidence in the diagnosis may be high pending pivotal confirmatory evidence from test/procedures/consultants
- History and Physical suggestive of a diagnosis and some test results may be back but still awaiting further evidence
- management may be in process – it is often the case that in the ED patient be started on treatment even though all the evidence is not back yet
- May have a presumptive diagnosis but that may change.

iii) **Late**

Late cases are ones in the reassessment phase where some intervention has occurred, and the patient is being reviewed to assess whether the treatment has succeeded, or further tests or interventions are required

- Diagnosis highly likely –pending test/procedure/consultant to close the case
- Confidence in clinical impression is high
- Definitive diagnosis may still be unclear
- Sufficient tests/exams are complete to exclude serious pathology and patient is stable for disposition

iv) **Book End**

These are cases where the disposition has been agreed or the patient has just arrived, and no workup has been started.

- Either the case is so close to disposition that no further diagnostic decision making is required
- Or the case is so early that no information useful for decision making has been gathered

The control cases did not contain test conditions. The various control cases were set to contain a sign out that occurred at different stages in the diagnostic process for the patient. The stages were either early, middle, late or end cases.

6.6.2 Test Condition Constructs for Stages in the Diagnostic Process

If the diagnostic process were considered in the form of a linear representation, then cases that were later in the process could be considered as longer cases and cases that were earlier in the process could be considered as shorter cases.

Following this conceptual frame, the test conditions to represent the stage in the diagnostic process were expressed as a construct of length. So, for the purpose of developing the stimuli case vignettes late cases were labelled as long cases and middle stage cases were labelled short cases. The terms late and middle corresponded to the definitions stated above.

In other words, the term ‘long’ was a proxy for a case where sign out occurred at a late stage in the diagnostic process. Correspondingly, the term ‘short’ represented a case where the sign out is approximately in the early to middle stage of the diagnostic process.

This concept was carried through to the design of the patient case vignettes by assigning the theoretical sign out at the point in the patient case vignette to reflect the length of the case and thereby the stage of the diagnostic process.

Long = Late = sign out assigned at a late stage in the diagnostic process

Short = Middle = sign out assigned at the middle stage of the diagnostic process.

6.6.3 Definition of Explicitness and Implicitness

Explicitness referred to the instances where pertinent data and its potential meaning is presented in packaged up or in an unpacked way. Explicit information is presented in an unpacked manner and implicit information is presented in a packed-up form. These constructs are based on the Unpacking Principle that is part of Support Theory proposed by Tversky and Koehler (1994) who suggest that the more explicitly information is conveyed is likely to influence the weight the information is given during decision making.

Implicitness/Explicitness of SO information- refers to the instances where pertinent data and its potential meaning are presented in packed up or unpacked way. The resultant SO information may contain more (unpacked – explicit) information or less (packed – implicit) information. This feature is based on the unpacking principle, which is explained by Support Theory

Unpacking Principle and Support Theory

Support Theory is a descriptive theory, proposed by Tversky and Koehler (1994), which suggests that the unpacking principle may influence anchoring bias. The Unpacking Principle suggests that the more unpacked information about an event is, i.e. the provision of a more detailed description, the more likely it is to increase the judged probability of the event by the recipient. (Chapman & Elstein, 2000; Tversky & Koehler, 1994)

In other words, the more unpacked format is for the information provided about an event, results in an increased number of items of detail within the information about that event. This in turn influences the recipient of the information to consider that there is a higher

likelihood of the event. (Sperber & Wilson, 1986.) So, the more explicitly information is conveyed enhances the perception of its likelihood of being reliable and thereby relevant to the decision being made. This principle suggests the explicitness or implicitness of the information may contribute to its influence in anchoring on said information during decision making.

6.6.4 Test Condition Constructs for Explicitness

The two instances of the construct of explicitness were the test conditions of Explicit and Implicit.

Explicit

- An assumption communicated by an utterance of its logical form – details are unpacked
- Lends itself to natural comparative interpretation
- Example: "We have ordered a full panel workup including BMP, CBC, Coags including D-dimer. Most of them are back but not the D-dimer."

Implicit

- Not explicitly stated – not unpacked into components or details
- Lends itself to inference to a plausible explanation
- Example: "We have ordered a full panel workup. Most of them are back."

Design and Development of Study Materials

Reference Materials.

To begin the process of developing patient case vignettes a selection of potential candidate cases were shortlisted from the medical literature. The two main texts used for reference were:

- Levis, J. T., & Garmel, G. M. (2009). Clinical emergency medicine casebook. Cambridge University Press.
- Okuda, Y., & Nelson, B. P. (2009). Emergency medicine oral board review illustrated. Cambridge University Press.

After an initial review for content and complexity within the cases of the shortlist, ten cases from either the Clinical emergency medicine casebook (CEMC) or the Emergency medicine oral board review illustrated (EMOBRI) were selected. Seven of the cases selected were the templates for The seven control patient vignettes and three were selected for the stimuli conditions. A further case was for a stimuli case vignette was provided by one of the EM Physicians and was based on real patient case from their own experience.

6.7 Patient Case Vignette Set.

The cases were chosen following the design criteria, in that they were a selection of cases that were relatively balanced for gender, had a representative range of ages as well as having diverse and frequent chief complaints with differing levels of complexity at diverse stages in the process.

Table 12

EDSO Study Summary of Patient Vignettes

Name	Control/ Stimuli	Chief Complaint	Stage	Gender	Age	Source
Epsilon	Control #1	Lung – Pneumothorax	Early	Male	47	CEMC – case24
Mu	Control #2	Abdomen - Appendicitis	Early	Male	16	EMOBRI – case 37
Zeta	Control #3	MVC – Fractured ulna	Late	Female	32	CEMC – case 67
Beta	Control #4	Head – Stroke	Middle	Male	66	EMOBRI – case 78
Kappa	Control #5	Abdomen – DKA + UTI	Middle	Female	53	EMOBRI – case 42
Gamma	Control #6	AMS – hyponatremia	Late	Male	57	EMBORI – case 11
Lambda	Control #7	Trauma – Jones Fracture	End	Male	14	CEMC – case 65
Alpha	Stimuli – Head	AMS – UTI + skin ulcer	Randomized	Female	84	EMBORI – case 38
Sigma	Stimuli – Lung	Pulmonary Embolism	Randomized	Female	45	EMBORI – case 35
Omega	Stimuli – Abdomen	Upper GI bleed	Randomized	Male	59	SME – EM Physician
Delta	Stimuli – Cardiac	Aortic Dissection	Randomized	Male	73	CEMC – case 18

Note: case Mu is a case of a motor vehicle crash (MVC), case Kappa is a case of diabetic ketoacidosis (DKA, case Gamma is a case of Altered Metal Status (AMS) due to hyponatremia, which is very low levels of sodium.

6.7.1 Control Vignettes:

1. **Case Epsilon**: was a case of a tall thin 47-year-old man who came in with shortness of breath after a severe bout of coughing. He is a smoker with a history of COPD, but his right lung showed an absence of breath sounds. His diagnosis was a pneumothorax.
2. **Case Mu**: this was a case of a 16-year-old male who presented with nausea and severe abdominal pain. He was first thought to have a viral syndrome when he saw his primary care physician and his urinalysis was negative, but his symptoms got worse. His exam showed right lower quadrant tenderness. His diagnosis was acute appendicitis.
3. **Case Zeta**: this was a case of a 32-year-old woman, who was involved in a motor vehicle crash with another car, when she was forced to drop her motorbike. She had a helmet on and was alert when she arrived, but she had a laceration and fracture of her left arm. Her diagnosis was fracture of left ulna
4. **Case Beta**: this was a case of a 66-year-old man brought in by EMS called by his wife, as he was unable to get out of bed when woke after a nap. He showed right side paralysis and right sided facial droop. He currently smokes and drinks alcohol in a moderate amount; chest x-ray and labs were normal. His diagnosis was stroke.
5. **Case Kappa**: a 53-year-old woman came in with diffuse abdominal pain and vomiting. She has type 1 diabetes and does not smoke or drink. Her fingerstick was 435, her anion gap was 29, her EKG shows tachycardia, and her urinalysis is positive for infection. Her diagnosis was diabetic keto acidosis due to a urinary tract infection (UTI)
6. **Case Gamma**: this was 57-year-old man who was brought in by police concerned for his altered mental status and saw him shaking violently. He shows no trauma and chest x-ray

and EKG were normal but toxicology showed alcohol intoxication and sodium of 110. His diagnosis was severe hyponatremia.

7. **Case Lambda**: this was a 14-year-old male who had an inversion injury during soccer and was unable to bear weight. He was otherwise healthy with no significant medical history. His foot showed a fracture on x-ray, and he was awaiting a fracture boot. His diagnosis was a Jones fracture. This was an end case where no further diagnostic decision making was required, and the disposition was set.

6.7.2 Stimuli Vignettes: -

1. **Case Alpha**: this was a case of an 84-year-old woman who was disoriented and forgetful with a slight fever. She did have a fall 3 days prior but had no loss of consciousness. She showed no focal neuro deficits. Her EKG, chest x-ray and labs were all normal, but her urinalysis was only weakly positive for infection. This diagnosis was UTI + an undetected decubitus skin ulcer
2. **Case Sigma**: this was a 45-year-old woman who presented with chest pain and shortness of breath after she had an argument with her son. No recent trauma but did have ankle surgery two weeks prior. Chest x-ray and ultrasound were clear, but EKG showed tachycardia. Labs were normal except for an elevated D-Dimer. The diagnosis for this case was a pulmonary embolism
3. **Case Omega**: this was a case of a 59-year-old man who came in with shortness of breath which his primary care physician suspected was “holiday heart syndrome”, as his symptoms developed after going on a fishing trip with his friends, during which he consumed six beers a day. His labs from two weeks prior were normal with hemoglobin of 14. His EKG showed tachycardia and dysrhythmia, but chest x-ray was normal. His

latest labs showed a normal troponin, normal white blood cell count and hemoglobin of 9.2. His diagnosis was an upper gastrointestinal bleed.

4. **Case Delta**: this case was of a 73-year-old man who came in complaining of sudden headache, neck and chest pain and severe pain radiating down his back. He had a history of hypertension, and coronary artery disease which is controlled with medication. He was alert with no focal neuro deficits, but his pulse was weak, and he had a diastolic murmur when auscultating his heart. His EKG showed tachycardia but no ST elevation. His labs were all normal, but his chest x-ray showed a widened mediastinum. His diagnosis was an aortic dissection.

These descriptions of the cases are brief summaries. The stimuli cases had varying amounts of information based on the test condition permutations. The full case details including the other materials produced for developing the case vignettes and case sign out narratives are included in the appendices.

6.7.3 Within-Subjects Blocking

The two test conditions of length and explicitness were considered to be not truly independent. Long cases may have more information and cases that are explicitly conveyed may also have more information. To determine which condition contributed to any effect observed within an experiment, the design would have to be blocked for potential interaction effects of the two test conditions.

6.7.4 Blocking for Interaction Effect of Stimuli Conditions

The two test conditions each have two instances of the condition i.e., long or short and explicit or implicit. In order to block for any potential interaction effect, the stimuli cases should evenly present all the permutations of the combination of these two test

conditions. Hence the full set of permutations resulted in four test conditions as illustrated in the 2x2 table 12.

Table 13

Blocked Design for Non-independent Test Conditions

TEST CONDITIONS	**Short**	**Long**
Implicit	Short & Implicit	Long & Implicit
Explicit	Short & Explicit	Long & Explicit

The 2x2 design of test conditions resulted in four counter blocked test conditions: long/explicit, long/implicit, short/explicit, short/implicit, which in their abbreviated form would be S/I, S/E, L/I and L/E respectively.

6.7.5 Blocking for Potential Confounding Effect of Chief Complaint

The control cases were developed to provide a set of standard cases against which the diagnostic decision making results of the stimuli test condition cases could be compared. However, the chief complaint is recognized as having the potential to significantly influence how a case diagnosis is considered. This suggests the chief complaint had the potential to confound the effect of study measures in the patient case vignette experiment. To control for the potential confounding effect of the chief complaint, the experimental design was blocked for the most common chief complaints. The chief complaints were specified to be cases related to the head, lung, abdominal and cardiac conditions by the emergency medicine physician SMEs. This knowledge was based on observing the patient population presenting in the hospital emergency departments over their working

experience. These chief complaints also correspond to the most common emergency department chief complaints reported in the literature (McCaig & Nawar, 2006) Hence four of the patient cases in the experiment would be stimuli cases to accommodate the even presentation of each of the four chief complaint types.

6.7.5 Test Condition Permutations

The design of the stimuli patient cases had been blocked for the test condition interaction effect resulting in the four permutations of the combination test conditions. There was also the need to block for the four main chief complaints.

In summary, there were 4 test conditions i.e., the permutations for length and explicitness:

- short/implicit
- short/explicit
- long/implicit
- long/explicit

Then there is blocking for the 4 chief complaints: Head, Lung, Abdomen, Cardiac. Hence, the total number of blocked test condition stimuli cases was a 4x4 design of 16 permutations.

Table14

Test Condition Permutations for Length, Explicitness & Chief Complaint

CHIEF COMPLAINT	Short Implicit	Short Explicit	Long Implicit	Long Explicit
Head	H-S/I	H-S/E	H-L/I	H-L/E
Lung	L-S/I	L-S/E	L-L/I	L-L/E
Abdominal	A-S/I	A-S/E	A-L/I	A-L/E
Cardiac	C-S/I	C-S/E	C-L/I	C-L/E

The total number of permutations included in the stimuli test conditions were the sixteen instances listed above, that have been blocked for chief complaint and test conditions.

6.8 Patient Case Vignette Development Process

The development of the patient case vignettes that would be part of the experimental study involved a multi-step process. Having identified the cases that would be used for the controls and the stimuli cases, the clinical notes for each case were collated and supplemented from other sources such as SME suggestions. The information from the clinical notes were converted into case vignette diagrams, which were used to visually represent the point at which the sign out occurred. The diagrams facilitated discussion with SMEs while having the case details easily visible. This allowed discussion and agreement on matters like the placement of the sign out. Once this was agreed for a case the sign out narrative was developed to reflect all information pertinent to the case and design considerations mentioned earlier e.g., concise language and medical abbreviations. The process resembled this flow.

Patient Case Clinical Notes ➔ Case Vignette Diagram ➔Agree Sign Out Stage➔Sign Out Narrative

This same initial processes for building the patient case vignettes were followed for the development of each of the control and stimuli cases. In addition, the stimuli cases had further development to incorporate the various stages in the process i.e., long, and short cases and the explicitness or implicitness of the information.

The component materials produced for the purpose of developing the case vignettes for all the control and stimuli cases, including all the test conditions are described below. All

the component materials, e.g., clinical notes, vignette diagrams and sign out narratives are included in the appendices.

6.8.1 Patient Case Clinical Notes

A set of clinical notes for each patient case whether control or stimuli was created and included all the relevant clinical details such as test results and imaging (figures 8 and 9). The information was used to develop vignette chronology diagrams e.g., figure 10.

2. 16 y/o male with nausea, and abdominal pain (p123 CEMC - #37) – appendicitis

A 16 year old male presented to the ED complaining of abdominal pain, nausea.

B.Vital Signs

Temperature	38.2C
Pulse	90 beats/minute
Blood pressure	120/70 mmHg
Respiration	18 breaths/minute
Oxygen saturation	98% on room air

C. Physical Examination

GENERAL APPEARANCE: The patient was lying supine on the gurney, appeared uncomfortable due to pain

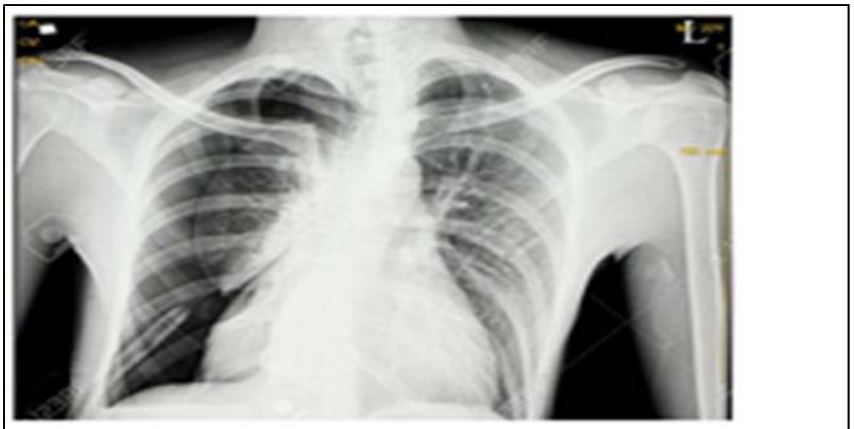
D. Primary survey

a. Airway: speaking in full sentences
b. Breathing: no apparent respiratory distress, no cyanosis, clear lungs
c. Circulation: RRR, pale and cool skin, radial pulses 2+, normal capillary refill

E. History

a. HPI: A 16-year old male brought in complaining of abdominal pain. The patient was seen one week prior to his ED visit by his primary care provider (PCP), at which time he described the previous complains as well as subjective fevers. At that time, the patient's temperature was 37.9C. He was noted to be well appearing and in no acute discomfort. The abdominal examination revealed suprapubic tenderness to palpitations without the presence of rebound or guarding, no costovertebral angle tenderness (CVAT) was noted and his genitourinary (GU) examination was normal. A urinalysis was negative for infection and the patient was diagnosed with a viral syndrome.
Three days prior to his ED presentation, the patient reported a temperature of 103F (39.4C) and severe suprapubic pain. The following day, the intensity of his pain diminished somewhat and his fever resolved. In the ED, the patient continued to complain of crampy abdominal pain at a level of 6 (on a scale 0 to 10), with associated dysuria, nausea and vomiting. He denied diarrhea, constipation or penile discharge, and was tolerating oral liquids.
b. PMHx: none
c. PSHx: none
d. Allergies: none
e. Meds: none
f. Social: lives with parents at home; denies alcohol, smoking, or drugs; not sexually active
g. FHx: not relevant
h. PMD: Dr Miller

Figure 7: Example 1 of Case Clinical Notes – control case 2: Appendicitis



c. Monitor: BP: 95/54 HR: 110 RR: 24 T: 37°C Sat: 96% on non-rebreather mask

H. Results

Complete blood count:	
WBC	9.9x10³/µL
Hct	39.1%
Plt	142x10³/µL
Basic metabolic panel:	
Na	134 mEq/L
K	4.3 mEq/L
Cl	99 mEq/L
CO_2	26 mEq/L
BUN	12 mEq/dL
Cr	1.1 mg/dL
Gluc	79mg/dL
Coagulation panel:	
PT	13sec
PTT	26sec
INR	1.0
Liver function panel:	
AAST	23U/L
ALT	26U/L
Alk Phos	42 U/L
T bili	1.0mg/dL
D bili	0.5mg/dL
Amylase	51 U/L

Figure 8: Example 2 of Case Clinical Notes – control case 1: Pneumothorax

6.8.2 Patient Case Diagrams and Stages in the Process

The details of the clinical notes were converted into diagrams using Microsoft Visio to be able to visualize the chronology of events and the patient course. This visual format facilitated discussion with SMEs about the timing of the sign out, with the benefit of seeing the information available at the various points during the course of the case. The diagram timelines were not to scaled to time but rather designed to show key information.

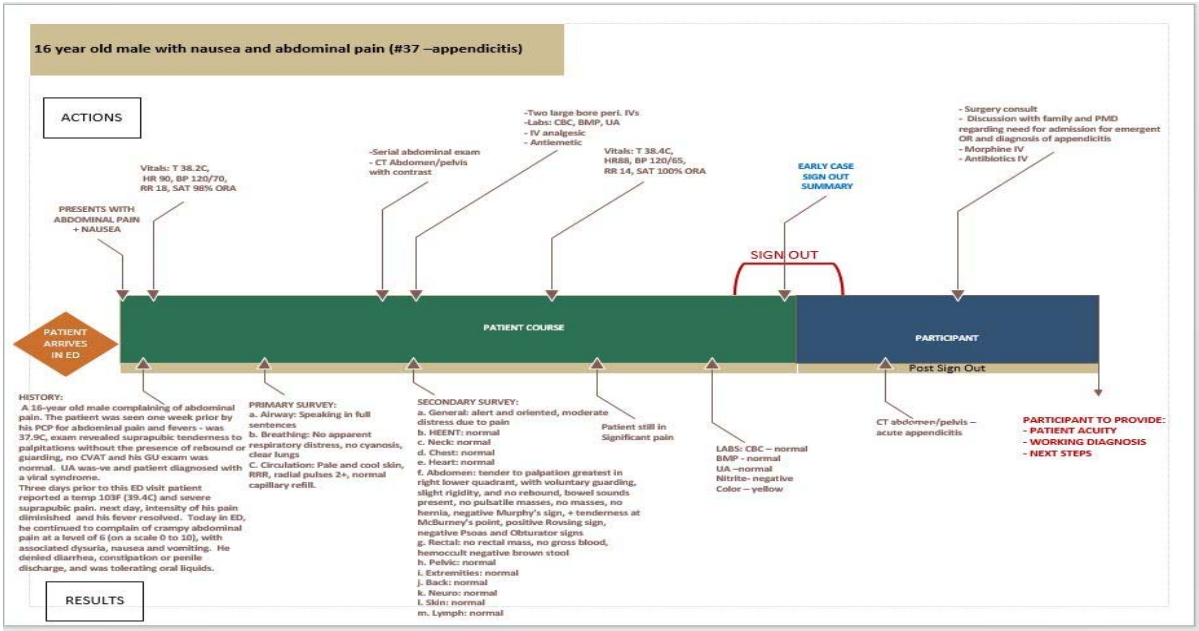


Figure 9: Example Case Vignette Diagram: Case 2– Control case of Appendicitis– Early stage

6.8.3 Development of Sign Out Narratives

The details from the case clinical notes were considered in conjunction with the case vignette diagrams to identify the key information that would be included in the sign out narrative for that case. The narrative would include the relevant information up to the point of the sign out. The linguistic style and content reflected the language that EM

clinicians use during sign out. This resulted in a sign out narrative for each of the cases e.g., fi, including one narrative for each permutation of the stimuli cases, i.e. four narratives for each chief complaint. All the narratives were reviewed with two SMEs iteratively to ensure consistency for language and clinical accuracy within the case summary sign out narrative.

Case 2: Appendicitis - EARLY

In room 7 we have a young 16-year old[1] male[2] who is here with nausea[3] and severe abdominal pain[4]. He previously presented to his PCP one week[5] earlier with fever[6], cramps[7] and pain[8] but the intensity[9] and fever diminished[10] the following day and with a negative urinalysis for infection[11] and he was diagnosed with a viral syndrome.[12] He is in the ER with nausea[13] and return of the suprapubic pain[14] and his temperature is raised[15]. He's received IV fluids,[16] pain meds[17] and an antiemetic.[18] Also, labs for CBC[19], BMP[20] and Urinalysis[21] have been ordered. His exam did show right lower quadrant tenderness[22] but no masses[22] or hernias detected[23] and his rectal exam was negative.[24] A CT of the abdomen and pelvis has been ordered.[25]

Figure 10 Example Sign Out Narrative: Case 2 Control case of Appendicitis – Early stage

6.8.3 Construct Validity: Measurement of Information Volume in Case Narratives

While the volume of information may be expected to increase with cases that are later in the diagnostic process, how explicitly or implicitly information is stated may also affect the volume of information. As a check for face validity of these constructs, a key data word count for clinically relevant words related to clinical decision making was conducted for each case narrative

Typically, more explicitly stated information could be considered to contain more information units and therefore would be associated with greater information volume. Thus, in terms of the test condition constructs an implicit case would have less information volume than explicit case of the same type.

Similarly, the stage of the case will affect the volume of information with cases later in the process predicted to be associated with a greater volume of information. Cases earlier in the diagnostic process would be associated with smaller volumes of information. So, this would translate for the test condition constructs. As for cases of the same type, a short case would have less information than a long case, which would have more information at the time of sign out.

i. Case Narrative Annotation for Key Data Word Counts

Direct comparison of key data word counts, as an indicator of information volume, cannot be made between different case types. This is because it does not consider other factors like the complexity of the case. However, comparing instances of the same case at different stages or described in more or less explicit terms may affect the key data words that are counted for each variation.

With this in mind, all the case narratives were annotated for key data word counts, including for each of the permutations of the test condition combinations in the stimuli cases. (see figure 12)

SIGN OUT NARRATIVES LUNG CASE

SHORT & IMPLICIT

In room 3, is a 45-year-old[1] female[2] symptomatic of chest pain[3] under the right breast[4] accompanied by dyspnea[5] for duration of one day.[6] She states the pain started in the afternoon[7] after she had an argument with her son earlier[8] and describes it as stabbing[9] and radiating to her back.[10] Past medical history is not significant.[11] No recent trauma[12] or fever[13], she had ankle surgery 2 weeks ago.[14] Her pain has resolved since arrival, so she thinks it was just stress from the argument. So, in her evaluation, her chest x-ray was clear[15] and EKG showed sinus tachycardia[16] and bedside echo was unremarkable[17] with a grossly normal EF.[18] Her pregnancy test was negative[19] but her other labs aren't back yet[20].

SHORT & EXPLICIT

In room 3, is a 45-year-old[1] female[2] symptomatic of chest pain[3] under the right breast[4] accompanied by dyspnea[5] for duration of one day.[6] She states the pain started in the afternoon[7] after she had an argument with her son earlier[8] and describes it as stabbing[9] and radiating to her back.[10] Past medical history is not significant.[11] No recent trauma[12] or fever,[13] she had ankle surgery 2 weeks ago.[14] Her pain has resolved since arrival, so she thinks it was just stress from the argument. So, in her evaluation, her chest x-ray was clear[15] and EKG showed sinus tachycardia[16] and bedside echo did not show any right heart strain[17] and showed a grossly normal EF.[18] Her pregnancy test was negative.[19] We ordered labs[21] for her including a D-dimer[22]but they are not back yet, so you need to follow them up.[23]

LONG & IMPLICIT

In room 3, is a 45-year-old[1] female[2] symptomatic of chest pain[3] under the right breast[4] accompanied by dyspnea[5] for duration of one day.[6] She states the pain started in the afternoon[7] after she had an argument with her son earlier[8] and describes it as stabbing[9] and radiating to her back.[10] Past medical history is not significant.[11] No recent trauma[12] or fever,[13] she had ankle surgery 2 weeks ago.[14] Her pain has resolved since arrival, so she thinks it was just stress from the argument. So, in her evaluation, her chest x-ray was clear[15] and EKG showed sinus tachycardia[16] and bedside echo was unremarkable[17] with a grossly normal EF.[18] We ordered labs[19] for her including a D-dimer.[20] Her CBC,[21] BMP[22] and coags[23] were all negative and her urinary pregnancy test was negative[24]but her D-dimer was positive.[25]

LONG EXPLICIT

In room 3, is a 45-year-old[1] female[2] symptomatic of chest pain[3] under the right breast[4] accompanied by dyspnea[5] for duration of one day.[6] She states the pain started in the afternoon[7] after she had an argument with her son earlier[8] and describes it as stabbing[9] and radiating to her back.[10] Past medical history is not significant.[11] No recent trauma[12] or fever[13], she had ankle surgery 2 weeks ago.[14] Her pain has resolved since arrival, so she thinks it was just stress from the argument. So in her evaluation, her chest x-ray was clear[15] and EKG showed sinus tachycardia[16] and bedside echo did not show any right heart strain[17] and showed a grossly normal EF.[18] We ordered labs[19] for her including a D-dimer.[20] Her CBC,[21] BMP[22] and coags[23] were all negative and her urinary pregnancy test was negative[24]but her D-dimer was positive.[25] The plan is to do a chest CT[25] and then pending result treat accordingly[26] and she may need admission to medicine depending on the results.[27]

Figure 11: Example of Stimuli Narrative Permutation with Data Counts – Stimuli Lung Case of Pulmonary Embolism

6.8.4 Results: Information Volume as Case Narrative Key Data Counts

The assumption was that in terms of word count as measure of volume of information, for a given stimuli case type, the short/implicit cases would be at one end of the range with the smallest number of words. This would be followed by short/explicit, then long/implicit and then long/explicit case would be at the other end of the scale, with the highest key word count (see full counts in appendices).

The word count annotation for all the cases were checked with an EM physician SME for accuracy of the annotation. The results of the key data word count for all the cases are shown below in table 15.

Table 15

Summary of All Vignette Cases and Key Data Word Counts

Name	**CONTROL CASES**	**Stage**	**Count**
Epsilon	Pneumothorax	Early	17
Mu	Appendicitis	Early	25
Zeta	MVC-Fractured ulna	Late	22
Beta	Stroke – Ischemic	Middle	18
Kappa	DKA +UTI	Middle	19
Gamma	Alcohol Intoxication	Late	25
Lambda	Foot Fracture	End	15

Name	**STIMULI CASES**	**Stage – Count**	**Stage – Count**	**Stage - Count**	**Stage - Count**
Sigma	Lung: Pulmonary Embolism	SI – 20	SE – 23	LI – 25	LE – 27
Alpha	Head: UTI + Decubitus ulcer	SI – 27	SE – 30	LI – 33	LE – 37
Omega	Abdomen: Upper GI Bleed	SI – 20	SE – 22	LI – 26	LE – 29
Delta	Cardiac: Aortic Dissection	SI -35	SE – 37	LI – 39	LE – 41

6.8.5 Conclusion: Volume of Information as Case Narrative Key Data Counts

The volume of information presented as counts of key data items was observed to have pattern of increasing volume for the test conditions S/I through to L/E as anticipated. This pattern was observed consistently across all the stimuli cases, confirming that short implicit cases do have the smallest volume of information and long explicit case have the greatest volume of information.

6.9 SME Review of Vignettes and Study Constructs

The concept of Stage in the Diagnostic Process represented a temporal construct that signified the point at which the sign out occurred in the timeline of the workup and information gathering for a given patient case vignette.

In order to ascertain whether there was a common understanding of the construct of stage in the diagnostic process and its instances, a group of four emergency medicine SMEs reviewed the construct via a review survey that was created for this purpose. The review survey also enabled the case narratives for the patient vignettes to be reviewed by the SMEs for clinical accuracy and appropriate format.

i. Vignette and Construct Review Survey

A representative subset of patient cases was selected to develop a survey for the purpose of SME review of the definitions of the construct of stage in the diagnostic process. A combination of three control patient cases and four stimuli cases, were used to develop a survey in Qualtrics^{XM} software. The survey contained the same seven standard questions that was posed for each patient case presented in the survey.

TEXT: SIGN OUT NARRATIVE ABOUT THE PATIENT CASE

QUESTIONS

1. Cases progress through different stages. In the early stages, the patient is being evaluated and initial information is being gathered. During the middle of the case, information is coming in e.g., lab results and you understanding is evolving. The later stage of cases include final disposition of the case. Please identify the stage of the above case.
    - Early
    - Middle
    - Late

2. What diagnosis would you provide for the above patient?

3. For the above cases, is there any definitive bit of information (a test of lab result) that cinches your diagnosis? Please include information included in the case that is important and/or any information that is needed (but not available) above.

4. If you were to carry this case to completion, what would your next steps be?

Figure 12: Questions for SME Review of Constructs – each patient case in survey

Four emergency medicine physicians were provided with the URL link to the review survey and asked to review the cases presented in the survey and complete their responses to the questions in their own time.

ii. Results: Review of Stage in Diagnostic Process

The results for diagnosis and stage in the case identified by the SMEs were reviewed (table 16).

Table 16

SME Vignette Reviews

Chief Complaint	Patient Info	Condition	SME Diagnosis	Case type	Stage	SME Stages Summary	Comments
1. Lung – chest pain	45yr female	Pulmonary Embolism	3 X PE 1 x Chest pain	Short/ Explicit	Early/ Middle	4 x Middle	Consensus
2. Head– AMS	66yr male	Stroke – Ischemic	4x Stroke	Control	Middle	1x Early 1x Middle 2 x Late	Early was input error – should be Middle
3. Cardiac– chest pain	73yr male	Aortic Dissection	4x Aortic Dissection	Long/ Implicit	Late	2 x Middle 2 x Late	Even split among SMEs
4. Head - AMS	57yr male	Alcohol related hyponatremia	4 x Hyponatremia	Control	Late	2x Middle 2x Late	Even split among SMEs
5. Head - AMS	84yr female	UTI & Decubitus ulcer	2 x AMS 1X subdural hematoma 1x Sepsis	Short/ Implicit	Early/ Middle	3 x Early 1 x Middle	Most agreed on the intended stage
6. Lung – chest pain	27yr male	Pneumo-thorax	3x Pneumo-thorax 1x Chest pain	Control	Early	2x Early 2x Late	Greatest divergence in responses
7.Abdomen – chest pain	59yr male	Upper GI Bleed	4x Upper GI Bleed	Long/ Explicit	Late	1x Early 1x Middle 2x Late	Mixed variation in responses

iii.Conclusion

The results in table 16 indicated that for the diagnosis, most of the responses were correct or within the right diagnostic scope. The stimuli head case was associated with the lowest diagnostic accuracy, but this was a short implicit case. It was an earlier case with a greater level of ambiguity. So, this finding was somewhat expected, and during discussions, the SMEs stated an elderly patient with a fall has a wide differential diagnosis.

Identification of the stage in the process showed a greater level of variance in responses from the SMEs. In particular, for the stimuli case of upper GI bleed showed the greatest amount of variation in response. Also, the control case of pneumothorax, the response was split for identification of the stage in the diagnostic workup, which was intended to be an early case, but two responses stated it as late. The majority of the stimuli cases showed some consensus for amongst the SMEs for the stage of the case. Total consensus may be difficult to achieve due to the subjective nature of assessing the stage of a case may be in given that the case is presented as a snapshot of information.

iv. Revisions to Control Case Vignettes from Feedback from Two SME EM Physicians

The case narratives were reviewed in detail with two SME separately, and based on their findings, adjustments were made to the sign out narratives to address information issues that may have contributed to the variances in responses observed (see amended narratives in Appendix C)

The sign out narrative for each of the cases was reviewed with SMEs and edits were made to refine the information so that the cases would be more likely to align with the intended stages for the case as per the experimental design. The changes suggested by the SMEs are shown in red or blue. This was an iterative process conducted with multiple SMEs to ensure the greatest levels of consensus about the case stages was achieved.

With the sign out narratives reviewed, edited and finalized; they were ready to be used for the development of the sign out experiment.

6.10 Development of Final Survey for ED Sign Out Study

The within-subjects design involved random assignment for the stimuli permutation instances to the participants. To enable this the following randomization matrix was created to reflect all the permutations of the test conditions and chief complaints. (see table 17)

6.10.1 Randomization Across and Within Surveys

Table 17

Randomization Matrix of Test Condition Permutations

Rand.	Head	Cardiac	Abdominal	Lung
1	Short implicit	Long implicit	Long explicit	Short explicit
2	Long explicit	Short explicit	Short implicit	Long implicit
3	Long implicit	Short implicit	Short explicit	Long explicit
4	Short explicit	Long explicit	Long implicit	Short implicit
5	Long explicit	Short explicit	Long implicit	Short implicit
6	Short explicit	Long explicit	Short implicit	Long implicit
7	Long implicit	Short implicit	Short explicit	Long explicit
8	Short implicit	Long implicit	Long explicit	Short explicit
9	Long implicit	Short implicit	Long explicit	Short explicit
10	Short implicit	Long implicit	Short explicit	Long explicit
11	Long explicit	Short explicit	Long implicit	Short implicit
12	Short explicit	Long explicit	Short implicit	Long implicit
13	Long implicit	Long explicit	Short implicit	Short explicit
14	Short implicit	Short explicit	Long implicit	Long explicit
15	Long explicit	Long implicit	Short explicit	Short implicit
16	Short explicit	Short implicit	Long explicit	Long implicit

The randomization matrix provided a template for building the surveys that would be needed for the experiment. To ensure randomized assignment of the stimuli conditions across all the participants each permutation instance was developed into an individual survey instance. Each survey instance contained the seven control cases and the four

stimuli cases in a specified permutation instance of the test conditions and the chief complaints. (Table 18)

Table 18

Survey Builds with Control and Randomized Stimuli Cases

Survey #	Controls	Alpha	Delta	Omega	Sigma
1	#1-7	Short implicit	Long implicit	Long explicit	Short explicit
2	#1-7	Long explicit	Short explicit	Short implicit	Long implicit
3	#1-7	Long implicit	Short implicit	Short explicit	Long explicit
4	#1-7	Short explicit	Long explicit	Long implicit	Short implicit
5	#1-7	Long explicit	Short explicit	Long implicit	Short implicit
6	#1-7	Short explicit	Long explicit	Short implicit	Long implicit
7	#1-7	Long implicit	Short implicit	Short explicit	Long explicit
8	#1-7	Short implicit	Long implicit	Long explicit	Short explicit
9	#1-7	Long implicit	Short implicit	Long explicit	Short explicit
10	#1-7	Short implicit	Long implicit	Short explicit	Long explicit
11	#1-7	Long explicit	Short explicit	Long implicit	Short implicit
12	#1-7	Short explicit	Long explicit	Short implicit	Long implicit
13	#1-7	Long implicit	Long explicit	Short implicit	Short explicit
14	#1-7	Short implicit	Short explicit	Long implicit	Long explicit
15	#1-7	Long explicit	Long implicit	Short explicit	Short implicit
16	#1-7	Short explicit	Short implicit	Long explicit	Long implicit

6.10.2 Building the Simulated Sign Out Surveys

Once all the case vignettes, sign out narratives and randomization were all completed and checked the process of building the survey-based experiment commenced. The survey experiment was built in Qualtrics[XM]

The within-subjects design can increase the statistical power of a study with fewer units to test the study test conditions. To ensure even presentation of the test conditions, all the possible permutations of the test conditions should be presented to participants. This requires randomized assignment of the test conditions. In this case this would be the sixteen permutations of test condition combination and chief complaints.

i. Survey Build

1) Case Blocks, Randomized Assignment of Surveys

The complexity of this level of randomization to be executed within Qualtrics required the building of sixteen surveys, each containing a permutation instance as per the randomization matrix and survey build in tables 18 and 19. This was achieved by constructing a block, the term used in Qualtrics, for every instance of a case vignette. This was done for both the control cases and all the instances of the stimuli test condition cases.

Each block contained the sign out narrative for that case instance and seven questions pertaining to diagnostic decision making and the sign out information. These seven questions were the same for every block for each case in the study.

The case blocks were then used to construct sixteen individual surveys following the sequencing shown in Table 19. Once all the surveys were built, the overall survey experiment was set to randomly assign one of the sixteen surveys to each participant as they enrolled into the study. The random assignment was set to evenly assign the instances without replacement, so that there would be even presentation of each instance within set of sixteen.

2) Within Survey Randomization

However, within-subjects experimental designs can be prone to fatigue effects and carry over effects. To mitigate for the confounding impact of these issues, the individual units of a study can be randomized. To incorporate this approach into this experiment, the individual surveys were set to randomize the presentation order of the eleven cases within each survey. This means that each participant would get a random assignment of a survey instance from the set of sixteen surveys and the eleven cases within the survey would be presented in a randomized order also. So, if more than sixteen participants joined the study, the randomized assignment from the set of sixteen would start over again. Then the second participant to receive a survey instance would be presented with the eleven cases in a different order than the first participant who received that instance of the survey.

3) Survey Configurations

The other features that were included in the study were:

(1) Online consent – the first page of the survey contained an online informed consent form that was approved by the UTH and UMB IRB Boards. Participants were required to complete the online consent to be able to continue with the survey. If participants chose not to provide consent, they were exited from the survey.

(2) Demographic Data – the second page contained questions for participants to provide demographic information, which were their role and the number of years of emergency medicine experience they had.

(3) Question parameters – each of the seven questions on each of the eleven case blocks were set so that participants were required to provide an answer to be able to proceed through the rest of the survey. This was done to prevent the skipping over of questions and thus minimize the number incomplete or partially completed surveys.

(4) Bespoke exit page – a bespoke exit page was developed to enable provision of the gift card that was offered to participants who completed the entire survey. The two sites of UTH McGovern and UMB had different bespoke survey exit pages to enable the collection of appropriate participant information in compliance with each individual institution's reporting requirements.

(5) Incomplete and Aborted Responses – as participants were able to complete the survey at their own pace and in their own time the survey could be started and left open but not finished. Emergency medicine clinicians are extremely busy and have limited free time. For this reason, the surveys were set to allow participants a period of one week to

complete the survey, during which time it would show as 'in progress'. After the one-week period the survey would automatically close and exit the participant from the study. This response would show as incomplete status in the metadata of the survey response data file.

ii. Two Sites – Two Surveys

This experiment was conducted over two sites, UTH medical school and UMB medical school. To enable flexibility in management of the surveys, the two sites had mirrored experiments. Qualtrics allows the duplication of an entire survey project and so to accommodate different timelines and different consent processes, two mirror projects were created. Each project had a dedicated link to access the survey.

The two survey projects were identical in terms of content and the data being collected. The only parts that were different were the consent and exit pages of each survey.

This twinned project arrangement enabled ease of monitoring the progress of data collection, survey completion and the data quality of the responses at an individual site level. It also allowed ease of management of the participant recruitment and participant communication processes that were specific to each site.

The two sites each had their own participant response data files from the individual data collection process for each site. At the end of the study the two data files were downloaded and both projects closed. The two data files were joined to enable data analysis of the entire dataset as one data file.

6.11 Testing

During the survey build process, some testing of individual units and functions was conducted to ensure different components of the survey worked before moving on to joining the components together.

However, as the survey build process was near completion, systematic testing of the survey was conducted. There were two sets of testing that was conducted: testing for technical issues and logistics and testing of the randomization function.

6.11.1 Survey Flow and Function

A number of features of the survey were tested to identify and address any logistical or technical issues prior to survey launch. This reduces participant attrition and partially completed surveys due to issues experienced by participants when completing the study.

- Consent form & non-consent – the function of the informed consent screen was tested to ensure that it was not possible to progress without providing consent
- Exit screen working – survey completion and exit page messaging was tested
- Exit screen email launching – the UTH-McGovern survey exit page automatically launched a preformatted email so that participants could easily provide the information required to comply with institutional reporting functions. This correct function of this bespoke exit page was tested along with the email functions.

6.11.2 Randomization Testing

The reliable and consistent randomization function of the survey was systematically tested over three rounds of testing.

The first round of testing was conducted towards the completion of the build phase when all the structural elements of the survey were complete but before features like the question configuration had been switched on. This enabled the rapid testing of the survey flow and the randomization of the survey in production mode similar to a prototyping approach.

The second round of testing of the randomization was conducted when the survey project build for UTH McGovern was completed and was published but before it was distributed to participants. This enabled testing of the survey flow and randomization in a live set up. The survey was then duplicated so randomized assignment was reset and launched to initiate data collection from participant responses.

The third round of testing was conducted once the UMB survey project build was completed and published but prior to launching the study at that site. The testing was repeated with the UMB survey as this survey was a mirrored survey but with different participant consent and exit pages. Testing ensured that the survey flow and randomization functioned without issues similar to the UTH McGovern survey.

The rounds of testing consisted of running through the survey a minimum of eighteen times and making note of the order of the cases presented for each. This was to ensure completion of a least one complete randomization cycle through the set of sixteen survey permutations. The lists of cases and the details of the cases from the notes were checked off against the randomization matrix to check that all instances of the randomization had indeed been presented. The randomized assignment in the live surveys and the case

randomization within the survey cases were found to be working accurately and consistently.

6.12 Data Collection

The data collection phase of this project involved launching this study across two sites. The first site was UTHealth McGovern Medical School, Houston, Texas and the second site was UMB in Baltimore, Maryland. This required two participant recruitment processes and two data collections. Participant recruitment was enabled and coordinated by co-investigators at UTH McGovern and UMB sites, who were named on the IRB protocols, Amit Mehta, MD and Brent King, MD respectively.

6.11.1Participant Recruitment and Survey Launch

Dr. Mehta and Dr. King communicated to potential participants directly at departmental meetings and using email with fliers via departmental email groups, for their respective sites. The participant pool contacted included attendings, residents, fellows, and APPs at both sites. Once the provision to offer gift cards to participants completing the survey was approved as compliant with each institution's regulatory policies, the surveys were launched.

The participant recruitment communications were coordinated to align with the survey launch dates:

The UTH McGovern Survey launched on November 15th, 2021.

The UMB Survey launched on November 19th, 2021.

The surveys were planned to close on January 7th, 2022, allowing data collection to proceed for a duration of 7 weeks. Follow up reminder recruitment communications were sent by Dr. King and Dr. Mehta at two intervals during the data collection period. Additionally, this researcher attended the McGovern Department of Emergency Medicine weekly resident didactic conference on December 9th, 2021 and gave a participant recruitment presentation to the residents and faculty.

As part of the participant recruitment email an URL link for the survey was included. The survey link was specific for each survey at each site and enabled the participant to anonymously complete the survey. No personal identifying information about the participants was collected via the survey. Those participants who wished to receive the gift card upon completion of the survey, contacted the researcher directly via the email that was set to launch and autofill from the survey at the point of exiting.

6.12.2 Survey Close

The responses and participant numbers were monitored throughout the data collection period to check that no technical or other issues with the surveys that might interfere with the continuity of the data collection.

Participants received a final communication notifying them of the upcoming close date of the surveys. Both surveys were closed on January 7th, 2022, as planned and the participant response data files for each site survey were downloaded from Qualtrics onto a UTH encrypted computer drive as per the IRB protocol.

6.12.2 Data Management

The survey response data showed that at the time of closing there were:

UTH McGovern	60 responses	35 completed responses
UMB	50 responses	34 completed responses

The response data were examined, and it was noted that nearly all the incomplete responses for both sites showed participants left at the consent page of the survey. So, in conclusion the Emergency Department Sign Out (EDSO) project had a total of 69 participants overall. The data files were stored and managed on UTH laptop computers only and stored for access by study co-investigators on UTH Secure-share drives.

6.13 Survey Questions

In Qualtrics, the questions in the survey were grouped onto blocks, with a dedicated block for each patient vignette case. While each case was presented on a separate block, the questions to the participants for that case were the same for each case vignette i.e., every block had the same seven question for every case in the survey.

For each of the eleven cases in a survey the following questions were asked of the participant:

1) Cases progress through different stages in the diagnostic reasoning process. In the early stages, the patient is being evaluated and initial information is being gathered. During the middle of the case, information is coming in e.g., lab results and you understanding is evolving. The later stage of cases includes final disposition of the case. Please identify the stage of the above case.

- Early – Patient is being evaluated and initial information is gathered
- Middle – Information is coming in and understanding is evolving

- Late – Confidence in clinical impression is high or disposition is known

2) What diagnosis and any alternatives are you considering for the above patient? (with your primary diagnosis being first)

3) For the above case, what are the definitive pieces of information or data included in the case that confirms your diagnosis? (e.g., lab results or imaging study)

4) If you were presented with this case, what would your immediate next steps be?

5) We are interested in how you felt answering the tasks for this survey. Please indicate your level of confidence (by %) for the following:

6) My confidence in determining a diagnosis with the available information is ___%.

7) With the information available I was __% confident to sufficiently consider my scope of differential diagnoses

8) I am _____ % ready to plan treatment or disposition for this patient based on the information provided in the sign out.

For the first question required participants to select one option from the radio buttons of either Early, Middle or Late for the stage of the case they have been presented. The second question required completion of a text box field with a list of one or more diagnoses for the case.

The third question was more about what information about the case the participant considered. Their response required completion of a text box field where they can specify none, one or more items.

Question four was about treatment planning and asked the participant fill in a text box field with one or more of the follow-up steps they would do next for the case presented.

The last three questions, five to seven, each had a slider function that the participant could adjust to indicate their level of confidence for the questions asked about their information needs. The participant was able to move the slider bar to indicate a level of confidence expressed as percentage value.

Upon completion of all the questions for each case the participant could progress to end of the survey and exit. At this point they could decide if they wished to follow up receiving the gift card or not.

6.13.1 Study Measures

The data collected were the answers that participants provided to the survey questions. The types of data varied by the question and the questions were structured to elicit data for the measures of interest related to diagnostic decision making and patient management. (see table 19)

Table 19

EDSO Study Measures

Survey Question	Participant Response	Method of response	Data type	Measure
1.	Select either Early, Middle or Late for Stage in the diagnostic process	Select one option from radio buttons	categorical variable	What stage in the diagnostic process is the case perceived to be at from the information provided
2.	Participant's diagnosis and differential	list with one or more values	categorical variable	Presumed diagnosis for the patient and any others they are still considering
3.	Key data for the case	list with zero or more values	categorical variable	Information in the case that aided their decision making
4.	Next steps consistent with diagnosis or differential	list with one or more values	categorical variable	Other data still needed to proceed or finalize the case
5.	Percentage confidence in diagnosis provided	Moves slider to indicate numerical value	continuous variable	How sure that they identified the correct diagnosis
6.	Percentage confidence in scope of diagnostic	Moves slider to indicate numerical value	continuous variable	Did they feel they had sufficient information to develop their diagnostic impression
7.	Percentage ready to plan treatment	Moves slider to indicate numerical value	continuous variable	Did they get what they need to complete the diagnosis and treatment for the case

6.13.2 What is Being Measured

1) Stage in the process – is there consistency in how participants perceive where a case is in the diagnostic process when a sign out occurs from the information provided?

2) Diagnosis and differential – Did they identify the correct diagnosis and/or is it part of their differential?

3) Key Data – what and how much data was important for decision making. Was it provided in the case?

4) Next Steps planning – Does the participant want to do more data gathering e.g., more tests or are they ready to move to treatment for the case

5) Confidence in Diagnosis – how sure are they that they identified the correct diagnosis or are they still considering other options

6) Confidence in Scope of Differential Diagnoses – how did they feel about the diagnostic they were presented with?

7) Confidence to Treat – Do they feel they have enough information to go ahead and manage the patient?

6.14 Data Analysis

6.14.1 Data Processing

Once the survey closed, the data files for the participant responses for the UTH and UMB surveys were downloaded from the Qualtrics^{XM} software. The Qualtrics generated .csv files were for complete responses to the survey and after they were downloaded, they were converted into spreadsheets in Microsoft Excel.

As the survey required participants to provide a response for every question to be able to progress through the survey, all the data files contained complete records only. There were no incomplete or missing data or partial records in the participant response data.
A number of additional columns representing created variables were added to the spreadsheet. This allowed for the recording of data synthesized from coding the responses in the data from participants to questions posed in the survey.
A number of created variables were inserted as new columns into the datafile. The first set of created variables were to transfer descriptive details about the data present in the original files such as:

- Location – identifier whether the data was from either the UTH site or the UMB site
- Case – the type of case – control or stimuli
- Chief Complaint – was the type of body system associated with the case that the patient presented with in the ED
- Length – indicating whether the case was a Short (Middle) or Long (Late) case
- Explicitness – was the case an Explicit or Implicit case

A second set of created variables were inserted as new columns to record data synthesized by interpreting the responses to the questions. The created variables were:

- Resp Match – compared the participant's response for the variable Stage to the actual stage of the case as it was designed.
- Diagnosis – represented diagnostic accuracy by comparing the first item in the participant's differential diagnosis for the response to question 2 in the survey with the actual diagnosis associated for the case.

- Included – represented whether the correct diagnosis was included in the participant's differential. This was done by looking to the recorded variable *Differential*, the response to question 2 in the survey, and checking whether the correct diagnosis was present in the list.
- More Than One – represented whether more than one diagnosis was provided in the differential and looked to the response to question 2 for the recorded variable Differential to see if one diagnosis was provided or more than one in the participant's response.
- Count – was a simple count of the number of different diagnoses provided for the variable Differential by participants in response
- Data Match – compared the participant's response to survey question 3; items listed for recorded variable Key Data were compared to data items in the case sign out narratives to note whether any matched
- Next step match – the contents of the recorded variable *Next Steps* for survey question 4, were compared for a match to the recommended next steps in the Case Clinical Notes document, which were ultimately derived from the two medical reference texts used to develop the control and stimuli case vignettes.

For some of these created variables this involved a simple binary decision of a Y/N measure for the presence of the data, for others it involved some interpretation of the response data. For the created variable called *Diagnosis*, it required reviewing the responses for the second question in the survey that asked participants to list the possible diagnoses they were considering, i.e., their differential, with their presumed diagnosis

listed first. Then for each record the first item listed for this field was copied into the *Diagnosis* column for that record.

Data synthesis for the following created variables involved some interpretation of the participant's responses to assess whether it matched an appropriate response. It was then coded using a binary measure of Yes for a match or No there was no match to the appropriate response respectively.

This was conducted for the responses to all the cases, controls, and stimuli. Where a cell in the spreadsheet contained no data, as would be the case for the responses related to the randomized stimuli conditions, a value of 999 was entered.

6.14.2 Data Variables

The data processing resulted in a dataset containing the following columns of data listed in Table 20

Table 20

Survey Data: Responses and Created Variables

#	Column name	Survey/created?	Description	Data Type
1.	Role	Survey	The job title/role of the participant	Categorical
2.	Experience	Survey	How many years of experience in EM	Continuous
3.	Random ID	Survey	Unique ID generated by Qualtrics – for tracking provision of gift card	Integer
4.	Location	Created – transferred	Location of participant affiliation – UTH or UMB	Categorical
5.	Case Length	Created – transferred	Is the case Short (Middle) or Long (Late)	Categorical
6.	Explicitness	Created – transferred	Is the case implicit or explicit	Categorical
7.	Chief Complaint	Created – transferred	type of condition the patient presented with	Categorical
8.	Stage	Created – transferred	The participant's stated stage for the case	Categorical
9.	Resp Match	Created – coded	Did the participant correctly identify the stage	Categorical
10.	Differential	Survey	The items included in differential diagnoses	Categorical
11.	Diagnosis	Created – coded	The primary presumed diagnosis (listed first)	Categorical
12.	Accurate	Created – coded	Is the diagnosis specified correct	Categorical
13.	Included	Created – coded	Is the correct diagnosis in the differential?	Categorical
14.	More than one	Created – coded	Are there more than one differential	Categorical
15.	Count	Created – coded	How many items listed in differential	Integer
16.	Key Data	Survey	Key data to aid in diagnosing the case	Categorical
17.	Next Steps	Survey	What next steps would they do for the case	Categorical
18.	Conf in Diag.	Survey	Confidence in their diagnosis for the case	Continuous
19.	Diagnostic Scope	Survey	Confidence in considering scope of diagnosis	Continuous
20.	Treatment	Survey	Confidence in proceeding with treatment	Continuous
21.	Anchor	Created – coded	Did the participant anchor in their diagnosis	Categorical
22.	Data Match	Created – coded	Is key data the case provided in SO	Categorical
23.	Next Step Match	Created – coded	Consistent with diagnosis?	Categorical

6.14.3 Outcome Variables

The primary outcomes of interest were:

i. Diagnostic Accuracy was determined via assessing the created variable of *Accurate* across cases and by location, by role, by the case, by chief complaint, by length and by explicitness

ii. How confident the participant felt in the diagnosis they provided was determined by analyzing the recorded continuous variable of percentage confidence for question 5 in the survey, Confidence in their primary diagnosis was determined by analyzing across the cases and by location, by role, by case type, by chief complaint, by length and by explicitness

iii. How confident participants felt in the scope of their diagnostic differential was determined via analysis of the recorded continuous variable of percentage confidence response for question 6 in the survey. Confidence in considering their scope of differential diagnoses was determined by analyzing the recorded variable *Diagnostic Scope* across the cases and by location, role, by case type, by chief complaint, by chief complaint, by length and by explicitness

iv. Participant's inclination to proceed with their selected treatment via analysis of the recorded continuous variable of percentage in response to question 7 in the survey. Percentage ready to proceed with treatment or disposition was determined by analyzing the recorded variable *Treatment* across the cases and by location, role, by case type, by chief complaint, by chief complaint, by length and by explicitness.

The secondary outcomes of interest were:

i. To determine whether participants identified the stage of the case correctly, the created variable *Response Match* was analyzed first across all the cases and then by location, by case type, by chief complaint, and then by length of case and by explicitness

ii. To determine whether participants had identified the correct diagnosis within their differential, the created variable Included was analyzed across all the case and then by location, by case type, by chief complaint, and then by length of case and by explicitness

iii. To determine the extent to which the participants were committed to a single diagnostic path the created variable *More Than One*, which indicated whether their differential contained one or more diagnoses, was analyzed is a similar manner across all the case and then by location, by case type, by chief complaint, and then by length of case and by explicitness

iv. To understand how broad participants' differential diagnoses were, the created variable *Count* was analyzed for all the cases and then by location, by case type, by chief complaint, and then by length of case and by explicitness

v. To determine whether the data the participants listed as key data for their decision making for question 3 in the survey corresponded to the data provided in the case the created variable *Data Match* was analyzed for all cases and by location, by case type, by chief complaint, and then by length of case and by explicitness

vi. To determine whether participants specified the appropriate next steps for the case the created variable of *Next Steps Match* was analyzed for all cases and by location, by case type, by chief complaint, and then by length of case and by explicitness.

6.15 Statistical Analysis

To determine whether the test conditions that were considered risk factors for anchoring had any effect on the diagnostic and treatment decision making in participants, statistical analyses was conducted. The statistical software R (version 3.5.1) with RStudio (version 1.1.463) was used to conduct this analysis. For the statistical analyses the bivariate effects of categorical variables on the outcomes were assessed using Pearson's Chi-square test or Fisher's exact test. Continuous variables were assessed using a two-sample t-test or one-way analysis of variance (ANOVA).

Chapter 7: ED Sign Out Study Findings

Analyses were conducted on the response data for the primary and secondary outcomes using appropriate statistical methods. For each analysis, p-values less than 0.05 were considered statistically significant.

7.1 Analysis of Study Population Demographics

The first analysis was for demographic features of the participant population and table 21 below presents the demographic descriptive statistics.

The three groups are attending physicians, trainees, who include residents and fellows and the label other for advance practice providers (APPs) which includes nurse practitioners and physicians' assistants.

Table 21

Role of Participants by Location

ROLE	UMB	UTH	Overall
Attending Physician			
n	23	8	31
Overall %	33%	12%	45%
Row %	74%	26%	
Column%	68%	23%	
Residents & Fellows			
n	10	14	24
Overall %	14%	20%	35%
Row %	42%	58%	
Column%	29%	40%	
APPs			
n	1	13	14
Overall %	1%	19%	20%
Row %	7%	93%	
Column %	3%	37%	

N = 69 participants

Pearson $\chi^2 = 18.2$, df = 2, p-value <0.01

The distribution of roles of participants in the study by location was found to be statistically significant as not similar.

Of the participants who were attending physicians 74% were from UMB and 26% were from UTH. For participants who were trainees, that is either a resident or a fellow, 42% were from UMB and 58% were from UTH, while the corresponding figures are that of the APPs group most (93%) were from UTH. In terms of the overall participants, UMB had proportionally more attendings (68%) than residents (29%) and only one APP. Whereas with UTH there was more even distribution across the roles with 23% as attendings, 40% as residents, and 37% as APPs.

7.1.2 Years of ED Experience by Location

The overall mean length of experience was 6.8 (SD 6.49) years. The overall range was 1–35 years. As such, the participants from UMB had a longer mean experience (8.5; SD 5.92) years compared to participants from UTH (5.1; SD 6.67) years. ($t = 2.272$, $df = 67$, p-value = 0.03).

Finding: The demographic analysis suggests that the participants from the two sites are not similar populations, in terms of role or years of experience in emergency medicine.

7.2 Stage in the Diagnostic Process

The next set of results examined was whether participants correctly identified the stage in the diagnostic process of the case was in the experimental case.

Table 22

Accurate Identification of Stage of Case

Was Stage of the Case Correctly Identified	**Yes**	**No**	**Pearson χ^2 df**	**p-value**
Role			$\chi^2 = 0.109$ df =2	**0.95**
Attendings	62%	38%		
Residents & Fellows	64%	36%		
APPs	62%	38%		
Location			$\chi^2 = 0.0538$ df =1	**0.82**
UMB	62%	38%		
UTH	63%	37%		
Case			$\chi^2 = 1.9552$ df =1	**0.16**
Controls	65%	35%		
Stimuli	59%	41%		

N=759

Findings: Differences in the identification of the stage of the case by role was not statistically significant.

Difference in the identification of the stage of the case was not found to be statistically significant for the location or for the type of case, as in between controls or stimuli cases. The stage in the diagnostic process was identified correctly in 65% of control cases and 59% of stimuli cases.

7.3 Diagnostic Accuracy

The results for the statistical analyses to determine whether participants provided the correct diagnosis for the cases, are presented in table 23 Diagnostic Accuracy Results.

Table 23

Diagnostic Accuracy Results

Was Diagnosis Correct	**Yes**		**No**	**p-value**
Role				**P=0.99**
Attendings	75%		25%	
Residents and Fellows	75%		25%	
APPs	76%		24%	
Location				**p= 0.48**
UMB	77%		23%	
UTH	74%		26%	
Case				**P<0.01**
Controls	92%		8%	
Stimuli	54%		46%	
Chief Complaint – Overall				**p<0.01**
1. Stroke (Middle)	100%		0%	
2. DKA+UTI (Middle)	97%		3%	
3. Pneumothorax (Early)	96%		4%	
4. Appendicitis (Early)	91%		9%	
5. MVC- Fractured ulna (Late)	91%		9%	
6. Jones fracture (End)	90%		10%	
7. Lung – Stimuli	86%		14%	
8. AMS – Alcohol hyponatremia (Late)	80%		20%	
9. Cardiac – Stimuli	77%		23%	
10. Abdomen – Stimuli	22%		77%	
11. Head – Stimuli	1%		9%	
Chief Complaint – Controls				**p<0.01**
1. Stroke (Middle)	100%		0%	
2. DKA+UTI (Middle)	97%		3%	
3. Pneumothorax (Early)	96%		4%	
4. Appendicitis (Early)	91%		9%	
5. MVC- Fractured ulna (Late)	91%		9%	
6. Jones fracture (End)	90%		10%	
7. AMS-alcohol Hyponatremia (Late)	80%		20%	
Chief Complaint – Stimuli				**p<0.01**
1. Lung	86%		14%	
2. Cardiac	77%		23%	
3. Abdomen	22%		78%	
4. Head	1%		99%	
Stimuli – Explicitness				**P=0.55**
Explicit	49%		51%	
Implicit	44%		56%	
Stimuli – Length				**P<0.01**
Long	60%		40%	
Short	33%		67%	

7.3.1 Interpretation of Analyses for Diagnostic Accuracy

Finding: Differences in diagnostic accuracy for the cases was not found to be statistically significant when examined by role or between the two locations.

Finding: The difference observed in diagnostic accuracy for the cases was statistically significant by the type of case as to whether the case was control or stimuli case.
To explore which cases may be contributing to this finding the analysis was repeated with the control cases as a group and the stimuli cases as a group.

Finding: The differences observed in diagnostic accuracy were found to be statistically significant across all the control cases, when compared by chief complaint. The results suggest that the chief complaint did have an affect the diagnostic accuracy for the control cases. Of all the control cases with an incorrect diagnosis, the responses for the case of altered mental state (AMS) due to alcohol related hyponatremia had the highest level of incorrect diagnosis at 37%. Indeed, across all the control cases this particular case of alcohol related hyponatremia had a percentage of responses that were correct for diagnosis at 80%, which was the lowest for all control cases.

Finding: The differences observed in diagnostic accuracy were found to be statistically significant across all the stimuli cases when compared by chief complaint. The result suggests that the chief complaint did have an affect the diagnostic accuracy for the stimuli cases. This effect may be likely due to the very low level of diagnostic accuracy observed in the results for the stimuli head case, for which 99% of participants did not provide a correct diagnosis.

These results were considered in terms of hypotheses of how the test conditions of explicitness of the case and stage of the case influenced diagnostic accuracy:

Hypothesis 1: The explicitness of how a case is presented would affect how accurately participants could identify the correct diagnosis.

Finding: Differences in diagnostic accuracy for the cases was not statistically significantly different when examined by the type of case, whether the case was control or stimuli case and therefore the hypothesis was rejected. The results indicated that explicitness of the sign out information in a case did not affect participants' ability to accurately diagnose the case.

Hypothesis 2: The length of the case would affect how accurately participants could identify the correct diagnosis.

Finding: Differences in diagnostic accuracy for the cases was statistically significant for the construct of the length of the case. The results indicate that length of the case, in other words the volume of information presented in a sign out related to the length of the case the stage of the case in these stimuli cases, did affect participants' ability to accurately diagnose the case

7.3.2 Analysis of Diagnostic Accuracy and Identification of Stage of the case

Table 24

Matched Stage by Accurate Diagnosis (Stimuli Cases Only)

Was Diagnosis Accurate Identified Stage of Case Correctly	**Accurate Dx: No**	**Accurate Dx: Yes**
Matched Stage: No		
N	46	66
Overall %	17%	24%
Row %	41%	59%
Column %	31%	52%
Matched Stage: Yes		
n	102	62
Overall %	37%	22%
Row %	62%	38%
Column %	69%	48%

N = 276 responses Pearson $\chi^2 = 11.107$, df = 1, p-value <0.01

Finding: The variations observed in diagnostic accuracy when compared to whether participants were able to identify the stage in the diagnostic process for the case were found to be statistically significant.

The results suggest that of those who did not provide an accurate diagnosis, 69% correctly identified the stage in the diagnostic process and 31% did not. Whereas those who were accurate for the diagnosis were no better at identifying the stage in the diagnostic process, with 49% correct for stage in the process, compared to 52% who were incorrect.

Table 25

Additional Details for Diagnoses and Stage

Study Measure	**Count**
All case instances	759
→All instances of incorrect diagnosis for the case	186
→Incorrect diagnosis + incorrect stage	57
→incorrect diagnosis + incorrect stage + Earlier case identified as Late	50
→ incorrect diagnosis + incorrect stage + Late case identified as Earlier	7

7.4 Analyses of the Participants' Differential Diagnoses

A series of analyses were run on the responses to question 3 in the survey, where participants were asked about the diagnoses, they considered for the case with their primary diagnosis listed first. The analyses include i. whether the correct diagnosis was listed in their differential, ii. Did they include more than one diagnosis in their differential, iii. The number of diagnoses included in their differential.

i. Is Correct Diagnosis included in the participants' differential diagnoses

The variable of whether the correct diagnosis had been included within the differential was analyzed and presented as the measure Correct Diagnosis Included in Differential in table 26.

ii. Are there More than One Diagnosis in Participants' Differential Diagnoses

The response for question 3 was reviewed for the presence of a single diagnosis or multiple diagnoses in the differential provided and was represented by the measure More than One Diagnosis in Differential in table 27.

iii. The Number of Diagnoses in Participants' Diagnostic Differential

The number of diagnoses in the participants' differential diagnosis for question 3 were analyzed for the mean diagnoses across the cases and presented as the measure Number of Diagnoses in Differential in table 28.

Table 26

Correct Diagnosis Included in Differential

Was Diagnosis included in Differential	**Yes**	**No**	**p-value**
Role			**p =0.83**
Attendings	83%	17%	
Residents and Fellows	81%	19%	
APPs	81%	19%	
Location			**p = 0.30**
UMB	83%	17%	
UTH	80%	20%	
Case			**p<0.01**
Controls	96%	4%	
Stimuli	57%	43%	
Chief Complaint – Overall			**p<0.01**
1. Stroke (Middle)	100%	0%	
2. Appendicitis (Early)	99%	1%	
3. DKA +UTI – (Middle)	99%	1%	
4. MVC- Fractured ulna (Late)	97%	3%	
5. Pneumothorax (Early)	97%	3%	
6. Lung – Stimuli	96%	4%	
7. Jones fracture (End)	91%	9%	
8. Cardiac – Stimuli	91%	9%	
9. AMS Alcohol Hyponatremia (Late)	90%	10%	
10. Abdomen – Stimuli	30%	70%	
11. Head – Stimuli	10%	90%	
Chief Complaint – Controls			**p = 0.01**
1. Stroke (Middle)	100%	0%	
2. Appendicitis (Early)	99%	1%	
3. DKA+UTI – (Middle)	99%	1%	
4. MVC- Fractured ulna (Late)	97%	3%	
5. Pneumothorax (Early)	97%	3%	
6. Jones fracture (End)	91%	9%	
7. AMS Alcohol Hyponatremia (Late)	90%	10%	
Chief Complaint – Stimuli			**p<0.01**
Lung	96%	4%	
Cardiac	91%	9%	
Abdomen	30%	70%	
Head	10%	90%	
Stimuli – Explicitness			**p=0.63**
Explicit	59%	41%	
Implicit	55%	45%	
Stimuli – Length			**p<0.01**
Long	70%	30%	
Short	44%	56%	

Table 27

More than One Diagnosis in Differential

More than One Diagnosis in Differential	Yes	No	p-value
Role			**p = 0.01**
Attendings	78%	22%	
Residents and Fellows	80%	20%	
APPs	68%	32%	
Location			**P = 0.11**
UMB	79%	21%	
UTH	74%	26%	
Case			**p<0.01**
Controls	96%	4%	
Stimuli	57%	43%	
Chief Complaint – Overall			**p<0.01**
Stimuli – Head	93%	7&	
Appendicitis (Early)	88%	12%	
Stimuli – Abdomen	88%	12%	
AMS- Alcohol- Hyponatremia	87%	13%	
DKA +UTI (Middle)	87%	13%	
Stimuli – Cardiac	82%	18%	
Stimuli – Lung	80%	20%	
Pneumothorax (Early)	78%	22%	
Stroke	65%	35%	
MVC- Fractured ulna (Late)	52%	48%	
Jones Fracture (End)	40%	60%	
Chief Complaint – Controls			**p<0.01**
Appendicitis (Early)	88%	12%	
AMS- Alcohol- Hyponatremia	87%	13%	
DKA +UTI (Middle)	87%	13%	
Pneumothorax (Early)	78%	22%	
Stroke	65%	35%	
MVC- Fractured ulna (Late)	52%	48%	
Jones Fracture (End)	40%	60%	
Chief Complaint – Stimuli			**p =0.12**
Head	93%	7%	
Abdomen	88%	12%	
Lung	80%	20%	
Cardiac	82%	18%	
Stimuli – Explicitness			**p = 0.71**
Explicit	85%	15%	
Implicit	87%	13%	
Stimuli – Length			**p = 0.80**
Long	82%	18%	
Short	90%	10%	

Table 28

Number of Diagnoses in Differential

Number of Diagnoses in the Differential	**Mean (SD)**	**p-value**
Role		**p<0.01**
Attendings	2.8(1.45)	
Residents and Fellows	3.0(1.59)	
APPs	2.4(1.33)	
Location		**p = 0.93**
UMB	2.8(1.44)	
UTH	2.8(1.54)	
Case		**p<0.01**
Controls	76.9(20.1)	
Stimuli	56.5(24.4)	
Chief Complaint – Overall		**p<0.01**
1. Stimuli – Head	3.7(1.82)	
2. Appendicitis (Early)	3.3(1.61)	
3. Stimuli – Abdomen	3.1(1.34)	
4. AMS- Alcohol- Hyponatremia	3.1(1.23)	
5. Pneumothorax (Early)	3.0(1.59)	
6. DKA +UTI (Middle)	3.0(1.34)	
7. Stimuli – Lung	2.8(1.28)	
8. Stimuli – Cardiac	2.7(1.36)	
9. Stroke (Middle)	2.4(1.36)	
10. MVC- Fractured ulna (Late)	1.9(1.11)	
11. Jones Fracture (End)	1.6(0.91)	
Chief Complaint – Controls		**p<0.01**
1. Appendicitis (Early)	3.3(1.61)	
2. AMS- Alcohol- Hyponatremia	3.1(1.23)	
3. Pneumothorax (Early)	3.0(1.59)	
4. DKA+UTI – (Middle)	3.0(1.34)	
5. Stroke (Middle)	2.4(1.36)	
6. MVC- Fractured ulna (Late)	1.9(1.11)	
7. Jones Fracture (End)	1.6(0.91)	
Chief Complaint – Stimuli		**p<0.01**
1. Head	3.7(1.82)	
2. Abdomen	3.1(1.34)	
3. Lung	2.8(1.28)	
4. Cardiac	2.7(1.36)	
Stimuli – Explicitness		**p = 0.53**
Explicit	3.0(1.48)	
Implicit	3.1(1.55)	
Stimuli – Length		
Long	2.7(1.36)	**p<0.01**
Short	3.4(1.59)	

7.4.1 Interpretation of Analyses of Differential Diagnoses

i. Correct Diagnosis Included in Differential

Finding: Any differences in whether the correct diagnosis was included in the differential diagnoses between the two locations of UMB and UTH were not found to be statistically significant, nor were they significant when compared by role. So, the location or role of participants had no effect on whether the correct diagnosis was included in the differential.

Finding: The differences between the control cases and stimuli cases for whether the diagnosis was included in the differential was found to be statistically significant.

The results suggest that the type of case had a significant effect on whether the correct diagnosis for the case was mentioned in the differential provided by the participants. In 96% of control cases the differential diagnosis included the correct diagnosis, whereas in only 57% of stimuli cases was this the case.

Finding: Any differences in whether the correct diagnosis was included in the differential diagnosis across the cases by chief complaint was found to be statistically significant

To explore which cases may be contributing to this finding, the analysis was repeated with the control cases as a group and the stimuli cases as a group.

Finding: The differences in whether the correct diagnosis was included in the differential found to be statistically significant for the control cases when compared by chief complaint.

The results suggest that the chief complaint influenced whether the correct diagnosis was considered as part of the differential for the case when looking at the control cases. This

effect may be due to the lower incidence of the correct diagnosis being in the differential for the case of AMS due alcohol intoxication related hyponatremia and the case of the Jones fracture, which were 90% and 91% respectively. Whereas for the other chief complaints the inclusion of the correct diagnosis in the differential was significantly higher at 97% or above.

Finding: For the stimuli cases the differences in whether the correct diagnosis was included in the differential were also found to be statistically significant when compared by chief complaint.

The results suggest that for the stimuli cases the chief complaint influenced whether the diagnosis was identified within the differential. The abdomen and head cases had considerably lower levels of the correct diagnosis being included in the differential for the case. The abdomen case and head case were at 30% and 10% respectively, compared to that observed for the heart and lung cases, which both had very high levels at over 90%

The stimuli cases were further analyzed to determine whether the test conditions had any effect on whether the correct diagnosis had been included in the diagnostic differential provided by the participant.

Finding: The results show that the correct diagnosis being within the differential diagnoses was not statistically significant for the explicitness of the information in the case.

Finding: The results showed that the correct diagnosis being in the differential was statistically significant for the length of the case in stimuli cases. In other words, the

length of the case, whether long or short, did influence whether the correct diagnosis was included within the differential for the case.

ii. More than One Diagnosis Included in Differential

Finding: The results showed that the differences observed for whether the participant provided more than one diagnosis in their differential compared by the role of the participant was statistically significant.

While the majority of all participants gave more than one diagnosis in their differential, within the APPs group a greater proportion gave only one diagnosis, 32%, compared to the proportion of those who gave one diagnosis in the other groups of attendings and trainees with 22% and 20% respectively. In terms of those providing one diagnosis in their differential, 42% of responses were from attendings, while the trainees (residents and fellows), and APPs were roughly similar at 30% and 28% respectively.

Finding: The results for any differences in whether there was more than one diagnosis in the differential showed no statistically significant difference by the location. In other words, there were no difference between responses from UMB or UTH.

Finding: Whether more than one diagnosis was in the differential was found to be statistically significant by the type of case, with stimuli cases showing a higher proportion of differentials that contained more than one diagnosis.

Finding: The differences in whether there was more than one diagnosis in the differential was found to be statistically significant when compared by chief complaint across all the case types overall, as well as by chief complaint in the control cases as a group and also by chief complaint in the stimuli cases as a group.

This finding was examined further by analyzing whether more than one diagnosis in the differential participants provided was influenced by the test conditions of explicitness and length. The results are as follows:

Finding: Any differences in whether more than one diagnosis was included in the differential were not found to be statistically significant for the condition of explicitness.

Finding: Differences in whether the differential contained more than one diagnosis were not found to be statistically significant for the condition of length of case.

iii. Number of Diagnoses in Differential for the Case

Finding: Variations in the number of diagnoses included in the differential was found to not be statistically significant for the location of the participant**.**

Finding: Variations in the number of diagnoses included in the differential however was found to be statistically significant for both the role of the participant and also for the type of case, whether control or stimuli.

Finding: The differences in the number of diagnoses included in the differential was found to be statistically significant when compared by chief complaint across all the case types, as well as by chief complaint in the control cases as a group and by chief complaint in the stimuli cases as a group.

The stimuli cases were further examined for the test conditions of explicitness and length of case to determine whether these conditions had any effect on the number of diagnoses counted in the differential for the case.

Finding: Differences in the number of diagnoses provided in the differential was found to not be statistically significant when compared by the explicitness of the case information.

Finding: Differences in the number of diagnoses provided in the differential diagnoses were statistically significant when compared by the length of the stimuli case.

7.5 Analyses of Participants' Responses about Their Confidence Levels for Diagnosis and Treatment for Cases

The following results are related to the responses for question 5 in the survey, which was comprised of three additional questions. The three questions were answered by moving a slider button to indicate their level of confidence expressed as percentage. The three questions were to estimate for each case, given the information they were provided:

1. Their level of confidence in the diagnosis they provided for the case
2. Their level of confidence in considering the diagnostic scope of the differential diagnoses for the case
3. Their level of confidence with proceeding to either treat the patient or arrange disposition for the case

7.5.1 Participants' Confidence in Their Diagnosis for the Case

The results for the analyses of participants' confidence in the diagnosis they provided for each case were as follows mean (SD):

Overall mean confidence in diagnosis: 69.5(23.85)

7.5.2 Participants' Confidence in Scope of Diagnostic Differential

The results for the analyses of participants' confidence in considering the scope of their differential diagnoses given the information provided for each case were as follows:

Overall mean confidence in scope of diagnostic differential: 76.3(20.65)

7.5.3 Participants' Readiness to Plan Treatment or Disposition

The results for the analyses of participants' readiness to plan treatment or disposition for the case given the information provided were as follows:

Overall mean readiness to plan treatment or disposition: 65.5(29.74)

Table 29

Confidence in Diagnosis

Confidence in Diagnosis Provided	**Mean**	**p-value**
Role		**p<0.01**
Attendings	67.5(25.9)	
Residents and Fellows	68.6(23.2)	
APPs	75.5(18.9)	
Location		**p = 0.70**
UMB	69.8(23.1)	
UTH	69.2(24.6)	
Case		**p <0.01**
Controls	76.9(20.1)	
Stimuli	56.5(24.4)	
Chief Complaint – Overall		**p <0.01**
1. Jones Fracture (End)	92.9(8.96)	
2. MVC- Fractured ulna (Late)	85.3(12.0)	
3. DKA+UTI (Middle)	78.9(15.4)	
4. AMS-Alcohol-Hyponatremia (Late)	77.3(16.2)	
5. Stroke (Middle)	74.9(19.5)	
6. Pneumothorax (Early)	70.2(16.5)	
7. Cardiac – Stimuli	60.4(24.2)	
8. Lung – Stimuli	60.3(23.1)	
9. Appendicitis (Early)	58.9(26.5)	
10. Head – Stimuli	55.9(25.7)	
11. Abdomen – Stimuli	49.4(23.5)	
Chief Complaint – Controls		**p <0.01**
1. Jones Fracture	92.9(8.96)	
2. MVC- Fractured ulna	85.3(12.0)	
3. DKA +UTI (Middle)	78.9(15.4)	
4. AMS-Alcohol-Hyponatremia	77.3(16.2)	
5. Stroke (Middle)	74.9(19.5)	
6. Pneumothorax (Early)	70.2(16.5)	
7. Appendicitis (Early)	58.9(26.5)	
Chief Complaint – Stimuli		**p = 0.03**
1. Cardiac	60.4(24.2)	
2. Lung	60.3(23.1)	
3. Head	55.9(25.7)	
4. Abdomen	49.4(23.5)	
Stimuli – Explicitness		**p = 0.61**
Explicit	57.3(24.2)	
Implicit	55.8(24.7)	
Stimuli – Length		**p <0.01**
Long	64.1(21.9)	
Short	48.9(24.6)	

Table 30

Confidence in Considering the Scope of the Differential Diagnoses

Confidence in Considering the Scope of the Differential Diagnosis	**Mean**	**p-value**
Role		**p = 0.29**
Attendings	77.1(21.9)	
Residents and Fellows	74.4(19.3)	
APPs	77.3(20.0)	
Location		**p = 0.56**
UMB	76.7(20.5)	
UTH	75.9(20.8)	
Case		**P<0.01**
Controls	80.8(18.5)	
Stimuli	68.4(21.9)	
Chief Complaint – Overall		**p<0.01**
1. Jones Fracture (End)	91.8(12.5)	
2. MVC- Fractured ulna (Late)	85.6(14.4)	
3. AMS-Alcohol Hyponatremia (Late)	80.1(16.1)	
4. Stroke (Middle)	79.1(19.3)	
5. DKA + UTI (Middle)	79.1(17.4)	
6. Pneumothorax (Early)	78.2(19.0)	
7. Lung -Stimuli	74.4(17.0)	
8. Appendicitis (Early)	72.0(22.9)	
9. Cardiac- Stimuli	71.3(23.2)	
10. Head -Stimuli	65.5(22.8)	
11. Abdomen – Stimuli	62.3(22.1)	
Chief Complaint – Controls		**p<0.01**
1. Jones Fracture	91.8(12.5)	
2. MVC – Fractured ulna	85.6(14.4)	
3. AMS-Alcohol Hyponatremia	80.1(16.1)	
4. Stroke (Middle)	79.1(19.3)	
5. DKA +UTI (Middle)	79.1(17.4)	
6. Pneumothorax (Early)	78.2(19.0)	
7. Appendicitis (Early)	72.0(22.9)	
Chief Complaint – Stimuli		**p = 0.17**
1. Lung	74.4(17.0)	
2. Cardiac	71.3(23.2)	
3. Head	65.5(22.8)	
4. Abdomen	62.3(22.1)	
Stimuli – Explicitness		**p = 0.52**
Explicit	70.6(20.0)	
Implicit	66.1(23.4)	
Stimuli – Length		**p <0.01**
Long	73.3(19.1)	
Short	63.4(23.4)	

Table 31

Readiness to Plan Treatment or Disposition

Ready to Plan Treatment or Disposition	Mean	p-value
Role		**P = 0.03**
Attendings	62.4(32.5)	
Residents and Fellows	67.9(26.9)	
APPs	68.3(27.3)	
Location		
UMB	67.0(28.9)	**p = 0.20**
UTH	64.2(30.5)	
Case		**p<0.01**
Controls	72.7(27.3)	
Stimuli	52.9(29.7)	
Chief Complaint – Overall		**p<0.01**
1. Jones Fracture (End)	94.8(9.84)	
2. MVC- Fractured ulna (Late)	84.3(18.2)	
3. AMS-Alcohol Hyponatremia (Late)	80.6(22.0)	
4. DKA + UTI (Middle)	73.3(24.5)	
5. Stroke (Middle)	68.0(27.5)	
6. Stimuli – Head	58.7(29.2)	
7. Pneumothorax (Early)	58.2(26.1)	
8. Stimuli – Cardiac	53.6(32.8)	
9. Stimuli – Lung	52.1(28.3)	
10. Appendicitis (Early)	49.9(29.4)	
11. Stimuli – Abdomen	47.4(27.9)	
Chief Complaint – Controls		**p<0.01**
1. Jones Fracture	94.8(9.84)	
2. MVC- Fractured ulna	84.3(18.2)	
3. AMS-Alcohol Hyponatremia	80.6(22.0)	
4. DKA+UTI (Middle)	73.3(24.5)	
5. Stroke (Middle)	68.0(27.5)	
6. Pneumothorax (Early)	58.2(26.1)	
7. Appendicitis (Early)	49.9(29.4)	
Chief Complaint – Stimuli		**p = 0.17**
1. Head	58.7(29.2)	
2. Cardiac	53.6(32.8)	
3. Lung	52.1(28.3)	
4. Abdomen	47.4(27.9)	
Stimuli – Explicitness		**p = 0.52**
Explicit	54.1(29.7)	
Implicit	51.5(29.9)	
Stimuli – Length		**p<0.01**
Long	61.4(28.0)	
Short	44.5(29.1)	

7.5.4 Interpretation of Analyses Confidence in Diagnosis

Finding: The variation in participants' confidence in their diagnosis was not statistically significant when compared by location.

Finding: Variation in participants' confidence in their diagnosis was statistically significant when compared for the role of the participants and also when compared for the type of case.

Finding: Variation in participants' confidence in their diagnosis was found to be statistically significant for cases when compared by chief complaint. This was the case for all the cases overall and for the control cases only as a group, as well as for the stimuli cases only as a group.

To explore this finding further the results for the participants' confidence in their diagnosis was analyzed for any potential effects related to the test conditions of explicitness of the information and for length of the case.

Finding: Differences in participants' confidence in their diagnosis was found to not be statistically significant when compared by the explicitness of the case information.

Finding: Differences in participants' confidence in their diagnosis was statistically significant when comparing by the length of the stimuli case. So, the length of the case was associated with differences in participant's confidence in their diagnosis with long cases having higher levels of confidence than short cases.

7.5.5 Confidence in considering the scope of Differential Diagnoses

Finding: Variation in participants' confidence in their ability to consider their scope of differential diagnoses was not found to be statistically significant for either the role or the location of the participants.

Finding: The variation in participants' confidence in their ability to consider their scope of differential diagnoses was statistically significant for the type of case. The confidence levels were much higher for control cases at around 81% level of confidence compared to stimuli cases at 68%.

Finding: Variations in the levels of participants' confidence when considering the scope of differential diagnoses with the information provided, were found to be statistically significant for chief complaint in cases overall. In addition, differences in their level of confidence were also found to be statistically significant by chief complaint in control cases as a group. Similarly, statistically significant differences for confidence in considering the scope of the differential diagnoses were also observed in stimuli cases as a group.

To explore this further the confidence levels for considering the scope of differential diagnosis were analyzed for the test conditions of explicitness and length of case for the stimuli cases. The results were as follows:

Finding: Variations in the levels of participants' confidence for the scope of differential diagnosis were not statistically significant in stimuli cases when compared by explicitness.

Finding: Variations in the levels of participants' confidence when considering the scope of differential diagnosis with the information provided was found to be statistically significant for in stimuli cases depending on length of the case with long cases associated with higher levels of confidence.

7.5.6 Readiness to Plan Treatment or Disposition

The results for the analyses of participants' readiness to plan treatment or arranging their disposition given the information provided for each case were as follows:

Finding: Variation in participants' readiness to plan treatment or disposition was not statistically significant for either the role of the location of the participants.

Finding: The variation in participants' readiness to plan treatment or disposition with the information provided was found to be statistically significant for the type of case.

Finding: Variations in the levels of participants' readiness to plan treatment or disposition with the information provided was found to be statistically significant for chief complaint in cases overall as well as for chief complaint in the control cases as a group

Finding: Variations in the levels of participants' readiness to plan treatment or disposition was not found to be statistically significant in stimuli cases as a group when compared by chief complaint.

To explore whether the test conditions of explicitness and length of case for the stimuli cases had any effect on participants' readiness to plan treatment or disposition the results were analyzed:

Finding: Variations in the levels of participants' readiness to plan treatment or disposition were not statistically significant in stimuli cases by explicitness.

Finding: Variations in the levels of participants' readiness to plan treatment or disposition with the information provided were found to be statistically significant for stimuli cases by length of the case.

7.6 Participants' Response for Key Data Match to Data in Sign Out

Participants' responses to question 3: what key data participants identified as important to their decision making from the sign out information provided was compared for matching with the details in the sign out narrative and represented as *Key Data Match*. Results for the measure *Key Data Match* were analyzed by comparing the various study parameters. The results showed the following findings (see table 32):

7.6.1 Interpretation of Analyses for Key Data Match

Finding: The variation in whether participants' key data matched to the sign out narrative was found to not be statistically significant when compared by either the location or by the type of role of the participants.

Finding: The variation in whether participants' key data matched to the sign out narrative was found to be statistically significant when compared by the type of case i.e., whether control cases or stimuli cases.

Finding: The variation in whether participants' key data matched to the sign out was found to be statistically significant when compared by chief complaint across all the cases overall and also for control cases as a group compared by chief complaint but was

found to not be statistically significant when compared by chief complaint across the stimuli cases only as a group.

Further analyses were conducted to determine the effect of the test conditions of explicitness and length on the measure of *Key Data Match*:

Finding: The variation in *Key Data Match* was found to not be statistically significant for explicitness.

Finding: The variation in *Key Data Match* was found to be statistically significant when compared for the test condition of length

Table 32

Key Data Match

Key Data in Sign Out	Yes	No	p-value
Role			**p = 0.84**
Attendings	80%	20%	
Residents and Fellows	81%	19%	
APPs	82%	18%	
Location			**p = 0.04**
UMB	82%	18%	
UTH	81%	19%	
Case			**p<0.01**
Controls	88%	12%	
Stimuli	69%	31%	
Chief Complaint – Overall			**p<0.01**
1. Jones fracture (End)	97%	3%	
2. DKA (Middle)	96%	4%	
3. AMS Alcohol Hyponatremia	93%	7%	
4. MVC – Fractured ulna (Late)	91%	9%	
5. Pneumothorax (Early)	88%	12%	
6. Stroke (Middle)	83%	17%	
7. Lung – Stimuli	72%	28%	
8. Cardiac – Stimuli	71%	29%	
9. Head – Stimuli	71%	29%	
10. Appendicitis (Early)	70%	30%	
11. Abdomen – Stimuli	61%	39%	
Chief Complaint – Controls			**p<0.01**
1. Jones fracture (End)	97%	3%	
2. DKA (Middle)	96%	4%	
3. AMS Alcohol Hyponatremia	93%	7%	
4. MVC Fractured ulna (Late)	91%	9%	
5. Pneumothorax (Early)	88%	12%	
6. Stroke (Middle)	83%	17%	
7. Appendicitis (Early)	70%	30%	
Chief Complaint – Stimuli			**p = 0.44**
1. Lung	72%	28%	
2. Cardiac	71%	29%	
3. Head	71%	29%	
4. Abdomen	61%	39%	
Stimuli – Explicitness			**p = 0.93**
Explicit	68%	32%	
Implicit	69%	31%	
Stimuli – Length			**p = 0.01**
Long	76%	24%	
Short	61%	39%	

Table 33

Next Steps Match

Next Steps Consistent with Diagnosis	**Yes**	**No**	**p-value**
Role			**p = 0.06**
Attendings	91%	9%	
Residents and Fellows	88%	12%	
APPs	84%	16%	
Location			**p = 0.04**
UMB	91%	9%	
UTH	86%	14%	
Case			**p<0.01**
Controls	96%	4%	
Stimuli	57%	43%	
Chief Complaint – Overall			**p<0.01**
Pneumothorax (Early)	100%	0%	
Appendicitis (Early)	99%	1%	
Stroke (Middle)	99%	1%	
MVC – Fractured ulna (Late)	97%	3%	
Cardiac – Stimuli	97%	3%	
Jones fracture (End)	96%	4%	
DKA (Middle)	93%	7%	
Lung – Stimuli	90%	10%	
AMS Alcohol Hyponatremia (Late)	84%	16%	
Abdomen – Stimuli	62%	38%	
Head – Stimuli	62%	38%	
Chief Complaint – Controls			**p<0.01**
Pneumothorax (Early)	100%	0%	
Appendicitis (Early)	99%	1%	
Stroke (Middle)	99%	1%	
MVC-Fractured ulna (Late)	97%	3%	
Jones fracture (End)	96%	4%	
DKA (Middle)	93%	7%	
AMS alcohol Hyponatremia (Late)	84%	16%	
Chief Complaint – Stimuli			**p<0.01**
Cardiac	97%	3%	
Lung	90%	10%	
Abdomen	62%	38%	
Head	62%	38%	
Stimuli – Explicitness			**p= 0.77**
Explicit	79%	21%	
Implicit	77%	23%	
Stimuli – Length			**p = 0.77**
Long	79%	21%	
Short	76%	24%	

7.7 Participants' Response for Next Steps for the Case

Participants' responses to question 4: what their next steps would be for the case given the information provided in the sign out compared to correct next steps according to the reference texts. Results for the measure Next Steps Match were analyzed by comparing the various study parameters. The results showed the following findings (see table 33):

7.7.1 Interpretation of Analyses for Next Steps

Finding: Results show that variations in whether participants' responses matched the next steps specified in medical reference texts expressed as the measure Next Steps Match was not found to be statistically significant when compared by the role of the participant.

Finding: Participants' responses for Next Steps Match did show a slightly statistically significant difference when compared for the two locations of UMB and UTH.

Finding: The variation in whether participant specified next steps matched those in medical reference texts was statistically significant for the type of case. Control cases had a 96% of matching whereas stimuli cases matched at 57%.

Finding: Variations in Next Steps Match for responses by participants were found to be statistically significant by chief complaint for all the cases. The differences in Next Steps Match were also statistically significant for control cases when compared by chief complaint as well as for stimuli cases as a group when compared by chief complaint.

This effect was further investigated to determine whether the test conditions of explicitness of the information or the length of the case had any influence on the measure of Next Steps Match.

Finding: Neither of the test conditions of length of case nor the explicitness of sign out information had any statistically significant effect on the variations observed in the measure Next Steps Match, which equated to whether the next steps specified by participants matched the appropriate steps specified in the medical reference text for the stimuli cases.

7.8 Interaction Effects Analysis for Stimuli Cases: Abdomen and Head Cases

To better understand the factor that may be at play for the two stimuli cases appearing to have variation in the data, further analyses were conducted. These two stimuli cases were the abdomen and the head cases. Each case was analyzed for effects in diagnostic accuracy, confidence in diagnosis, count of the number of diagnoses in the differential and participants' readiness to plan treatment.

7.8.1 Interaction Effects of Abdomen Case: Upper Gastrointestinal Bleed

Table 34

Diagnostic Accuracy by Explicitness – Abdomen Case

Diagnosis Accurate	Explicit	Implicit
No		
n	28	26
Overall %	41%	38%
Row %	52%	48%
Column %	78%	79%
Yes		
n	8	7
Overall %	12%	10%
Row %	53%	47%
Column %	22%	21%

N = 69 responses Pearson χ^2 = 5.8545e-32, df = 1, p-value = 1.0

Finding: Variations in the levels of participants' diagnostic accuracy between explicit and implicit abdominal cases was found to be not statistically significant.

Table 35

Diagnostic Accuracy by Case Length – Abdomen Case

Diagnosis Accurate	Long	Short
No		
n	21	33
Overall %	30%	48%
Row %	39%	61%
Column %	60%	97%
Yes		
n	14	1
Overall %	20%	1%
Row %	93%	7%
Column %	40%	3%

N = 69 responses Fisher's Exact p-value <0.01

Finding: Variations in the levels of participants' diagnostic accuracy for abdominal cases by length was found to be statistically significant when comparing long and short cases.

Table 36

No. of Diagnoses for Abdomen Cases by Explicitness

Case	Mean (SD)
Explicit	3.1(1.40)
Implicit	3.0(1.31)

Two sample t-test: t = 0.34223, df = 66, **p-value =** 0.73

Table 37

No. of Diagnoses for Abdomen Cases by Length

Case	Mean (SD)
Long	2.5(1.08)
Short	3.7(1.31)

Two sample t- test t = -4.2335, df = 66 **p-value <** 0.01

Finding: Variations in the number of diagnoses in the differential was not statistically significant in stimuli abdomen cases by explicitness.

Finding: Variations in the number of diagnoses in the differential was statistically significant in stimuli abdomen cases when compared by length of the case.

Table 38

Confidence in Diagnosis for Abdomen Case by Explicitness

Case	**Mean (SD)**
Explicit	49.2(22.3)
Implicit	49.6(25.0)

Two sample t- test t = -0.67766, df =7, **p-value** = 0.95

Table 39

Confidence in Diagnosis for Abdomen Case by Length

Case	**Mean (SD)**
Long	59.6(21.4)
Short	38.9(21.0)

Two sample t-test: t = 4.052, df = 67, **p-value** <0.01

Finding: Variations in the number of diagnoses in the differential was not statistically significant in stimuli abdomen cases by explicitness.

Finding: Variations in the number of diagnoses in the differential was statistically significant in stimuli abdomen cases when compared by length of the case.

Table 40

Readiness to Plan Treatment Abdomen Case by Explicitness

Case	Mean (SD)
Explicit	46.6(27.4)
Implicit	48.2(29.0)

Two sample t-test: t = -0.23643, df = 67, **p-value** = 0.81

Table 41

Readiness to Plan Treatment Abdomen Case by Length

Case	Mean (SD)
Long	59.4(26.0)
Short	35.1(24.6)

Two sample t-test: t = 3.9835, df = 67 **p -value** <0.01

Finding: Variations in the number of diagnoses in the differential was not statistically significant in stimuli abdomen cases by explicitness.

Finding: Variations in the number of diagnoses in the differential was statistically significant in stimuli abdomen cases when compared by length of the case.

Table 42

Summary Details from Abdomen Case

Diagnosis	N	Long	Short
Correct	15	14	1
Incorrect	54	21	33
Cardiac related diagnoses	22	4	18
Pulmonary related diagnoses	16	3	13

7.9.2 Interaction Effects of Head Case: UTI + Decubitus Ulcer

The head case caused particular difficulty for participants in terms of providing an accurate diagnosis. Thus, the case was further analyzed to determine other potential factors that may have been in play for this case.

Table 43

Diagnostic Accuracy by Explicitness – Head Case

Diagnosis Accurate	**Explicit**	**Implicit**
No		
n	32	36
Overall %	46%	52%
Row %	47%	53%
Column %	97%	100%
Yes		
n	1	0
Overall %	2%	0%
Row %	100%	0%
Column %	3%	0%

N = 69 responses Fisher's Exact p-value = 0.48

Finding: Variations in the levels of participants' diagnostic accuracy between explicit and implicit head cases was found to not to be statistically significant.

Table 44

Diagnostic Accuracy by Case Length – Head Case

Diagnosis Accurate	**Long**	**Short**
No		
n	33	35
Overall %	48%	51%
Row %	49%	51%
Column %	97%	100%
Yes		
n	1	0
Overall %	1%	0%
Row %	100%	0%
Column %	3%	0%

N = 69 responses Fisher's Exact p-value = 0.49

Finding: Variations in the levels of participants' diagnostic accuracy between long and short head cases was found to not to be statistically significant.

Table 45

No. of Diagnoses for Head Case by Explicitness

Case	**Mean (SD)**
Explicit	3.5(1.92)
Implicit	3.9(1.73)

Two sample t-test: t = -0.98215, df = 67, **p-value** = 0.33

Table 46

No. of Diagnoses for Head Case by Length

Case	**Mean (SD)**
Long	3.3(1.63)
Short	4.1(1.95)

Two sample t-test: t = = -1.7621, df = 67, **p-value** = 0.08

Finding: Variations in the number of diagnoses in the differential was not statistically significant in stimuli head cases when compared by the explicitness of the information or by length of the case.

Table 47

Confidence in Diagnosis for Head Case by Explicitness

Case	**Mean (SD)**
Explicit	60.5(25.2)
Implicit	51.6(25.8)

Two sample t-test: t = = 1.4497, df = 67, **p-value** = 0.15

Table 48

Confidence in Diagnosis for Case Head by Length

Case	**Mean (SD)**
Long	61.8(21.3)
Short	50.2(28.5)

Two sample t-test: t = 1.9046, df = 67, **p-value** = 0.06

Finding: Variations in participants' confidence in diagnosis were not statistically significant in stimuli head cases when compared either by the explicitness of the information or the by length of the case

Table 49

Readiness to Plan Treatment for Head Cases by Explicitness

Case	**Mean (SD)**
Explicit	63.5(29.4)
Implicit	54.2(28.6)

Two sample t-test: t = 1.3344, df = 67, **p-value** = 0.18

Table 50

Readiness to Plan Treatment for Head Cases by Length

Case	**Mean (SD)**
Long	68.5(24.0)
Short	49.1(30.8)

Two sample t-test: t = 2.9039, df = 67, **p-value** <0.01

Finding: Variations in participants' readiness to plan treatment or disposition were not statistically significant in stimuli head cases when compared for explicitness of the information.

Finding: Variations in participants' readiness to plan treatment or disposition was statistically significant in stimuli head cases when compared by length of the case.

Table 51

Summary Details from Head Cases with Incorrect Diagnosis

Diagnosis		**Count**	**Proportion**
UTI		29	43%
Head injury/event		14	20%
Other infection		14	20%
Non-specific Diagnosis		12	17%
Correct for Stage of Head case		48	70%

Table 52

Data Analysis Summary Matrix

Parameter / Measure	X= not significant, p value = where significant							
	Location	**Role**	**Case**	**Chief Complaint Overall**	**Chief Complaint Controls**	**Chief Complaint Stimuli**	**Explicitness**	**Length**
Identified the Stage of Case	**X**	**X**	**X**	--	--	--	--	--
Diagnosis was Accurate	**X**	**X**	**p<0.01**	**p<0.01**	**p<0.01**	**p<0.01**	**X**	**p<0.01**
Diagnosis included in Differential	**X**	**X**	**p<0.01**	**p<0.01**	**p=0.01**	**p<0.01**	**X**	**p<0.01**
More than One Diagnosis	**X**	**p=0.01**	**p<0.01**	**p<0.01**	**p=0.01**	**X**	**X**	**X**
No. of Diagnoses in Differential	**X**	**p<0.01**	**p<0.01**	**p<0.01**	**p<0.01**	**p<0.01**	**X**	**p<0.01**
Confidence in Diagnosis	**X**	**p<0.01**	**p<0.01**	**p<0.01**	**p<0.01**	**p=0.03**	**X**	**p<0.01**
Confidence in diagnostic scope	**X**	**X**	**p<0.01**	**p<0.01**	**p<0.01**	**p<0.01**	**X**	**p<0.01**
Ready to plan treatment or disposition	**X**	**p=0.02**	**p<0.01**	**p<0.01**	**p<0.01**	**X**	**X**	**p<0.01**
Key Data Matches Sign Out	**X**	**X**	**p<0.01**	**p<0.01**	**p<0.01**	**X**	**X**	**P=0.01**
Next Steps Match with Diagnosis	**p=0.04**	**X**	**p<0.01**	**p<0.01**	**p<0.01**	**p<0.01**	**X**	**X**

CHAPTER 8: EDSO DISCUSSION, RECOMMENDATIONS AND LIMITATIONS

8.1 Discussion of Initial Findings

The EDSO survey intended to assess the affect that the two test conditions of explicitness of the information and length of the case had on the diagnostic accuracy and reasoning of EM clinicians when a series of patient cases were transferred in the form of a sign out communications. The test conditions were related to the volume of information and the experiment was designed to determine the impact of varying the amount of information contained in the sign out narrative. The analysis of the response data yielded findings about the primary outcomes as well as providing some insight about several secondary outcomes.

8.1.1 Demographic Differences

Firstly, initial analysis of the data was conducted to better understand the specific demographic details of the participant population. The data collection for the study from the two sites resulted in 35 participants from UTH and 34 from UMB, giving a total of 69 participants who completed the survey. While the number of responses from the two sites appeared similar, analysis of the survey data revealed the population of participants at the two sites were quite different in terms of the participants roles. (See table 22) The responses from UMB had a far higher proportion of attending physicians than UTH and indeed of the attending physician participant group, 74% were from UMB and 26% were from UTH. The respondents who were trainees, that is either a resident or a fellow, were

more evenly balanced for the two sites with 42% from UMB and 58% from UTH. However, of the APPs group, all but one, were from UTH. In terms of the respondents by site, UMB had proportionally more attendings (68%) than residents (29%). Whereas with UTH, there was more of an even distribution for the roles with 23% as attendings 40% as residents and 37% as APPs.

This suggests that there may have been an experience gradient within the study population, as UMB participants had an average of 8.5 years of experience compared to the average of 5.1 years for UTH participants. Also, UMB had only one APP complete the survey and so this group could effectively be considered as under-represented in the UMB participant population.

The main distinguishing factor between the two sites was that at UMB sign outs are conducted using a checklist-based handoff protocol, (see Appendix A) whereas at UTH routine use of protocols at sign out is not required and is not consistently practiced. Despite this major difference in terms of protocol use, analysis of the data revealed no significant differences for various study outcomes except for the *Next Steps Planning* for the case, when compared for Location. Hence, it was concluded that any effects observed in the data could not be attributed to site specific differences and therefore could not be attributed to the use of handoff protocols. The influence of participant role and experience will be discussed pertaining to the outcomes where significant differences were observed.

8.2 Analysis of Study Outcomes

The main study outcomes were analyzed to determine the influence of the test conditions that were included in the experiment. The test conditions were length of the case and explicitness of the information. Additionally, the experimental design was blocked for chief complaint in the stimuli cases due to its potential propensity to drive the clinical reasoning in EM clinicians when they considered a case.

8.2.1 The Effect of Test Conditions of Length and Explicitness

The primary outcomes for this study were how accurately the participants would be able to diagnose cases that include the test conditions of length and explicitness. In addition, the analysis sought to determine how participants felt about the information for their diagnostic and treatment planning decision making in terms of the measures of participants' confidence in their diagnosis, their confidence in considering the scope of their differential diagnoses, and their readiness to plan treatment or disposition of the case.

Hypothesis 1 was that the explicitness of sign out information would affect whether participants provide accurate diagnoses for vignette patient cases presented as either implicit or explicit cases.

The first test condition in this study was the explicitness of the sign out information in stimuli cases, which equated to how unpacked the information was and as such reflected the number of pieces of information in the sign out. The hypothesis was that the explicitness of the case information would influence the decision making of the recipient

of the sign out information with more explicit cases associated with higher diagnostic accuracy results than implicit cases.

The analysis findings indicated that the level of explicitness of the information in the case did not have any significant effect on the results for diagnostic accuracy for the stimuli cases ($p=0.55$). Explicitness did not affect diagnostic accuracy for the cases regardless of the type of stimuli case, in terms of whether it related to a cardiac, or lung, or abdominal or head type of chief complaint.

Additionally, none of the other measures related to the diagnostic differential for the cases provided by the participant were found to be statistically significant related to the explicitness of the case. So, whether the correct diagnosis was included in the differential, whether one or more diagnoses were part of the differential and the number of diagnoses in the differential, did not show any significant effects related to the implicit or explicit nature of the sign out information. From this it was concluded that explicitness, as it related to the number of information pieces did not affect providing a diagnosis or the reasoning for the differential diagnoses of participants in this study.

Hypothesis 2 was that the length of the sign out case would affect whether participants provide accurate diagnoses for vignette patient cases presented as either long (late) or short (earlier/middle) cases at sign out.

The second test condition was the length of the stimuli cases, which was equivalent to the stage in the diagnostic process for the patient in the case. The hypothesis was that the length of the case would influence the diagnostic accuracy for the case with long cases associated with improved diagnostic accuracy than short cases. The results of the survey

indicated that the length of a case, i.e., the stage at the time of sign out, did have a significant effect on diagnostic accuracy with long cases demonstrating higher diagnostic accuracy (60%) than short cases (30%). As long cases were associated with significantly higher levels of diagnostic accuracy, it was concluded that the length of a case and thereby the stage a case was in, did affect the accuracy of diagnostic decisions. Additionally, the other measures related to the diagnostic differential for the cases provided by the participants were found to be statistically significant related to the length of the case for two out of three measures. For the measures of whether the correct diagnosis was included in the differential and the count of diagnoses in the differential, the results demonstrated long cases had significantly different results than short cases. From this it was concluded that length as it related to the stage in the case and thus the amount of information for a case, significantly affected participants' ability to provide accurate diagnoses.

8.2.2 Additional Analyses

Further analyses investigated whether the explicitness of case information provided in a sign out scenario would affect how participants considered the sign out information for their diagnostic decision making and treatment planning in terms of:

i. their confidence in their diagnosis for the case,
ii. their confidence in considering the scope of their differential diagnoses, or
iii. their readiness to plan treatment or disposition for the case.

Additional primary outcomes of this study were related to how confident participants felt about the information provided in the case. The variables of participants' confidence in

their diagnosis, their confidence in considering the scope of their differential diagnoses and for their readiness to plan treatment or disposition for the case, were all measures related to how useful participants considered the information in the case for their decision making. (See tables 30-32)

The null hypothesis was that participants' consideration of the information in the case would not be affected by the explicitness of the information. The statistical analyses failed to reject the null hypothesis and showed that explicitness had no effect on any of the three measures related to participants consideration of the information in the case. Hence it was concluded that explicitness of sign out information had no influence on how participants felt about whether the information supported their reasoning; Confidence in their diagnosis ($p=0.61$), confidence in considering the diagnostic scope of the case ($p=0.52$), and their readiness for treatment planning ($p=0.52$).

Additionally, further analyses investigated whether the length of the case provided in a sign out scenario affected participants considered the sign out information for their diagnostic decision making and treatment planning in terms of:

i. their confidence in their diagnosis for the case,
ii. their confidence in considering the scope of their differential diagnoses, or
iii. their readiness to plan treatment or disposition for the case.

Similarly, further outcomes were related to how confident participants felt about the information provided in the case based on the length of the case. The variables of participants' confidence in their diagnosis for a case, their confidence in considering the scope of their differential diagnoses and for their readiness to plan treatment or

disposition for the case represented how useful the participants felt the information was in supporting their decision making. (See tables 30-32)

The null hypothesis was that whether a case was long or short would not affect how useful participants felt the information in the case was. The results indicated that the length of the case was significant for all three measures related to how the information supported the participants decision making. So, the null hypothesis was rejected, and it was concluded that the long case and thereby a later stage and more information, was highly influential in supporting the diagnostic decisions making considering the scope of the differential and treatment planning in participants.

8.2.3 Discussion

The rationale for the experiment was the length of a case equated to how far along the diagnostic process a case was, which would in turn determine the amount of information presented for the case at sign out. The term long case was a construct that denoted a case equivalent to the late stage observed in the control cases, while the term short was equivalent to a case in earlier stage. The other test variable was explicitness of the information provided, which proposed that information communicated explicitly contains a greater volume of information than if conveyed implicitly.

In terms of information volume for a case, while explicitness may represent the number of pieces of data associated with the information item, length determined whether the information item is even available, depending on when the cut off occurs in the form of a sign out. Long cases have progressed further allowing for more test results, imaging or consults to be completed and available. From the results of the analysis, we could see that

while explicitness had no effect on diagnostic decision making, the length of the case and thereby the amount of information presented at sign out did strongly influence the ability to accurately identify the diagnosis.

It could be argued that in this experiment participants were provided a static snapshot of the case. This did not allow for the interactive dynamic nature of the diagnostic process, as it occurs in reality, to inform their reasoning process when developing a diagnosis for the case. However, the diagnostic differential is the primary driver for this process, which involves ruling in or ruling out conditions as the test results for the patient come back. While the patient's response to initial treatments also feeds into the reasoning process, many of those initial treatments are also still driven by that initial presumptive differential. The results for the measure of number of diagnoses in the differential (see table 29) do indicate that short cases have a broader mean differential (3.4(1.59)) than long cases where the differential is narrowed (2.7(1.36)) as the case progresses, in keeping with the whittling down process of developing a diagnosis. However, the measure for whether the correct diagnosis was included in the differential (see table27) did show that for a short/earlier case, the correct diagnosis was included in the differential much less frequently (44%) than in long cases (70%). This suggests that stage of the case at the time of sign out poses a potential risk factor for how the recipient of the sign out uses the information to develop their diagnostic differential. This assessment might thereby influence the potential tests and treatments ordered and thereby set the diagnostic path. This is particularly so if the stage the patient is in is not explicitly highlighted during the sign out. Indeed, the results show (see table 23) that the recipients

of the sign out in this study were not able to accurately identify the stage in a large proportion of cases (35% for controls and 41% for stimuli cases) from the sign out narrative alone.

8.2.4 Diagnostic Accuracy and Stage of the Case

The results of the analyses of the relationship between accurately diagnosing the case and correctly identifying the case did show statistically significant differences. What was most interesting was that those who did not accurately diagnose the case were more likely to get the stage correct, whereas those who did get the correct diagnosis showed no significant ability to identify the stage with roughly equal split for correct and incorrect answer for stage. This potential inverse relationship in the responses for these two measures suggests the occurrence of premature closure. So those who stated the case was a late case did so because it aligned with their expectation that the diagnostic process was over, and they had reached their diagnosis. However, their assessment was incorrect. Review of the response data at item level (see table 26) revealed that of the cases where participants had provided an incorrect diagnosis and had incorrectly identified the stage, the majority of the time, 88% (50/57 instances), they had identified a case as being a late stage when it was an earlier stage case, compared to the 12% who had identified a late stage case as an earlier stage (7/57 instances). This suggests that premature closure may be a contributing factor in a large proportion of cases of incorrect diagnosis where the case is at the earlier stage of the diagnostic process at the time of sign out.

8.3 Chief Complaint

The study design included blocking for the chief complaint because findings from the interview study in Aim 1 and consultation with physician SMEs indicated that chief complaint influenced the diagnostic process. The significant finding from this study was that certain types of chief complaints did have a strong influence not just on accuracy of diagnosis but also on the confidence the clinician had in their diagnostic reasoning. The chief complaints that demonstrated the most differences were those involving the head and abdomen, both in stimuli and control cases.

8.3.1 For Control Cases

Three cases stood out from the set of control cases; the case of altered mental state (AMS) due to alcohol related hyponatremia case, the appendicitis case, and the pneumothorax cases. For the AMS Hyponatremia, which was a late stage case, the diagnostic accuracy level (see table 24) was lowest of the control cases (80%). In contrast, for the pneumothorax and appendicitis cases, which were both early cases, the diagnostic accuracy was much higher at 96% and 91% respectively. However, participants' confidence level in their diagnosis did not correlate with the diagnostic accuracy observed (see table 29). Participants' levels of confidence in their diagnosis in the AMS Hyponatremia case was 4th highest (77%), whereas for the pneumothorax case it was lower at 6th place (70%) and the appendicitis case at 9th place (58.9%). A possible explanation could be that both the appendicitis and pneumothorax cases were early cases. While these control cases were overall more straightforward, the cause of illness may have appeared more ambiguous due to their early stage of presentation. This may be why

we see the lower confidence in the diagnosis for these two cases despite the high accuracy in correctly identifying the diagnosis in each case. Studies have demonstrated that clinicians' confidence in their diagnosis does not always correspond to their diagnostic accuracy (Meyer et al., 2013)

With the AMS case, many of the responses for diagnosis that were incorrect listed alcohol withdrawal rather than hyponatremia, despite clear information on the salt level of the patient at admission. The participants in this case may have focused on the report that the patient was seen shaking by the police who called EMS, which was mentioned in the sign out narrative. Participants appear to have settled on the possible earlier seizure that occurred, despite there being no mention of any indication of seizures, a symptom of alcohol withdrawal, since the patient's arrival in the ED. This suggests that even within the controls, a case that has a chief complaint related to the head or AMS can prove challenging for clinicians to diagnose. They may have focused on ruling out life threatening complications like potential seizures because the consequences of missing something in a head case are far more serious than for a trauma case involving a fractured toe like the Jones fracture case. In this hyponatremia head case, many seemed to miss the extremely low salt levels in the blood, which has a relatively clear path of treatment but one that is different from that for a case of alcohol withdrawal.

8.3.2 For Stimuli Cases

Across the four stimuli cases a pattern of grouping was observed where the two cardiac and lung cases had similar levels consistently for most measures. The cardiac and lung cases typically had high levels for diagnostic accuracy (see table 24) and narrower

differential diagnoses (table 29) and higher levels of confidence in the diagnosis also (see table 30). Meanwhile the abdominal case and particularly the head case had significantly lower levels of diagnostic accuracy and confidence in diagnosis. All the stimuli cases were designed for similar levels of complexity and representativeness in terms of the chief complaints that typically present in the ED (McCaig & Nawar, 2006) The results showed that diagnostic decision making and clinical impression was statistically significant for these stimuli cases and the abdomen and head cases seemed to elicit very different responses from participants than for the other cases. In addition, the stimuli cases showed significantly better diagnostic accuracy, narrower differentials, and higher confidence levels for all the measures for long case over short cases

However, the results for the head and abdomen cases were so markedly different from the cardiac and pulmonary cases that closer examination of these two cases was conducted. While overall much lower than the cardiac and lung cases, the abdominal case and the head case showed marked differences in diagnostic accuracy as well as the other parameters.

Abdomen Case – The abdominal case involved a patient who presented with chest pain and shortness of breath shortly after a fishing vacation with his friends, when he had consumed higher amounts of alcohol than usual throughout the trip. While his EKG did show sinus tachycardia, he did not have any signs of cardiac dysrhythmia. The patient's blood tests were normal, except for a drop in hemoglobin levels which was mentioned in the longer cases. While this was a case of an upper gastro-intestinal bleed, this diagnosis was missed in the majority of instances (77%) (See table 24). Few participants specified a

rectal exam as part of the next steps, which would have indicated that they were considering some form of abdominal issue. Indeed, participants misdiagnosed the case most commonly as a potential cardiac complaint, while others as a lung case, e.g., pulmonary embolism. So, these incorrect diagnoses were not accurate even for the correct chief complaint type, which was abdomen.

Head Case – this case proved to be a very difficult case for participants to diagnose, with a 99% incorrect diagnosis rate. The case was of an elderly lady with altered mental state (AMS). She had a fall 3 days earlier but with no injury or neuro-focal deficits. She presented in the ED with a headache and slight fever. Her EKG and chest x-ray were normal, but her urine test was positive for a mild urinary tract infection (UTI). The correct diagnosis for this case would have been a potential decubitus skin ulcer, which would require a skin exam. However, most participants missed the need to conduct a skin exam to identify the cause of the AMS, as the UTI was too mild to cause the symptoms observed in the patient. The absence of the mention of a skin exam in participants key data, showed it possibly was not considered as data that had been ruled out, but it was also absent in the next steps provided. This would suggest that a skin ulcer had not been considered a potential diagnosis in the differential to follow up. The majority of the responses for diagnosis listed a UTI or cystitis as the diagnosis. A substantial number listed a head injury related diagnosis, e.g., intercranial hemorrhage (ICH). Both these conditions are related to information provided in the sign out narrative. This focus on the UTI or head injury occurred despite the positive urinalysis results for UTI being too weak to explain the severity of the symptoms and the neuro-

exam related information refuting the possibility of head injury resulting from the fall. The presence of the positive urinalysis and knowledge of the fall in this patient scenario could be viewed as possible information cues that had an anchoring effect on participants' reasoning concerning this case.

One explanation for the results of these two cases could be because the body systems involved contain a lot of ambiguity and therefore are difficult to diagnose. In both cases the correct diagnosis was also seldom mentioned in the differential, suggesting that they were not considered as potential diagnoses for follow up. Cases that are too ambiguous would be expected to have broader differentials and so would be more likely to include the correct diagnoses as options in the differential.

The other explanation was that these cases were designed to be too difficult. The abdomen case was adapted from a case provided by an EM physician SME based on a case from his own experience with an actual patient. Minor details had been changed to protect any possibility of identifying the actual patient but in general none of the vignettes contained any specific patient details other than age and gender. To verify these two cases as appropriate and comparable to the other stimuli cases, the head and abdomen case details were assessed and reviewed by a different group of two SMEs, who had not participated in the development of the case vignettes. This SME group verified these two cases were accurate in terms of representativeness of the case presentation and clinical details.

In order to understand possible factors about these two cases that might have contributed to the effects observed, the abdomen case and head case were further analyzed for

interaction effects for a number of parameters; diagnostic accuracy, the number of diagnoses in the differential, participants' confidence in their diagnosis and their readiness to plan treatment or disposition for the case. Additional details were provided about each case from examination of the specific items in the data.

8.4 Interaction Effects for Stimuli Cases

The analyses into the abdomen and head cases revealed some important differences between the two cases.

8.4.1 Abdomen Case – the results of the interaction analyses can be found in tables 35-43. Diagnostic accuracy for this stimuli case was examined for the test variables and was found to be significant for length but not for explicitness, which has been observed for throughout the stimuli cases. This pattern variation was observed for all the measures explored. Length of the case was associated with significant variation in the abdomen cases for the measure of *Number of Diagnoses in the Differential*. Short cases had a higher mean and thereby a broader differential than long cases. Similarly, length was significant factor for both participants' *Confidence in their Diagnosis*, as well as for their *Readiness to Plan Treatment* or *Disposition*, with long cases having higher levels for both. For all these measures, the test condition of explicitness had no effect on participants' responses.

So, from these analyses we can conclude the stimuli abdomen cases exhibited the same patterns that was observed in the other two stimuli cases for cardiac or lung chief complaints. The cases showed significant differences for long cases than short, where participants typically performed better for the measures in long cases than for the short

cases. The overall lower scores for diagnostic accuracy for this case may be related to the greater ambiguity associated with abdominal cases with symptoms that make them more difficult to diagnose, particularly for cases at an earlier stage. This was echoed in narratives from the EM physician interview study, where a clinician described the exact symptoms of this stimuli abdomen case:

> *"anyone who complains of chest pain and shortness of breath has a fairly broad differential."*

Indeed, a simple review of the item data for this case (see table 43) revealed that of the 54 instances of an incorrect diagnosis, 22 listed a diagnosis related to a cardiac chief complaint and 16 listed a pulmonary chief complaint. Also, even within these instances, short cases made up the majority of these incorrectly identified chief complaints, with 18 for the cardiac misdiagnoses and 13 for the lung misdiagnoses. In conclusion, this abdomen case may have scored lower for diagnostic accuracy overall, possibly due to a more atypical presentation than seen in other abdominal cases. However, the length of the case, thereby the stage and volume of information for the case, still showed the same significant effect consistently for all the measures in this case as seen throughout this study.

8.4.2 Head Case – unlike the results of the abdomen case, the results for this head case did not show the same pattern of variation related to the test conditions of explicitness and length. For all the measures of diagnostic accuracy, number of diagnoses in the differential and participants' confidence in their diagnosis, there were no significant differences for either explicitness or length. There was a significant difference noted for

the measure of readiness to plan treatment for the condition of length but not for explicitness. This result could be due to clinicians' awareness of the serious and potentially fatal consequences of not treating cases with a head related chief complaint. Emergency medicine clinicians are probably the most adept at operating under conditions of great uncertainty, as so often patients present with serious emergent conditions in the ED. In these cases, the primary goal of the clinician is to stabilize the patient and prevent any imminent deterioration in their condition, whether a definitive diagnosis has been reached or not. With head cases this would be particularly so. The poor diagnostic accuracy observed in the results for this case may be related to the ambiguity of this case and the fact that the presence of UTI had an anchoring effect as it is listed as the primary diagnosis for 43% of cases. (see table 52).

Also, while the length of the case did not have any statistically significant effect on the measures of *Number of Diagnoses in the Differential* and *Confidence in their Diagnosis*, the p-values for these analyses were only slightly above the cut off for significance of 0.05, suggesting there were differences between long and short cases but just not substantially so. What is interesting is that for this case participants did identify the stage of the case 70% of the time so they were aware of the unfolding of the diagnostic process with this case. The question is if the case had progressed even further, would they have at any point, identified that the cause of the AMS could not be explained by the results of the diagnostic tests already ordered? Then they might have realized an actual credible cause was needed to explain the AMS in a patient of this age and state and realized the need for a skin exam.

In this scenario it is the absence of evidence that is significant. People may find it easier to consider evidence that is provided, to rule in or rule out conditions as part of their reasoning process, than to consider things that are absent. The presence of the UTI, even at too mild a level, may have provided a readily available cue, as evidence to the explain the symptoms observed. This case also contained numerous other pieces of evidence; the earlier fall, the headache, the disorientation and confusion, the positive leukocyte esterase result (a sign of infection), and the raised temperature. All of which could have been discounted from the tests and exams conducted. Perhaps in the presence of a signal, albeit a weak one, amongst an array of other noise, it was too difficult for participants to detect the absence of a true signal, the decubitus skin ulcer, as the potential root cause of the AMS. Only four participants listed that the UTI was equivocal in their *Key Data* response and stated a skin exam for their Next Steps. Others listed much more invasive and time-consuming next steps such as lumbar punctures and MRIs which would not only be costly to the patient but also to the hospital when a relatively simple skin exam would obviate the need for further tests.

In conclusion, the chief complaint seems to be a very important factor in the development of diagnoses for cases in sign out. We see this particularly for head cases, both in the control and stimuli cases, which can present as a set of symptoms that are broad, vague, and confounding. From the physician interviews:

> *"Anyone who complains of headache can be anything from something that's catastrophic and emergent to something that is benign."*

These type of cases of altered mental state, whether due to a skin ulcer or an electrolyte imbalance, could be considered as a possible risk factor for anchoring. The absence of an obvious causal piece of evidence could be missed in the development of a differential diagnoses. With that comes the risk of failure to order test and conduct exams to identify them. This failure could thus lead to a delayed or even wrong diagnosis and ultimately to medical error.

8.5 Secondary Outcomes: Key Data and Next Steps

The results for the secondary outcomes of Key Data Match with Sign Out was no effect observed for location or role but there was an effect for the case type. (see table 33) However, further analyses showed this effect was seen only in the control cases and not the stimuli cases so it was concluded that participants did not have any differences in identifying key data for the stimuli cases.

In terms of the next steps that participant stated for the case, the results did not show any differences when examined for role but did show difference when examined by the case type. (see table 34) When examined for chief complaint there were significant differences indicated for all the cases together, for controls only and for stimuli cases only as a group. However, when the stimuli cases were examined for the test conditions of explicitness and length, there were no statistically significant differences found. What was interesting however, was that for location the analyses showed there was a site-specific variation, as UMB demonstrated higher levels for identifying the correct next steps than UTH. The analysis results were statistically significant but only just with a p-value of 0.04

indicating the difference observed between the two sites though not extreme was still substantial.

8.6 Location

So, throughout all the analyses the results of the participant responses did not show any differences when compared for location for any measure except one. Next Steps Match compared the response data for question 4 in the survey, when participants were asked to provide their next steps for each case, to the reference medical literature (Levis & Garmel, 2009; Okuda & Nelson, 2009) Next steps listed did show a statistically significant difference for location. (see table 53: Summary Matrix) The UMB site showed statistically significant higher match for next steps at 91% for UMB than the 86% for UTH.

Sign outs at the UMB site are conducted following a checklist protocol (see Appendix B figure 13), whereas UTH hospitals do not routinely use handoff protocols. It is possible that this site-specific difference in next steps planning may be due to protocol driven practices and communication. The UMB checklist protocol has three specific sections that could be considered as informing next steps planning; 1. 'Pending results consults', 2. 'Summary of what oncoming team needs to address vs what off-going team will address', and 3. 'Close the loop'. It could be argued that having these components included in every sign out, contributes to establishing stronger shared mental models between the oncoming and exiting providers as well as providing clarity in terms of responsibilities for the next steps in a case after sign out has completed. The results

observed for next steps planning for UMB may possibly be attributed to the benefits from the components in the sign out protocol used.

8.7 Role

The results of analyses did not indicate that there were differences in diagnostic accuracy by role. However, there were differences observed in how the information was considered and used. While most participants provided multiple diagnoses in their differentials, attendings were found to have slightly higher rates for providing a single diagnosis in their differential (22%) than Residents and Fellows (20%) but APPs were the highest for single diagnosis (32%) see table 28. However, in terms of confidence in their diagnosis (see table 30), the attendings had the lowest mean confidence (67.5(25.9)) compared the residents and fellows (68.6 (23.1)), while the APPs had the highest mean confidence (75.5(18.9)).

This was an interesting finding that might be explained by the concept of someone with less experience but with textbook knowledge who may recognize a case as a clear instance of a particular condition based on pattern recognition from previous typical cases. However, someone with increasing experience may also recognize the case as an instance of the condition but have greater uncertainty because they have had more exposure to the atypical presentations of similar cases. These experts are more aware of the ambiguity in cases and of what information may possibly be missing.

The greater level of confidence observed in APPs might be argued to be due to a form of the Dunning-Kruger effect, (Kruger & Dunning, 1999) where people with lower levels of expertise at a task may feel a greater sense of confidence in their ability. This confidence

may manifest so that they are unaware of the limits of their knowledge and expertise due to the limits of their knowledge and expertise. However, our results do not suggest that the APP group performed any differently than the expert and trainee physician groups, in terms of their ability to accurately diagnose the cases in this experiment. The results indicated that there was no statistical difference between the roles for diagnostic accuracy. (See table 24)

In UTH hospital emergency departments, APPs tend to be assigned to manage lower acuity patients in the ED. In addition, they also tend to staff the critical care or higher acuity areas less frequently than the attendings, residents and fellows. Consequently, while APPs are in job functions to see all type of patients all the time, they tend to see a higher proportion of less serious patients, than their physician counterparts. These patients may be more homogenous in their types of presentation. So, an alternative explanation for why mid-level providers were more confident in their decisions, could be because they normally deal with less complex cases that may be more clear-cut, over time. Consequently, this increased exposure to specific patient subsets may contribute to a type of availability bias in APPs.

Perhaps biases like the Dunning-Kruger effect and availability heuristic work to promote anchoring bias in some scenarios. The interaction effects of multiple cognitive biases are not well understood and the methods or tools to study them are not well developed. Other studies have used patient vignettes to investigate cognitive biases in diagnostic decision making. (Jenkins & Youngstrom, 2016a; Lutfey et al., 2009; Mohan et al., 2017) However, few studies aim to tease out the specific factors contributing to the individual

biases but rather acknowledge that multiple cognitive biases are playing out in the scenarios they are investigating.

Finally, it should be mentioned when describing the responses of APPs, we should take care to generalize about this participant group. Only one APP from UMB participated in the study and as such the APP group responses could be considered as more specific to UTH. The responses from APPs may be subject to a potential latent site-specific effect not yet identified in the data.

8.8 Conclusions of EDSO study

In conclusion, a number of key findings were identified from this study:

- The test condition of length of the case had a significant effect on nearly all measures in experiment but the explicitness of the information did not. As length equates to the stage in the diagnostic process, we can say that cases that were earlier in stage showed marked differences in outcomes like diagnostic accuracy, confidence in the diagnosis, whether the correct diagnosis was present in the differential, as well as how broad the differential was, than later cases.
- The chief complaint for a case is very important. The head cases presenting as altered mental state were associated with significantly lower diagnostic accuracy for both the controls and the stimuli cases. Head cases with multiple, diverse, and vague symptoms clearly pose a challenge for clinicians in terms of diagnosis and treatment planning. Certain scenarios of altered mental state caused by less immediately apparent sources, such as skin ulcers, when presented with a more readily available but weakly evidential cause, may be a serious potential risk for anchoring bias.

- Missing or absent information at the time of sign out has potential consequences on the reasoning of the recipient of the sign out information. While explicitness of information that is available was not a factor in decision making, explicitly highlighting absent information may be important. Explicitly mentioning information such as test and examination results that had not been done, e.g., the skin exam, or that is pending, may support the mental model of the oncoming clinician and help them to formulate their management plan.
- The role of the clinician did not affect their ability to provide accurate diagnoses for a variety of different cases. Attendings, residents, fellows, and APPs showed no significant differences in their diagnostic accuracy in this study.
- There were no site specific differences observed across this study except for *Next Steps Planning*. This suggests that the use of handoff protocols did not affect participants' clinical reasoning except at the point of treatment planning.
- Using a randomized survey format combined with detailed patient vignettes proved to be an effective way to assess outcomes for clinical decision making without the need to resort to simulations. Using an iterative design process, combined with regular consultation with subject matter experts, ensured important clinically relevant factors were reflected in the design. This resulted in achieving an effective and reliable method for conducting experimental assessment for clinical outcomes using readily available tools that could be implemented remotely without the need for in person interaction.

8.9 Application/Recommendations

8.9.1 Early Stage Cases

This result of the study identified that the length of the case had a significant effect on the diagnostic reasoning of EM clinicians. A recommendation derived from this research would be enhancements to current handoff practices to incorporate how early stage cases are handled and communicated during sign outs.

A potential intervention might involve highlighting the need for support around cases at an earlier stage at the time of sign out and raising awareness that these early cases may play out differently than later ones. Currently most sign out practices in most institutions are conducted in the order of the physical bed layout in the ED. In settings where protocols are not used, perhaps there could be a reminder added to the sign out practice or sign out training, to specifically mention the stage in the diagnostic process the patient is in or to highlight what diagnostic testing results are still pending or have not been done particularly for cases that are earlier the process. In settings using handoff protocols such as checklists, an additional section could be incorporated to the protocol that covered the stage of the case and what information is pending or required to be able to progress the case to the next stage.

These steps would help to make people more aware of the stage the case was in and whether further narrowing of the differential was needed. Various different approaches have already been developed to prompt clinicians into awareness of their metacognition as a means of avoiding biases during clinical decision making (Ludolph & Schulz, 2018; Sullivan et al., 2015) Highlighting the need to provide additional support for early stage

cases is a type of cognitive forcing strategy, as are many de-biasing interventions. However, the innovation that focusing on the stage of the case brings, is that it pivots from the perspective of the patient's course and length of stay in the ED, to the clinician's diagnostic reasoning process and the information required to support it.

8.9.2 Support for High Ambiguity Cases - Altered Mental State Cases

The other significant and unexpected finding was that patient cases involving an altered mental state significantly influenced clinicians' ability to identify the underlying source of the condition and thereby correct diagnosis. For two AMS cases in this study, participants suggested further tests or treatments that were incongruous with the actual diagnosis despite the presence of evidence that did not support their diagnostic path. Often these tests or treatments are expensive and invasive. In the case of the stimuli head case involving AMS due to a decubitus ulcer presenting with a mild UTI, some suggested further tests, such as lumber punctures, MRI, head CT as well as further blood tests. Lumber punctures and MRIs are not only sometimes distressing to patients but are also costly. All these procedures could be avoided by doing a simple skin exam that involves minimal cost and is typically faster. A study investigating EHR documentation showed that missed documentation of skin ulcers was an important source of lost revenue for many hospitals. CMS reimbursement for pressure ulcer related care is based on present-on-admission (POA) diagnosis documentation. If the ulcer is discovered later in the patient's stay, for example by a nurse providing care, the costs of this treatment will not be reimbursed to the hospital. The study found there was a 76% mismatch between nurse and physician documentation of pressure ulcers as present on admission. (Moerbe &

Kelemen, 2014) This type of information mismatch due to failure to do something as relatively simple as a skin exam, is not only costly to a hospital's operating budgets but has serious patient safety implications also. Missing the skin ulcer could lead to a delayed or even incorrect diagnosis and potential deterioration in the patient's condition, which might even prove fatal in an elderly patient.

One recommendation might be to have a reminder to first check and rule out possible hidden causes for a presenting condition. This reminder could be added to sign out checklists and be specifically for certain types of vulnerable patients, e.g. elderly or wheelchair bound or bed bound patients presenting with AMS should routinely be checked for evidence of pressure ulcers before initiating more involved testing procedures.

An intervention of this nature could be implemented relatively easily without the need or expense of a major intervention implementation. It could be implemented as easily as including a reminder item in a checklist or highlighting this issue by presentation of such a case during grand rounds. It could be reinforced during medical training by including it as part of the standard questions that should be asked about this type of case during sign out, rather than assuming a skin was done because it was not mentioned. By promoting explicit inquiry about omitted data would help clinicians develop a common understanding and shared mental model of a case during a sign out.

Many checklists have been developed to support prevention of diagnostic error in the ED and some even focus on the common missed conditions for particular chief complaints (Graber et al., 2014) Similar checklists could be drawn to focus on commonly missed or

overlooked pieces of evidence, which could be asked about during sign out. While many checklists mention tests as a category, they do not specifically suggest explicit communication about the content, such as which tests have been completed or have not been done and also tests that are pending or should be considered. This would ensure that there are no communication gaps regarding testing and results between the outgoing and oncoming clinicians.

8.9.3 Survey + Vignette Format – an Effective Assessment Tool

The initial plan for this study was to conduct a simulation of a sign out. However, the occurrence of the COVID-19 pandemic necessitated using a different experimental platform. The study was designed and conducted using an online survey format. The results of this study have shown to be consistently significant for the test condition of length of the case, while not significant for the condition of explicitness. Length equates to stage of the case and as such this study has demonstrated that the stage that a patient case at the time of sign out has a strong effect on diagnostic reasoning. The results also show other factors such as role specific or site-specific factors did not influence the effect on the study outcomes observed. The consistency of the results throughout the measures in the experiment is suggestive of low type 1 errors in the design. This suggests the study outcomes were sensitive to and reflective of the effect of the test conditions within the experimental design.

Hence it is proposed that this approach of using patient case vignettes in an online survey format could be an effective method for assessing clinical decision making outcomes and could be used in other clinical institutions without the need for full simulation based

methods. This is a particularly useful finding given that currently there are still restrictions around access to medical facilities and personnel, due to the ongoing COVID-19 pandemic. In addition, participants were able to complete this study in their own time. Therefore, a study using this type of online format, provides significant advantages in terms of participant recruitment and scheduling, as well as study resource management, especially given how busy clinicians can be. In conclusion this is a method for conducting an effective experimental study that yields data that reliably reflects actual effect size in study outcomes. The fact that it can be managed and conducted remotely and completed at the convenience of participants' personal agendas, makes it a potentially useful tool for designing clinical case based experiments in the near future.

8.9.4 Generalizability

In terms of generalizability, the two sites involved a protocol compliant and a non-protocoled institution. The analysis showed that the results we saw were not related to site specific factors. As such it was concluded protocol use did not influence the main study outcomes. The location did not have influence on any of the results, except for next steps planning. Potential modifications to the case design could be made to consider protocol driven tasks where applicable. One could argue that both the sites in this study were academic institutions. Hence the generalizability of this survey based method only applies to academic institutions. However, a future study could be conducted comparing a non-academic setting along with another academic institution to further assess how generalizable this experimental approach might be.

8.9.5 Vignette Staging Cases to Aid Training in Diagnosis

Cognitive forcing strategies like checklists and mnemonics have been effective but their effectiveness over time has been hard to demonstrate. (O'Sullivan & Schofield, 2019) Patient vignettes are a good way to present patient cases in a training setting.

Using a vignette case similar to the stimuli head case in this study during training by developing visual interactive tools containing patient vignette scenarios that allow adjusting the point of the sign out, may help with de-biasing training for clinicians. Greater need to include awareness of cognitive bias and mitigating solutions as part of medical training have been suggested already (Croskerry, 2014; Jenkins & Youngstrom, 2016) Experiential de-biasing efforts have been shown to achieve greater and longer lasting effects ((Dunbar et al., 2014; Mohan et al., 2017) In one study by Mohan and colleagues (2017), a video game was developed containing personally relatable patient vignettes as a de-biasing solution for trauma scenarios. The video game arms of the study were shown to be highly successful in reducing bias for both high and low cognitive load cases. In this study the under triaging of trauma cases was improved with a video gaming tool that achieved greater efficacy, which was sustained over time, than the traditional approaches of didactic trauma training materials used for trauma certification. (Mohan et al., 2017) Introducing more experiential approaches during training may help the development of heuristics and system 1 processes that clinicians, novice and experts alike use as part of their clinical reasoning. Using visual representations of patient cases like those developed in this study may provide a useful to tool to facilitate discussion about the clinical details of the case, as it did in the iterative

design phase of this study. Making the visual aid interactive would further enhance its possible usefulness.

8.10 Limitations

8.10.1 Selection bias of participants

The participants recruitment process involved clinician SME collaborators reaching out the faculty and clinical staff at their institutions via direct email, fliers, and departmental meetings. Participants for the experimental were effectively a convenience sample from a pool of ED clinicians, who chose to participate in survey-based study. However, limitations of convenience samples are that these participants may not be representative of the general target emergency medicine clinician population. However, the efforts were made to reach out to all types of ED clinicians including attendings, residents, fellows, and APPs. Only student clinicians were not eligible to participate.

Also, the study was conducted over two sites, UTH in Houston, Texas and UMB in Baltimore, Maryland, making the participant population more diverse. However, both institutions were academic facilities and as such the findings may only be applicable to EM clinicians and EDs in academic settings. Despite this, it is possible that the findings of this work may have some relevance and inform further research into cognitive bias for medical training and practice in emergency medicine settings.

8.10.2 Limited Sample size

The sample size of the participant groups may be limited in terms of providing ideal statistical power to the study. Due to the restricted time frame available within a doctoral degree and the limited funds available for participant recruitment activities, it was

necessary to recruit from institutions where SBMI faculty had collaborative relationships, these being UTH and UMB. Despite the time and resource constraints, the project was able to recruit reasonable numbers and almost equal sample sizes at both sites with an overall sample size of 69. This was in large part to the dedicated and continued efforts to the SME clinician collaborators at both sites. The randomized assignment of the test conditions in the experiment required a minimum of sixteen participants at each site. With 35 participants at UTH and 34 at UMB a sufficiently large sample size of participants was achieved to enable statistical analyses to be conducted.

8.10.3 Not standardized Cases

All the patient case vignettes were developed for this study, and none were taken from standardized clinical cases. This is because standardized cases for the purpose of research are few and access to them is limited. Also, the conditions that were being tested in the experiment would be required to be built in the patient case vignette of the experiment and these conditions may not have been compatible with any standardized cases available. To this end the cases used for this study were sourced from standard medical reference literature. These are texts that are used by EM clinicians as part of their medical training or board certification ((Levis & Garmel, 2009; Okuda & Nelson, 2009) The cases were built following a systematic and consistent process outlined in chapter 6. In addition, the cases were discussed and verified in an iterative manner throughout the design process via SME consultation. All efforts were made to ensure the cases were balanced and representative of real patient cases.

8.10.4 Randomized Assignment and Incomplete Responses

The randomized assignment of the survey instances was achieved by setting parameters in Qualtrics software to sample from the 16 surveys in the survey flow evenly and assign to participants as they initiated their survey attempt. The informed consent page of the survey required participants to provide consent to be able to proceed. If participants did not provide consent, they were exited from the survey. In such cases the aborted attempt would be recorded as an incomplete response. Also, the attempts where participants left the survey without completing all the questions would be marked as incomplete responses. Because the random assignment was done in real time from the 16-item survey pool, these incomplete responses did impact the automated random assignment process and perfectly even random assignment generation was affected. However, all computer-generated randomizations are never truly random, so this is a systemic issue that can affect all experimental studies that employ automated randomization methods.

Randomized studies represent an approximation of randomization for the conditions being tested. The purpose of randomization is to strive to achieve even presentation of all the instances of control and stimuli conditions to participants without influencing which participant receives which assignment, e.g., randomized controlled trials use randomized assignment to try to evenly assign participants to the control and intervention arms of a study in order to balance out the effect of participant characteristics across the study results. While the true randomization within the Qualtrics was affected by incomplete responses, the distribution of the stimuli case permutations was monitored to ensure that floor and ceiling extremes of assignment had not occurred. The overall assignment of

cases was found to be relatively balanced across all the test condition permutations. This is somewhat evident by the consistency in the results of the statistical analyses. The effects observed for the outcomes and the test conditions of explicitness and length were consistent for all the measures. This suggests that there were minimal type 1 errors and that the data distributions were sufficiently balanced across the case instances. The consistency of the finding across all the statistical analyses conducted suggest that the findings observed were due actual effects sizes attributed to the test conditions.

8.10.5 Participant Attrition

There were some cases of participant attrition; UTH had 60 attempted responses, of which 35 were completed responses and UMB had 50 attempted responses, of which 34 were complete responses. Review of the incomplete responses showed the majority were responses terminated at the point of the consent page plus a handful of attempts terminated after the first or second case was presented. It should be noted that the survey was demanding both in terms of time and cognitive effort. The target participants were EM clinicians, who are extremely busy professionals with limited time and sometimes unpredictable working hours. Unfinished attempts for this survey were expected as some people may have started the survey with the intent of completing it but either got distracted or found they did not have a convenient 30-60minute window of time to complete it. The survey assignment was randomized and the questions within each survey were randomized for presentation order. As such when the unfinished attempts were reviewed no case(s) or question(s) were observed to be associated with abandoned response attempts.

8.10.6 Stimuli Head Case Design

The stimuli head case of AMS due to a decubitus skin ulcer coupled with a mild case of UTI, had significantly lower diagnostic accuracy than all the other stimuli cases. A possible reason was thought to be that the case had been designed to be too difficult, as so few correctly identified the diagnosis or included it in their differential. However, consultation with SME confirmed that the case was accurate, and representative of AMS cases seen in the ED.

The other possible reason could be that because no mention of a skin exam was made either in terms of being needed or planned in the sign out narratives, participants may have assumed a skin exam had already been conducted but no significant findings made. Typically, during medical training oral board exams students know that if a pertinent piece of information about the case is not mentioned, they should make no assumptions about its status, and they will lose points if they fail to ask about it. So, if this head case had been presented during a board exam, students would lose points if they failed to ask whether a skin exam had been conducted. However, students taking board exams are briefed on these rules and their responsibility to ask about or mention key information.

A limitation of this study design was that no such ground rules were provided to participants about the absence or presence of key information. Consequently, some participants may have assumed because no mention of the skin exam was provided, that it was done and was not significant. Perhaps this could have been mitigated by providing brief instructions at the start of the survey. The other option would be to include a long case instance that hinted at the need for conducting/ruling out a skin exam. In such a

scenario, it would be interesting to see if this prompted more participants to consider a decubitus skin ulcer as the diagnosis or in their differential. It should be noted though that the SMEs did state that this type of case, i.e., an elderly patient with AMS and other weakly presenting issues, is very typical for the ED and it is not uncommon for the presence of skin ulcers to sometimes get overlooked.

8.10.7 Clinician Roles in Academic Settings

An important factor to consider is that because of the academic nature of the study sites, the participants all have different roles with which come different responsibilities. For example, attendings may list next steps differently than the other roles might, because they know that staff reporting to them, like residents, will perform some of the tasks and will consult with them for some cases before proceeding. Also, residents may respond to the confidence related questions such as confidence in diagnosis, in their diagnostic scope, readiness to treat, differently based on where they are in their training on the program. The analyses conducted was done by role but not broken down by years of experience. This might be an interesting measure study in future data analyses.

8.10.8 Subjective Data Synthesis for some Created Variables

Data Match and Next steps were interpreted created variables based on participant's responses to questions three and four in the survey. In question three they were asked what key data about the case influenced their decision making and in question four they were asked what their next steps would be for the case. The responses were compared by the researcher to the data presented in the case narrative and the next steps mentioned to the medical reference text respectively. For both these created variables, the data

synthesis involved a subjective interpretation of the response data. In addition, the researcher did not have clinical training. These factors may have influenced the accuracy of the data synthesized for these two created variables. With this in mind, a second review of the response data could mitigate for these limitations. In the interest of reducing the subjectivity of a second pass of the data, the second reviewer should be blinded to whether the participant had provided accurate diagnosis and ideally should possess some clinical training. The time constraints of this doctoral research project did not permit for this second level review, but this will be conducted in the near future in preparation for publication of this work.

8.11 Final Conclusion

This dissertation was exploratory research into the factors related to the presentation of information that might contribute to the influence of anchoring bias in diagnostic decision making and was studied within the context of emergency department sign outs. The results of the study suggest that volume of information as it pertains to the stage of the case in the diagnostic process does have a significant impact on clinician decision making. This work has laid the foundation for future research into the influence of information content and its communication on clinical decision making, to ultimately inform the development of informatics-based decision support solutions and may be applicable to other health care subdomains, such as medical education and training.

References

Abraham, J., Kannampallil, T., Brenner, C., Lopez, K. D., Almoosa, K. F., Patel, B., & Patel, V. L. (2016). Characterizing the structure and content of nurse handoffs: A Sequential Conversational Analysis approach. *Journal of Biomedical Informatics*, *59*, 76–88. https://doi.org/10.1016/j.jbi.2015.11.009

Abraham, J., Kannampallil, T. G., Patel, V. L., Patel, B., & Almoosa, K. F. (2016). Impact of Structured Rounding Tools on Time Allocation During Multidisciplinary Rounds: An Observational Study. *JMIR Human Factors*, *3*(2), e29. https://doi.org/10.2196/humanfactors.6642

Abraham, J., Kannampallil, T., Patel, B., Almoosa, K., & Patel, V. L. (2012). Ensuring patient safety in care transitions: An empirical evaluation of a Handoff Intervention Tool. *AMIA ... Annual Symposium Proceedings. AMIA Symposium*, *2012*, 17–26.

Abraham, J., Kannampallil, T., & Patel, V. L. (2014). A systematic review of the literature on the evaluation of handoff tools: Implications for research and practice. *Journal of the American Medical Informatics Association*, *21*(1), 154–162. https://doi.org/10.1136/amiajnl-2012-001351

Abraham, J., Nguyen, V., Almoosa, K. F., Patel, B., & Patel, V. L. (2011). Falling through the cracks: Information breakdowns in critical care handoff

communication. *AMIA ... Annual Symposium Proceedings. AMIA Symposium*, *2011*, 28–37.

Anderson, C. E., Nicksa, G. A., & Stewart, L. (2015). Distractions During Resident Handoffs: Incidence, Sources, and Influence on Handoff Quality and Effectiveness. *JAMA Surgery*, *150*(5), 396–401. https://doi.org/10.1001/jamasurg.2014.2459

Apker, J., Mallak, L. A., & Gibson, S. C. (2007). Communicating in the "Gray Zone": Perceptions about Emergency Physician–hospitalist Handoffs and Patient Safety. *Academic Emergency Medicine*, *14*(10), 884–894. https://doi.org/10.1197/j.aem.2007.06.037

Armstrong, M. B., & Thurber, J. (2014). Disseminated Kaposi's Sarcoma—A Missed Diagnosis. *Journal of Emergency Medicine*, *47*(5), 520–523. https://doi.org/10.1016/j.jemermed.2013.08.139

Arora, V., Johnson, J., Lovinger, D., Humphrey, H. J., & Meltzer, D. O. (2005). Communication failures in patient sign-out and suggestions for improvement: A critical incident analysis. *BMJ Quality & Safety*, *14*(6), 401–407. https://doi.org/10.1136/qshc.2005.015107

Arora, V. M., Eastment, M. C., Bethea, E. D., Farnan, J. M., & Friedman, E. S. (2013). Participation and Experience of Third-Year Medical Students in Handoffs: Time to Sign Out? *Journal of General Internal Medicine*, *28*(8), 994–998. https://doi.org/10.1007/s11606-012-2297-9

Augestad, L. A., Stavem, K., Kristiansen, I. S., Samuelsen, C. H., & Rand-Hendriksen, K. (2016). Influenced from the start: Anchoring bias in time trade-off valuations. *Quality of Life Research : An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *25*(9), 2179–2191. https://doi.org/10.1007/s11136-016-1266-x

Bates, D. W., & Singh, H. (2018). Two Decades Since To Err Is Human: An Assessment Of Progress And Emerging Priorities In Patient Safety. *Health Affairs*, *37*(11), 1736–1743. https://doi.org/10.1377/hlthaff.2018.0738

Blumenthal-Barby, J. S., & Krieger, H. (2015). Cognitive Biases and Heuristics in Medical Decision Making: A Critical Review Using a Systematic Search Strategy. *Medical Decision Making*, *35*(4), 539–557. https://doi.org/10.1177/0272989X14547740

Brixey, J. J., Tang, Z., Robinson, D. J., Johnson, C. W., Johnson, T. R., Turley, J. P., Patel, V. L., & Zhang, J. (2008). Interruptions in a level one trauma center: A case study. *International Journal of Medical Informatics*, *77*(4), 235–241. https://doi.org/10.1016/j.ijmedinf.2007.04.006

Campbell, S. G., Croskerry, P., & Bond, W. F. (2007). Profiles in Patient Safety: A "Perfect Storm" in the Emergency Department. *Academic Emergency Medicine*, *14*(8), 743–749. https://doi.org/10.1197/j.aem.2007.04.011

Chapman, G. B., & Elstein, A. S. (2000). Cognitive processes and biases in medical decision making. In *Decision making in health care: Theory, psychology, and applications* (pp. 183–210). Cambridge University Press.

Chapman, G. B., & Johnson, E. J. (1994). The limits of anchoring. *Journal of Behavioral Decision Making*, *7*(4), 223–242. https://doi.org/10.1002/bdm.3960070402

Chapman, G. B., Johnson, E. J., & Chapman, G. (2002). Incorporating the irrelevant: Anchors in judgments of belief and value. *In The Psychology of Intuitive Judgement: Heuristics*.

Chapman, G. B., & Sonnenberg, F. A. (2003a). *Decision Making in Health Care: Theory, Psychology, and Applications*. Cambridge University Press.

Chapman, G. B., & Sonnenberg, F. A. (2003b). *Decision Making in Health Care: Theory, Psychology, and Applications*. Cambridge University Press.

Cheung, D. S., Kelly, J. J., Beach, C., Berkeley, R. P., Bitterman, R. A., Broida, R. I., Dalsey, W. C., Farley, H. L., Fuller, D. C., Garvey, D. J., Klauer, K. M., McCullough, L. B., Patterson, E. S., Pham, J. C., Phelan, M. P., Pines, J. M., Schenkel, S. M., Tomolo, A., Turbiak, T. W., … White, M. L. (2010). Improving Handoffs in the Emergency Department. *Annals of Emergency Medicine*, *55*(2), 171–180. https://doi.org/10.1016/j.annemergmed.2009.07.016

Chew, K. S., Durning, S. J., & van Merrienboer, J. J. (2016). Teaching metacognition in clinical decision-making using a novel mnemonic checklist: An exploratory study. *Singapore Medical Journal*, *57*(12), 694–700. https://doi.org/10.11622/smedj.2016015

Clegg, B. A., McKernan, B., Martey, R. M., Taylor, S. M., Stromer-Galley, J., Kenski, K., Saulnier, E. T., Rhodes, M. G., Folkestad, J. E., McLaren, E., Shaw, A., & Strzalkowski, T. (2015). Effective Mitigation of Anchoring Bias, Projection Bias,

and Representativeness Bias from Serious Game-based Training. *Procedia Manufacturing*, *3*, 1558–1565. https://doi.org/10.1016/j.promfg.2015.07.438

Collyer, S. C., & Malecki, G. S. (1998). Tactical decision making under stress: History and overview. In *Making decisions under stress: Implications for individual and team training* (pp. 3–15). American Psychological Association. https://doi.org/10.1037/10278-016

Committee on Diagnostic Error in Health Care, Board on Health Care Services, Institute of Medicine, & The National Academies of Sciences, Engineering, and Medicine. (2015). *Improving Diagnosis in Health Care* (E. P. Balogh, B. T. Miller, & J. R. Ball, Eds.). National Academies Press (US). http://www.ncbi.nlm.nih.gov/books/NBK338596/

Cook, L. J., Knight, S., Junkins Jr., E. P., Mann, N. C., Dean, J. M., & Olson, L. M. (2004). Repeat Patients to the Emergency Department in a Statewide Database. *Academic Emergency Medicine*, *11*(3), 256–263. https://doi.org/10.1111/j.1553-2712.2004.tb02206.x

Cree, C., Pillai, A., Jones, B., & Blyth, M. (2007). Bilateral patellar tendon ruptures: A missed diagnosis. *Knee Surgery, Sports Traumatology, Arthroscopy*, *15*(11), 1350–1354. https://doi.org/10.1007/s00167-007-0350-8

Croskerry, P. (2002a). Achieving quality in clinical decision making: Cognitive strategies and detection of bias. *Academic Emergency Medicine : Official Journal of the Society for Academic Emergency Medicine*, *9*(11), 1184–1204.

Croskerry, P. (2002b). Achieving quality in clinical decision making: Cognitive strategies and detection of bias. *Academic Emergency Medicine : Official Journal of the Society for Academic Emergency Medicine*, *9*(11), 1184–1204.

Croskerry, P. (2003). The importance of cognitive errors in diagnosis and strategies to minimize them. *Academic Medicine : Journal of the Association of American Medical Colleges*, *78*(8), 775–780.

Croskerry, P. (2009a). A Universal Model of Diagnostic Reasoning. *Academic Medicine*, *84*(8), 1022–1028. https://doi.org/10.1097/ACM.0b013e3181ace703

Croskerry, P. (2009b). Context Is Everything or How Could I Have Been That Stupid? *Healthcare Quarterly*, *12*(sp), e171–e176. https://doi.org/10.12927/hcq.2009.20945

Croskerry, P. (2013). From Mindless to Mindful Practice—Cognitive Bias and Clinical Decision Making. *New England Journal of Medicine*, *368*(26), 2445–2448. https://doi.org/10.1056/NEJMp1303712

Croskerry, P. (2014). Bias: A normal operating characteristic of the diagnosing brain. *Diagnosis*, *1*(1), 23–27. https://doi.org/10.1515/dx-2013-0028

Croskerry, P., & Norman, G. (2008). Overconfidence in Clinical Decision Making. *The American Journal of Medicine*, *121*(5), S24–S29. https://doi.org/10.1016/j.amjmed.2008.02.001

Croskerry, P., Singhal, G., & Mamede, S. (2013). Cognitive debiasing 2: Impediments to and strategies for change. *BMJ Quality & Safety*, *22 Suppl 2*, ii65–ii72. https://doi.org/10.1136/bmjqs-2012-001713

Davis, J. H., & al, et. (1984). Order effects in multiple decisions by groups: A demonstration with mock juries and trial procedures. *Journal of Personality*, *47*(5), 1003–1012.

Dhingra, K. R., Elms, A., & Hobgood, C. (2010). Reducing Error in the Emergency Department: A Call for Standardization of the Sign-out Process. *Annals of Emergency Medicine*, *56*(6), 637–642. https://doi.org/10.1016/j.annemergmed.2010.02.004

Dunbar, N. E., Miller, C. H., Adame, B. J., Elizondo, J., Wilson, S. N., Lane, B. L., Kauffman, A. A., Bessarabova, E., Jensen, M. L., Straub, S. K., Lee, Y.-H., Burgoon, J. K., Valacich, J. J., Jenkins, J., & Zhang, J. (2014). Implicit and explicit training in the mitigation of cognitive bias through the use of a serious game. *Computers in Human Behavior*, *37*, 307–318. https://doi.org/10.1016/j.chb.2014.04.053

Elstein, A. S. (1978). *Medical problem solving: An analysis of clinical reasoning*. Harvard University Press.

Elstein, A. S. (1999). Heuristics and biases: Selected errors in clinical reasoning. *Academic Medicine*, *74*(7), 791–794. https://doi.org/10.1097/00001888-199907000-00012

Ely, J. W., Graber, M. L., & Croskerry, P. (2011). Checklists to Reduce Diagnostic Errors. *Academic Medicine*, *86*(3), 307–313. https://doi.org/10.1097/ACM.0b013e31820824cd

Endsley, M. (1995). A taxonomy of situation awareness errors, human factors in aviation operations; *Proceedings of the 21st Conference of the European Association for Aviation Psychology (EAAP)*, *3*, 287–292.

Endsley, M. R. (1995). Measurement of Situation Awareness in Dynamic Systems. *Human Factors*, *37*(1), 65–84. https://doi.org/10.1518/001872095779049499

Epley, N., & Gilovich, T. (2005). When effortful thinking influences judgmental anchoring: Differential effects of forewarning and incentives on self-generated and externally provided anchors. *Journal of Behavioral Decision Making*, *18*(3), 199–212. https://doi.org/10.1002/bdm.495

Evans, J. S. B. T. (1984). Heuristic and analytic processes in reasoning *. *British Journal of Psychology*, *75*(4), 451–468. https://doi.org/10.1111/j.2044-8295.1984.tb01915.x

Evans, J. St. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7*(10), 454–459. https://doi.org/10.1016/j.tics.2003.08.012

Franklin, A., Liu, Y., Li, Z., Nguyen, V., Johnson, T. R., Robinson, D., Okafor, N., King, B., Patel, V. L., & Zhang, J. (2011). Opportunistic decision making and complexity in emergency care. *Journal of Biomedical Informatics*, *44*(3), 469–476. https://doi.org/10.1016/j.jbi.2011.04.001

Frye, K. L., Adewale, A., Martinez, C. J. M., & Montero, C. M. (2018). Cognitive Errors and Risks Associated with Provider Handoffs. *Cureus*, *10*(10). https://doi.org/10.7759/cureus.3442

Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics*, *40*(1), 35–42. https://doi.org/10.1016/j.socec.2010.10.008

Galinsky, M. (2001). *Ovid: First Offers as Anchors: The Role of Perspective-Taking and Negotiator Focus.* http://ovidsp.dc1.ovid.com/sp-3.32.2a/ovidweb.cgi

Gandhi, T. K., Kachalia, A., Thomas, E. J., Puopolo, A. L., Yoon, C., Brennan, T. A., & Studdert, D. M. (2006). Missed and Delayed Diagnoses in the Ambulatory Setting: A Study of Closed Malpractice Claims. *Annals of Internal Medicine*, *145*(7), 488. https://doi.org/10.7326/0003-4819-145-7-200610030-00006

Garcia-Molina, M., & Chicaiza Becerra, L. A. (2015). Anchoring Bias in Tto Valuations. *Value in Health : The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, *18*(7), A718. https://doi.org/10.1016/j.jval.2015.09.2716

George, J. F., Duffy, K., & Ahuja, M. (2000). Countering the anchoring and adjustment bias with decision support systems. *Decision Support Systems*, *29*(2), 195–206. https://doi.org/10.1016/S0167-9236(00)00074-9

Glaser, B. G., Strauss, A. L., & Strutzel, E. (1968). The Discovery of Grounded Theory; Strategies for Qualitative Research. *Nursing Research*, *17*(4), 364.

Gogan, J. L., Baxter, R. J., Boss, S. R., & Chircu, A. M. (2013). Handoff processes, information quality and patient safety: A trans-disciplinary literature review. *Business Process Management Journal*, *19*(1), 70–94. https://doi.org/10.1108/14637151311294877

Gorman, J. C., Cooke, N. J., & Winner, J. L. (2006). Measuring team situation awareness in decentralized command and control environments. *Ergonomics*, *49*(12–13), 1312–1325. https://doi.org/10.1080/00140130600612788

Graber, M. L., Franklin, N., & Gordon, R. (2005). Diagnostic Error in Internal Medicine. *Archives of Internal Medicine*, *165*(13), 1493–1499. https://doi.org/10.1001/archinte.165.13.1493

Graber, M. L., Kissam, S., Payne, V. L., Meyer, A. N. D., Sorensen, A., Lenfestey, N., Tant, E., Henriksen, K., Labresh, K., & Singh, H. (2012). Cognitive interventions to reduce diagnostic error: A narrative review. *BMJ Quality & Safety*, *21*(7), 535–557. https://doi.org/10.1136/bmjqs-2011-000149

Graber, M. L., Sorensen, A. V., Biswas, J., Modi, V., Wackett, A., Johnson, S., Lenfestey, N., Meyer, A. N. D., & Singh, H. (2014). Developing checklists to prevent diagnostic error in Emergency Room settings. *Diagnosis*, *1*(3), 223–231. https://doi.org/10.1515/dx-2014-0019

Greenstein, M., & Velazquez, A. (2017). Not all Anchors Weigh Equally. *Experimental Psychology*, *64*(6), 398–405. https://doi.org/10.1027/1618-3169/a000383

Hazan, A., & Haber, J. (2018). Mindful EM: The Dos and Don'ts of Sign-Out Etiquette. *Emergency Medicine News*, *40*(3), 25. https://doi.org/10.1097/01.EEM.0000531140.21325.f6

Heilman, J. A., Flanigan, M., Nelson, A., Johnson, T., & Yarris, L. M. (2016). Adapting the I-PASS Handoff Program for Emergency Department Inter-Shift Handoffs.

Western Journal of Emergency Medicine, *17*(6), 756–761. https://doi.org/10.5811/westjem.2016.9.30574

Hern, H. G., Gallahue, F. E., Burns, B. D., Druck, J., Jones, J., Kessler, C., Knapp, B., & Williams, S. (2016). Handoff Practices in Emergency Medicine: Are We Making Progress? *Academic Emergency Medicine*, *23*(2), 197–201. https://doi.org/10.1111/acem.12867

Horsky, J., Suh, E. H., Sayan, O., & Patel, V. (2015). Uncertainty, Case Complexity and the Content of Verbal Handoffs at the Emergency Department. *AMIA Annual Symposium Proceedings*, *2015*, 630–639.

Horwitz, L., Moin, T., & Green, M. (2007). Development and Implementation of an Oral Sign-out Skills Curriculum. *Journal of General Internal Medicine*, *22*, 1470–1474. https://doi.org/10.1007/s11606-007-0331-0

Jenkins, M. M., & Youngstrom, E. A. (2016a). A randomized controlled trial of cognitive debiasing improves assessment and treatment selection for pediatric bipolar disorder. *Journal of Consulting and Clinical Psychology*, *84*(4), 323–333.

Jenkins, M. M., & Youngstrom, E. A. (2016b). A randomized controlled trial of cognitive debiasing improves assessment and treatment selection for pediatric bipolar disorder. *Journal of Consulting and Clinical Psychology*, *84*(4), 323–333. https://doi.org/10.1037/ccp0000070

Jones, L., Kannampallil, T., & Franklin, A. (2013). End-of-round time compression in physician handoff sessions. *JAMA Internal Medicine*, *173*(11), 1033–1034. https://doi.org/10.1001/jamainternmed.2013.412

Kachalia, A., Gandhi, T. K., Puopolo, A. L., Yoon, C., Thomas, E. J., Griffey, R., Brennan, T. A., & Studdert, D. M. (2007). Missed and delayed diagnoses in the emergency department: A study of closed malpractice claims from 4 liability insurers. *Annals of Emergency Medicine*, *49*(2), 196–205. https://doi.org/10.1016/j.annemergmed.2006.06.035

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*(9), 697–720. https://doi.org/10.1037/0003-066X.58.9.697

Kannampallil, T. G., Abraham, J., & Patel, V. L. (2016). Methodological framework for evaluating clinical processes: A cognitive informatics perspective. *Journal of Biomedical Informatics*, *64*, 342–351. https://doi.org/10.1016/j.jbi.2016.11.002

Kannampallil, T. G., Jones, L. K., Patel, V. L., Buchman, T. G., & Franklin, A. (2014). Comparing the information seeking strategies of residents, nurse practitioners, and physician assistants in critical care settings. *Journal of the American Medical Informatics Association*, *21*(e2), e249–e256. https://doi.org/10.1136/amiajnl-2013-002615

Kannampallil, T., Jones, L., Buchman, T., & Franklin, A. (2011). 638: LAST PATIENTS FINISH LAST: END OF ROUND TIME COMPRESSION DURING CT ICU CLINICAL ROUNDS. *Critical Care Medicine*, *39*(12), 176. https://doi.org/10.1097/01.ccm.0000408627.24229.88

Kaustia, M., Alho, E., & Puttonen, V. (2008). How Much Does Expertise Reduce Behavioral Biases? The Case of Anchoring Effects in Stock Return Estimates.

Financial Management, *37*(3), 391–412. https://doi.org/10.1111/j.1755-053X.2008.00018.x

Keeney, S., & Halalau, A. (2017). Anchoring bias in a case of recurrent abdominal pain. *BMJ Case Reports*, *2017*. https://doi.org/10.1136/bcr-2017-221027

Kitch, B. T., Cooper, J. B., Zapol, W. M., Hutter, M. M., Marder, J., Karson, A., & Campbell, E. G. (2008). Handoffs Causing Patient Harm: A Survey of Medical and Surgical House Staff. *The Joint Commission Journal on Quality and Patient Safety*, *34*(10), 563–570d. https://doi.org/10.1016/S1553-7250(08)34071-9

Klein, G. A. (2004). *The Power of Intuition: How To Use Your Gut Feelings To Make Better Decisions At Work*. Currency/Doubleday.

Klein, W. M. (1997). Objective standards are not enough: Affective, self-evaluative, and behavioral responses to social comparison information. *Journal of Personality and Social Psychology*, *72*(4), 763–774. https://doi.org/10.1037/0022-3514.72.4.763

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

Laxmisan, A., Hakimzada, F., Sayan, O. R., Green, R. A., Zhang, J., & Patel, V. L. (2007). The multitasking clinician: Decision-making and cognitive demand during and after team handoffs in emergency care. *International Journal of*

Medical Informatics, *76*(11–12), 801–811. https://doi.org/10.1016/j.ijmedinf.2006.09.019

Levis, J. T., & Garmel, G. M. (2009). *Clinical Emergency Medicine Casebook*. Cambridge University Press.

Long, B. (2015, November 18). *ED Handoffs – The problem and what we can do to improve*. EmDOCs.Net - Emergency Medicine Education. https://www.emdocs.net/ed-handoffs-the-problem-and-what-we-can-do-to-improve/

Ludolph, R., & Schulz, P. J. (2018). Debiasing Health-Related Judgments and Decision Making: A Systematic Review. *Medical Decision Making*, *38*(1), 3–13. https://doi.org/10.1177/0272989X17716672

Lutfey, K. E., Link, C. L., Grant, R. W., Marceau, L. D., & McKinlay, J. B. (2009). Is certainty more important than diagnosis for understanding race and gender disparities?: An experiment using coronary heart disease and depression case vignettes. *Health Policy*, *89*(3), 279–287. https://doi.org/10.1016/j.healthpol.2008.06.007

Makary, M. A., & Daniel, M. (2016). Medical error—The third leading cause of death in the US. *BMJ*, *353*, i2139. https://doi.org/10.1136/bmj.i2139

Mamede, S., Splinter, T. A. W., van Gog, T., Rikers, R. M. J. P., & Schmidt, H. G. (2012). Exploring the role of salient distracting clinical features in the emergence of diagnostic errors and the mechanisms through which reflection counteracts

mistakes. *BMJ Quality & Safety*, *21*(4), 295–300. https://doi.org/10.1136/bmjqs-2011-000518

Mamykina, L., Jiang, S., Collins, S. A., Twohig, B., Hirsh, J., Hripcsak, G., Stanley Hum, R., & Kaufman, D. R. (2016). Revealing structures in narratives: A mixed-methods approach to studying interdisciplinary handoff in critical care. *Journal of Biomedical Informatics*, *62*, 117–124. https://doi.org/10.1016/j.jbi.2016.03.025

Maragh-Bass, A. C., Fields, J. C., McWilliams, J., & Knowlton, A. R. (2017). Challenges and Opportunities to Engaging Emergency Medical Service Providers in Substance Use Research: A Qualitative Study. *Prehospital and Disaster Medicine*, *32*(2), 148–155. https://doi.org/10.1017/S1049023X16001424

Marcum, J. A. (2012). An integrated model of clinical reasoning: Dual-process theory of cognition and metacognition. *Journal of Evaluation in Clinical Practice*, *18*(5), 954–961. https://doi.org/10.1111/j.1365-2753.2012.01900.x

McCaig, L. F., & Nawar, E. W. (2006). National Hospital Ambulatory Medical Care Survey: 2004 emergency department summary. *Advance Data*, *372*, 1–29.

Meyer, A. N. D., Payne, V. L., Meeks, D. W., Rao, R., & Singh, H. (2013). Physicians' diagnostic accuracy, confidence, and resource requests: A vignette study. *JAMA Internal Medicine*, *173*(21), 1952–1958. https://doi.org/10.1001/jamainternmed.2013.10081

Moerbe, M., & Kelemen, A. (2014). Turning electronic health record data into meaningful information using SQL and nursing informatics. *Computers,*

Informatics, Nursing, *32*(8), 366–377. https://doi.org/10.1097/CIN.0000000000000079

Mohan, D., Farris, C., Fischhoff, B., Rosengart, M. R., Angus, D. C., Yealy, D. M., Wallace, D. J., & Barnato, A. E. (2017). Efficacy of educational video game versus traditional educational apps at improving physician decision making in trauma triage: Randomized controlled trial. *BMJ*, *359*, j5416. https://doi.org/10.1136/bmj.j5416

Montemayor, E. T., Long, B., Pfaff, J. A., & Moore, G. P. (2018). Patient with a Subarachnoid Headache. *Clinical Practice and Cases in Emergency Medicine*, *2*(3). https://doi.org/10.5811/cpcem.2018.5.38417

Mull, N., Reilly, J. B., & Myers, J. S. (2015). An elderly woman with 'heart failure': Cognitive biases and diagnostic error. *Cleveland Clinic Journal of Medicine*, *82*(11), 745–753. https://doi.org/10.3949/ccjm.82a.14087

Mumma, G. H., & Wilson, S. B. (1995). Procedural debiasing of primacy/anchoring effects in clinical-like judgments. *Journal of Clinical Psychology*, *51*(6), 841–853. https://doi.org/10.1002/1097-4679(199511)51:6<841::AID-JCLP2270510617>3.0.CO;2-K

Mussweiler, T., & Englich, B. (2005). *Organ. Behav. Hum. Decis. Process.* https://paperpile.com/view/38264ae4-d6b8-0788-b8e2-63fa9fb46c9f

Mussweiler, T., & Englich, B. (2005). Subliminal anchoring: Judgmental consequences and underlying mechanisms. *Organizational Behavior and Human Decision Processes*, *98*(2), 133–143. https://doi.org/10.1016/j.obhdp.2004.12.002

Norman, D. A. (1995). THE PSYCHOPATHOLOGY OF EVERYDAY THINGS. In R. M. Baecker, J. Grudin, W. A. S. Buxton, & S. Greenberg (Eds.), *Readings in Human–Computer Interaction* (pp. 5–21). Morgan Kaufmann. https://doi.org/10.1016/B978-0-08-051574-8.50006-6

Norman, G., Young, M., & Brooks, L. (2007). Non-analytical models of clinical reasoning: The role of experience. *Medical Education*, *41*(12), 1140–1145. https://doi.org/10.1111/j.1365-2923.2007.02914.x

Ogdie, A. R., Reilly, J. B., Pang, W. G., Keddem, S., Barg, F. K., Von Feldt, J. M., & Myers, J. S. (2012). Seen through their eyes: Residents' reflections on the cognitive and contextual components of diagnostic errors in medicine. *Academic Medicine : Journal of the Association of American Medical Colleges*, *87*(10), 1361–1367. https://doi.org/10.1097/ACM.0b013e31826742c9

Okafor, N., Payne, V. L., Chathampally, Y., Miller, S., Doshi, P., & Singh, H. (2016). Using voluntary reports from physicians to learn from diagnostic errors in emergency medicine. *Emerg Med J*, *33*(4), 245–252. https://doi.org/10.1136/emermed-2014-204604

Okuda, Y., & Nelson, B. P. (2009). *Emergency Medicine Oral Board Review Illustrated*. Cambridge University Press.

O'Sullivan, E. D., & Schofield, S. J. (2019). A cognitive forcing tool to mitigate cognitive bias – a randomised control trial. *BMC Medical Education*, *19*(1), 12. https://doi.org/10.1186/s12909-018-1444-3

O'Toole, Hepps Jennifer, Starmer Amy J., Patel Shilpa J., Rosenbluth Glenn, Calaman Sharon, Campos Maria-Lucia, Lopreiato Joseph O., Schnipper Jeffrey L., Sectish Theodore C., Srivastava Rajendu, West Daniel C., Landrigan Christopher P., Spector Nancy D., Yu Clifton E., & null null. (2020). I-PASS Mentored Implementation Handoff Curriculum: Frontline Provider Training Materials. *MedEdPORTAL*, *16*. https://doi.org/10.15766/mep_2374-8265.10912

Parekh, P., Oldfield, E. C., & Marik, P. E. (2013). Progressive outer retinal necrosis: A missed diagnosis and a blind, young woman. *Case Reports*, *2013*, bcr2013009333. https://doi.org/10.1136/bcr-2013-009333

Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of Persuasion. In R. E. Petty & J. T. Cacioppo (Eds.), *Communication and Persuasion: Central and Peripheral Routes to Attitude Change* (pp. 1–24). Springer New York. https://doi.org/10.1007/978-1-4612-4964-1_1

Polic, M. (2009). Decision Making: Between Rationality and Reality. *Interdisciplinary Description of Complex Systems - Scientific Journal*, *7*, 78–89.

Reilly, J. B., Ogdie, A. R., Von Feldt, J. M., & Myers, J. S. (2013). Teaching about how doctors think: A longitudinal curriculum in cognitive bias and diagnostic error for residents. *BMJ Quality & Safety*, *22*(12), 1044–1050. https://doi.org/10.1136/bmjqs-2013-001987

Richie, M., & Josephson, S. A. (2018). Quantifying Heuristic Bias: Anchoring, Availability, and Representativeness. *Teaching and Learning in Medicine*, *30*(1), 67–75. https://doi.org/10.1080/10401334.2017.1332631

Riesenberg, L. A., Leitzsch, J., & Little, B. (2009). *Systematic Review of Handoff Mnemonics Literature—Lee Ann Riesenberg, Jessica Leitzsch, Brian W. Little, 2009*. https://journals.sagepub.com/doi/abs/10.1177/1062860609332512

Rottenstreich, Y., & Tversky, A. (1997). Unpacking, Repacking, and Anchoring: Advances in Support Theory. [Miscellaneous Article]. *Psychological Review*, *104*(2), 406–415.

Royce, C. S., Hayes, M. M., & Schwartzstein, R. M. (2019). Teaching Critical Thinking: A Case for Instruction in Cognitive Biases to Reduce Diagnostic Errors and Improve Patient Safety. *Academic Medicine: Journal of the Association of American Medical Colleges*, *94*(2), 187–194. https://doi.org/10.1097/ACM.0000000000002518

Salas, E., Cooke, N. J., & Rosen, M. A. (2008). On teams, teamwork, and team performance: Discoveries and developments. *Human Factors*, *50*(3), 540–547. https://doi.org/10.1518/001872008X288457

Salas, E., & Klein, G. A. (Eds.). (2001). *Linking Expertise and Naturalistic Decision Making*. Psychology Press. https://doi.org/10.4324/9781410604200

Saposnik, G., Redelmeier, D., Ruff, C. C., & Tobler, P. N. (2016). Cognitive biases associated with medical decisions: A systematic review. *BMC Medical Informatics and Decision Making*, *16*(1), 138. https://doi.org/10.1186/s12911-016-0377-1

Schiff, G. D., Hasan, O., Kim, S., Abrams, R., Cosby, K., Lambert, B. L., Elstein, A. S., Hasler, S., Kabongo, M. L., Krosnjar, N., Odwazny, R., Wisniewski, M. F., &

McNutt, R. A. (2009). Diagnostic Error in Medicine: Analysis of 583 Physician-Reported Errors. *Archives of Internal Medicine*, *169*(20), 1881–1887. https://doi.org/10.1001/archinternmed.2009.333

Schiff, G. D., Kim, S., Abrams, R., Cosby, K., Lambert, B., Elstein, A. S., Hasler, S., Krosnjar, N., Odwazny, R., Wisniewski, M. F., & McNutt, R. A. (2005). *Diagnosing Diagnosis Errors: Lessons From A Multi-Institutional Collaborative Project*. AGENCY FOR HEALTHCARE RESEARCH AND QUALITY ROCKVILLE MD. https://apps.dtic.mil/docs/citations/ADA434037

Schraagen, J. M., & van de Ven, J. G. M. (2008). Improving Decision Making in Crisis Response through Critical Thinking Support. *Journal of Cognitive Engineering and Decision Making*, *2*(4), 311–327. https://doi.org/10.1518/155534308X377801

Sentinel Event | Joint Commission. (n.d.). Retrieved May 12, 2019, from http://www.jointcommission.org/sentinel_event.aspx

Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, *69*(1), 99–118. https://doi.org/10.2307/1884852

Singh, H. (2014). Editorial: Helping Health Care Organizations to Define Diagnostic Errors as Missed Opportunities in Diagnosis. *Joint Commission Journal on Quality and Patient Safety*, *40*(3), 99–101. https://doi.org/10.1016/S1553-7250(14)40012-6

Singh, H., Schiff, G. D., Graber, M. L., Onakpoya, I., & Thompson, M. J. (2017). The global burden of diagnostic errors in primary care. *BMJ Qual Saf*, *26*(6), 484–494. https://doi.org/10.1136/bmjqs-2016-005401

Singh, H., & Sittig, D. F. (2015). Advancing the science of measurement of diagnostic errors in healthcare: The Safer Dx framework. *BMJ Quality & Safety*, *24*(2), 103–110. https://doi.org/10.1136/bmjqs-2014-003675

Sperber, D., & Wilson, D. (n.d.). *(Blackwell's, Oxford)*.

Starmer, A. J., Spector, N. D., Srivastava, R., West, D. C., Rosenbluth, G., Allen, A. D., Noble, E. L., Tse, L. L., Dalal, A. K., Keohane, C. A., Lipsitz, S. R., Rothschild, J. M., Wien, M. F., Yoon, C. S., Zigmont, K. R., Wilson, K. M., O'Toole, J. K., Solan, L. G., Aylor, M., … Landrigan, C. P. (2014). Changes in Medical Errors after Implementation of a Handoff Program. *New England Journal of Medicine*, *371*(19), 1803–1812. https://doi.org/10.1056/NEJMsa1405556

Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology*, *73*(3), 437–446. https://doi.org/10.1037/0022-3514.73.3.437

Sullivan, C., Inboriboon, P. C., & Bridgford, S. (2015). Transitions in Care: When Words Can Save Lives. *Missouri Medicine*, *112*(3), 197–201.

Tews, M. C., Liu, J. M., & Treat, R. (2012). Situation-Background-Assessment-Recommendation (SBAR) and Emergency Medicine Residents' Learning of Case Presentation Skills. *Journal of Graduate Medical Education*, *4*(3), 370–373. https://doi.org/10.4300/JGME-D-11-00194.1

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Tversky, A., & Kahneman, D. (1985). The Framing of Decisions and the Psychology of Choice. In G. Wright (Ed.), *Behavioral Decision Making* (pp. 25–41). Springer US. https://doi.org/10.1007/978-1-4613-2391-4_2

Tversky, A., & Kahneman, D. (1986). Rational Choice and the Framing of Decisions. *The Journal of Business*, *59*(4), S251–S278. JSTOR.

Tversky, A., & Koehler, D. J. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review*, *101*(4), 547–567. https://doi.org/10.1037//0033-295X.101.4.547

van Geene, K., de Groot, E., Erkelens, C., & Zwart, D. (2016a). Raising awareness of cognitive biases during diagnostic reasoning. *Perspectives on Medical Education*, *5*(3), 182–185.

van Geene, K., de Groot, E., Erkelens, C., & Zwart, D. (2016b). Raising awareness of cognitive biases during diagnostic reasoning. *Perspectives on Medical Education*, *5*(3), 182–185. https://doi.org/10.1007/s40037-016-0274-4

Walmsley, S., & Gilbey, A. (2017). Debiasing visual pilots' weather-related decision making. *Applied Ergonomics*, *65*, 200–208.

Wansink. (1998). *An Anchoring and Adjustment Model of Purchase Quantity Decisions on JSTOR*. https://www.jstor.org/stable/3151931?sid=primo&origin=crossref&seq=1#metadata_info_tab_contents

Wason, P. C., & Evans, J. S. B. (1974). Dual processes in reasoning? *Cognition*, *3*(2), 141–154.

Wershofen, B., Heitzmann, N., Beltermann, E., & Fischer, M. R. (2016). Fostering interprofessional communication through case discussions and simulated ward rounds in nursing and medical education: A pilot project. *GMS Journal for Medical Education*, *33*(2). https://doi.org/10.3205/zma001027

Zsambok, C. E., & Klein, G. (2014). *Naturalistic Decision Making*. Psychology Press.

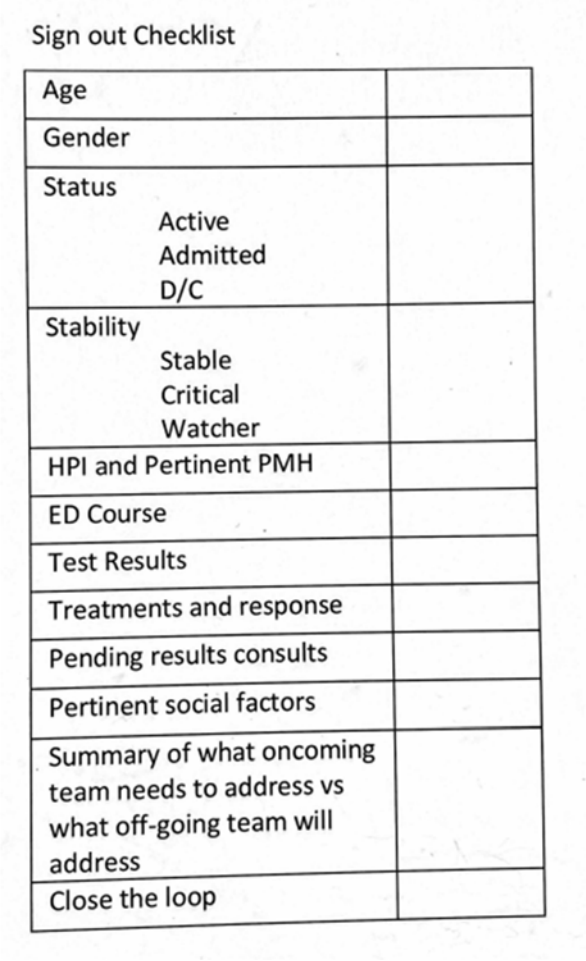
Zwaan, L., Monteiro, S., Sherbino, J., Ilgen, J., Howey, B., & Norman, G. (2017). Is bias in the eye of the beholder? A vignette study to assess recognition of cognitive biases in clinical case workups. *BMJ Quality & Safety*, *26*(2), 104–110. https://doi.org/10.1136/bmjqs-2015-005014

Appendix A: EM Physician Interview Guide

#	Question
1	How strongly do you agree that sign outs are intended to transfer patient information and a mental model of the case? 1=strongly disagree, 2= somewhat disagree 3 = neither agree or disagree, 4=somewhat agree, 5= strongly agree, 0 = undecided
	If you disagree, what do you feel is the purpose of sign outs?
2.	In your experience during a shift change did you receive a differential diagnosis for a patient that changed from what you received?
3.	Thinking about a recent sign out, what was good or bad, what ‹
4	What are you taking into account when you get a sign out? (e.g. a story)
5	When you receive a sign out are you considering the evidence they have presented at the time or later and how do you package it e.g. by the patient, task, source?
6	Does the way the information is delivered or who gives it make a difference in your consideration or is it solely the data?
7	When you receive a sign out do you what do focus on e.g. data about the patient or differential or anything else?
8	What amount of data or granularity of the data for each patient do you require? Do you have specific content that you require when you get a sign out?
9	When you have received a sign out what are the next steps that you do?
10	What do you feel affects a sign out? What are the risks of sign out?
11	Does the stage in the process the patient is in affect how you proceed?
12	There have been efforts to standardize sign-out e.g. hand off tools. What do you think of are the benefits of these tools?
13	Do you see the patients transferred to you in a sign out?
14	Do you think there is a chance of getting an anchoring bias from the differential provided during the sign out?
15	Regardless of whether anything went wrong or almost went wrong and thinking about what should be included in a sign out, is there anything about the sign out that you received that you think should have been better?

Figure 13: EM Physicians Interview Guide

Appendix B: UMB Handoff Checklist Protocol



Sign out Checklist

Age	
Gender	
Status Active Admitted D/C	
Stability Stable Critical Watcher	
HPI and Pertinent PMH	
ED Course	
Test Results	
Treatments and response	
Pending results consults	
Pertinent social factors	
Summary of what oncoming team needs to address vs what off-going team will address	
Close the loop	

Figure 14: Sign Out Checklist for University of Maryland Baltimore

Appendix C: Example of Edits from SME Review of Constructs

Revisions to Control Case Vignettes from Feedback from Two SME EM Physicians

Stage	Case	Initial SO Narrative 08/08	Edits from DR 08/30	Edits from BK 08/31
Early	#1 Control Pneumothorax	In room 4 we have a 27 year old male, who presented with chest pain on his right side and shortness of breath, which is worse on inspiration. The pain started four days ago after a bout of severe coughing though he denies alcohol or drug use. His exam was normal but did show absent breath sounds in the right lung field and he is tall and thin in appearance. His labs and a chest x-ray have been ordered but we are waiting on the results.	NO CHANGES	In room 4 we have a 47 year old male with a history of COPD, who presented with chest pain on his right side and shortness of breath, which is worse on inspiration. The pain started four days ago after a bout of severe coughing. He is a smoker but denies alcohol or drug use. His exam demonstrated wheezing and absent breath sounds in the right lung field and he is tall and thin in appearance. His labs and a chest x-ray have been ordered but we are waiting on the results.
Middle	# 4 Control Stroke	In room 5 we've got a 66 year old male who was brought in the by EMS who were called by his wife because he was unable to get out of bed when he woke after his midday nap because of right side paralysis. The patient shows right sided facial droop and states numbness in right arm and leg. Patient states he's a moderate smoker and social alcohol use and has a family history of cardiovascular disease and stroke. We have the initial CAT scan and the wet read shows no obvious hemorrhage but the radiology final read to determine if they want to do a TPA. His chest x-ray and labs were all normal	In room 5 we've got a 66 year old male who was brought in the by EMS who were called by his wife because he was unable to get out of bed when he woke after his midday nap because of right side paralysis. The patient shows right sided facial droop and states numbness in right arm and leg. Patient states he's a moderate smoker and social alcohol use and has a family history of cardiovascular disease and stroke. We have the initial CAT scan and the wet read shows no obvious hemorrhage but the radiology final read. Neurology to determine if they want to do a TPA. His chest x-ray and labs were all normal.	In room 6 we've got a 66 year old male who was brought in the by EMS who were called by his wife because he was unable to get out of bed when he woke after his midday nap because of right side paralysis. The patient shows right sided facial droop and states numbness in right arm and leg. Patient states he's a moderate smoker and social alcohol use and has a family history of cardiovascular disease and stroke. The initial plain CAT scan was fine. Waiting on the CT angiogram of head and neck ~~and the wet read shows no obvious hemorrhage but~~ need final read from radiology**. Neurology to determine if they want to do a TPA.** His chest x-ray and labs were all normal

Figure 15: Example of Edits from SME Review of Study Constructs - following SME review of the cases in the SME review survey.

The sign out narrative for each of the cases was reviewed with SMEs and edits were made to refine the information so that the cases would be more likely to align with the intended stages for the case as per the experimental design. The changes suggested by the SMEs are shown in red or blue. This was an iterative process conducted with multiple SMEs to ensure the greatest levels of consensus about the case stages was achieved.

Control Case 1: Pneumothorax Sign Out Narrative - EARLY

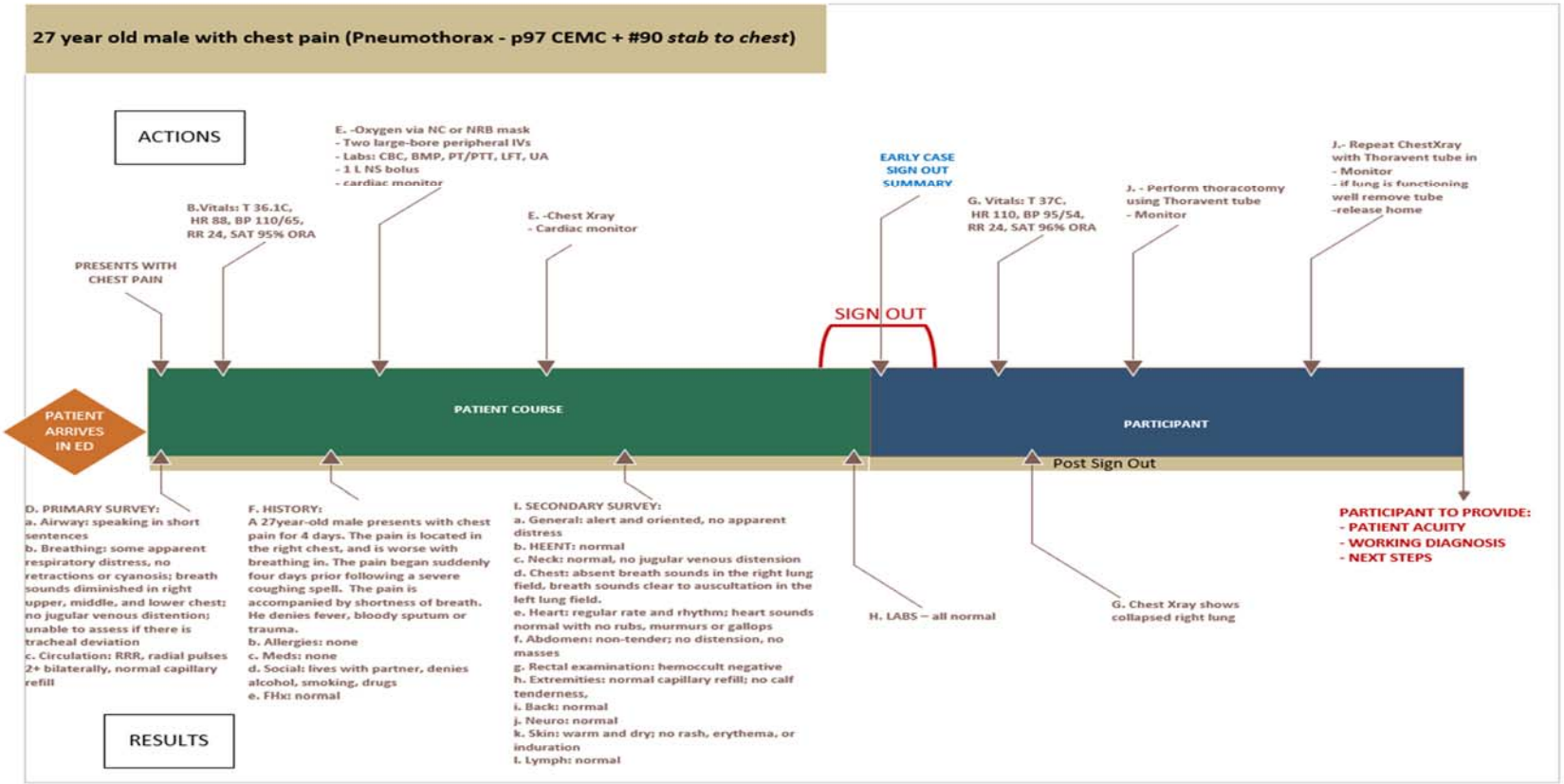


Figure 16: Control Case 1: Pneumothorax (Early)

Control Case 1: Pneumothorax Sign Out Narrative - EARLY

SIGN OUT NARRATIVE

In room 4 we have a 47 year old[1] male[2], who presented with chest pain[3] on his right side[4] and shortness of breath[5], which is worse on inspiration[6]. The pain started four days ago[7] after a bout of severe coughing[8]. He is a smoker[9] though he denies alcohol[10] or drug use[11]. His exam was normal[12] but did show absent breath sounds[13] in the right lung field[14] and he is a tall and thin in appearance[15]. His labs are unremarkable[16] and a chest x-ray has been ordered[17], but we are waiting on the results[18].

Control Case 2: Appendicitis – EARLY

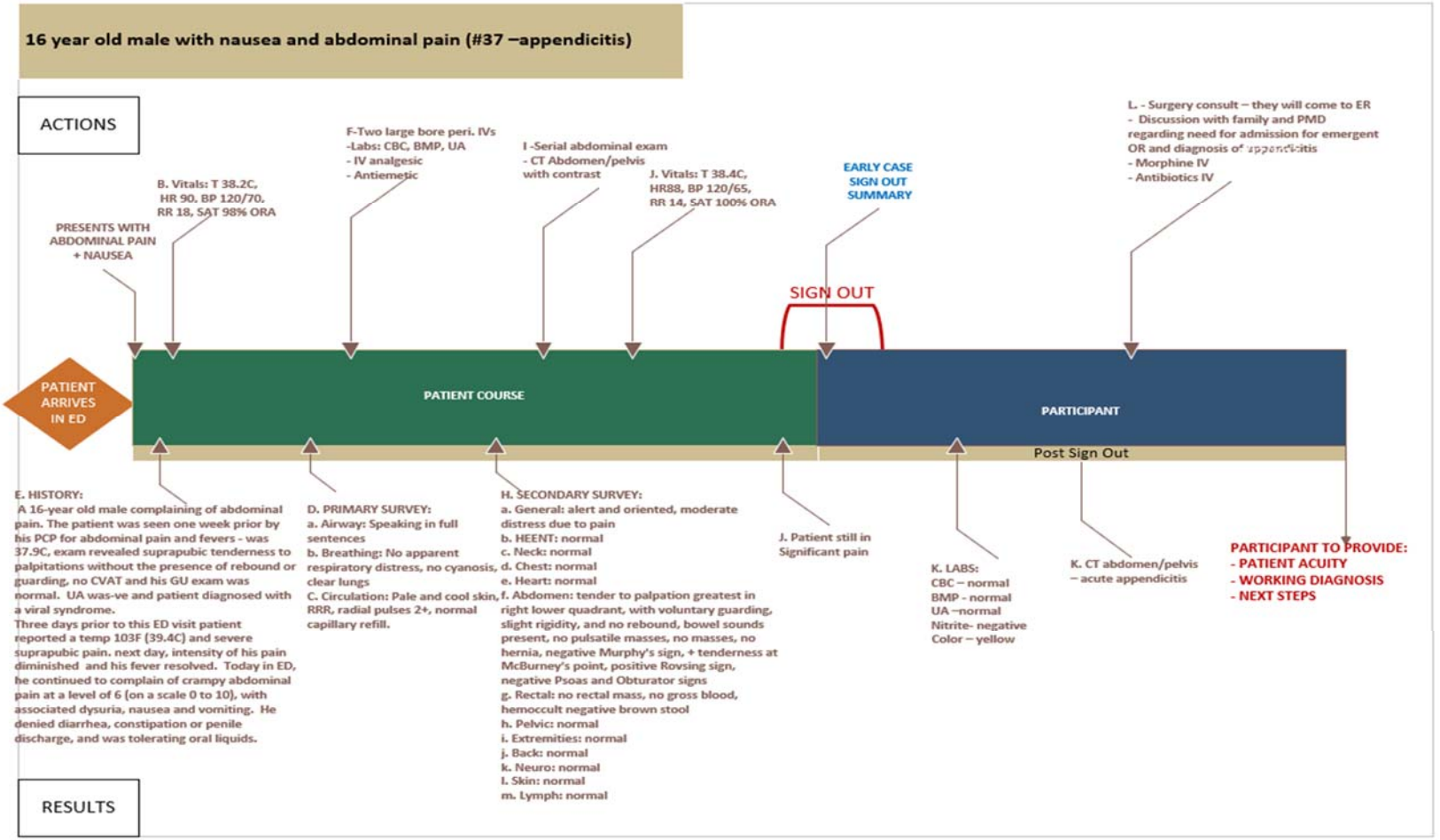


Figure 17: Control Case 2: Appendicitis (Early)

Control Case 2: Appendicitis – EARLY

SIGN OUT NARRATIVE

In room 7 we have a young 16-year old[1] male[2] who is here with nausea[3] and severe abdominal pain[4]. He previously presented to his PCP one week[5] earlier with fever[6], cramps[7] and pain[8] but the intensity[9]and fever diminished[10] the following day and with a negative urinalysis for infection[11] and he was diagnosed with a viral syndrome.[12] He is in the ER with nausea[13] and return of the suprapubic pain[14] and his temperature is raised[15]. He's received IV fluids,[16] pain meds[17] and an antiemetic.[18] Also, labs for CBC[19], BMP[20] and Urinalysis[21] have been ordered. His exam did show right lower quadrant tenderness[22] but no masses[22] or hernias detected[23] and his rectal exam was negative.[24] A CT of the abdomen and pelvis has been ordered.[25]

Control Case 3: MVC -Fractured ulna- LATE

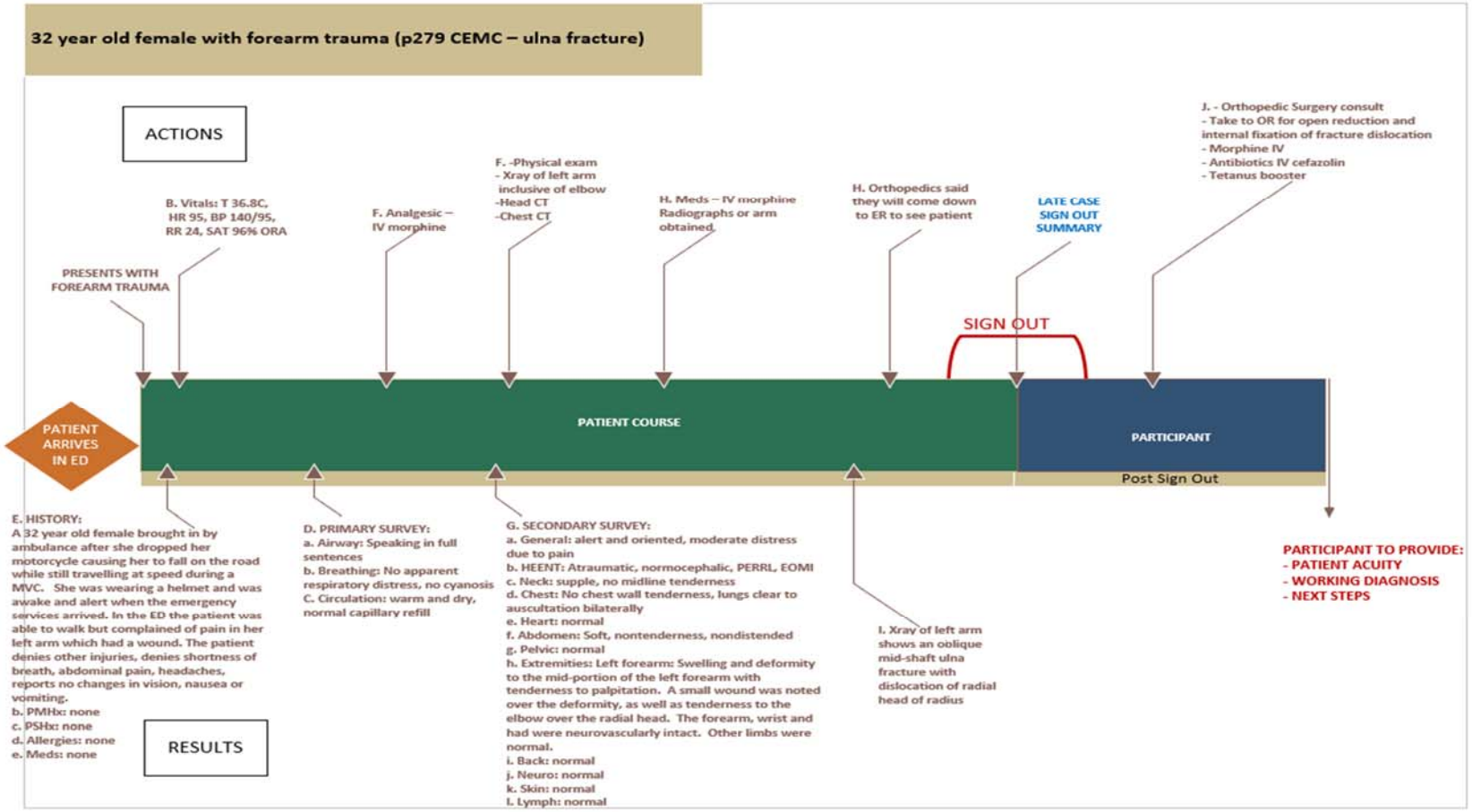


Figure 18: Control Case 3: MVC – Fracture ulna (Late) – motor vehicle crash with arm trauma

Control Case 3: MVC -Fractured ulna- LATE

SIGN OUT NARRATIVE

So, in room 2 we have a 32 year old[1] female[2] who was brought by EMT after road incident[3] involving her motorcycle and another vehicle which forced her to drop her bike.[4] The patient was wearing a helmet[5] at the time and was awake[6] and alert[7] when brought in but did complain of pain in their left arm[8] and elbow.[9] There is a small laceration on the left forearm[10] but physical exam was otherwise not significant[11]. Also, patient denies any changes in vision,[12] nausea[13] or vomiting,[14] shortness of breath[15] or any other injuries.[16] Patient has received some IV morphine.[17] The pan scans were negative,[18] labs are all normal[19] and they just have this fracture of the left ulna.[20] Ortho will come down to see the patient in the ER,[21] so you'll have to follow up with them.[22]

Control Case 4: Stroke – MIDDLE

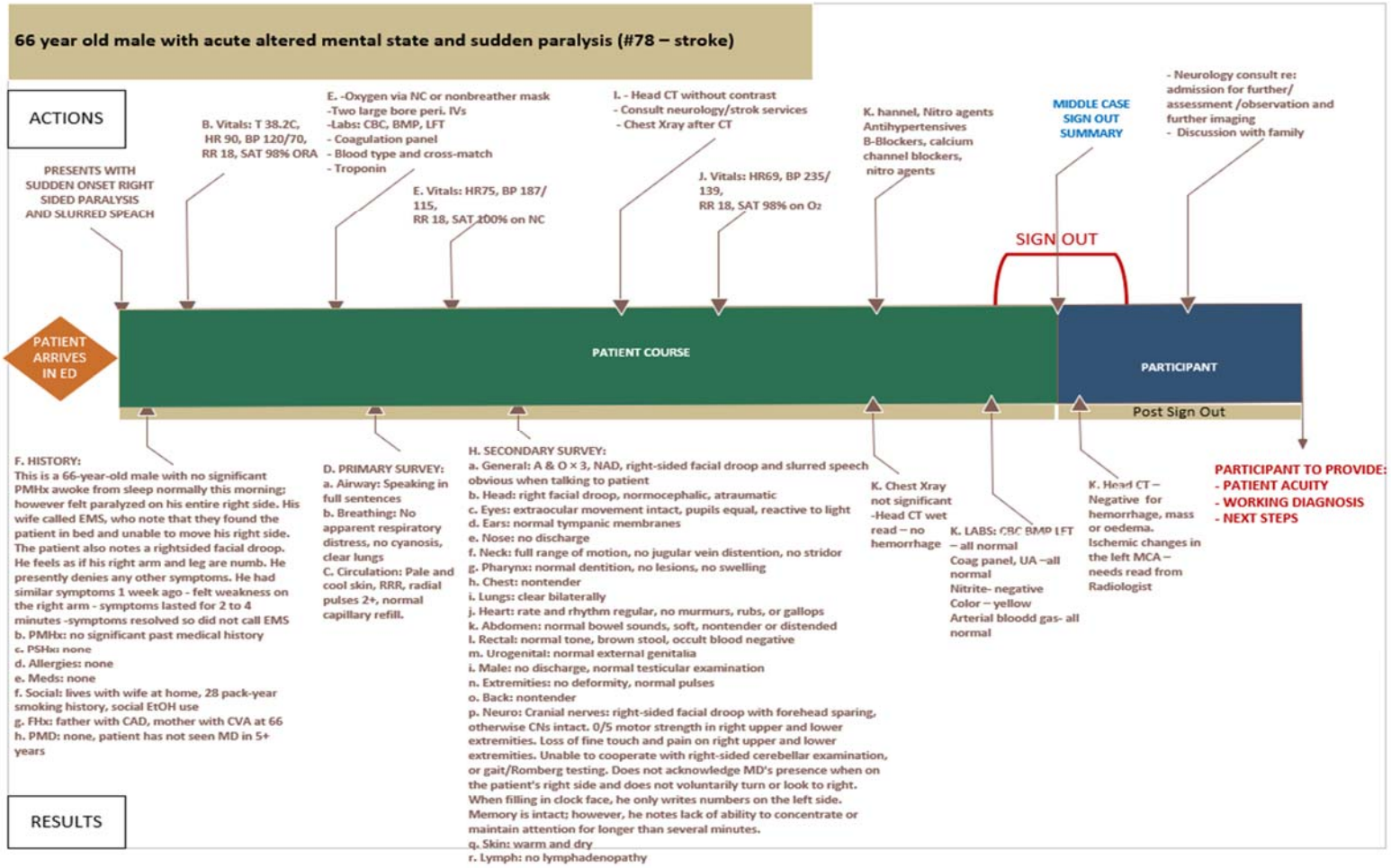


Figure 19: Control Case 4: Stroke (Middle)

Control Case 4: AMS - Stroke – MIDDLE

SIGN OUT NARRATIVE

In the room 6 we've got a 66-year old[1] male[2] who was brought in the by EMS who were called by his wife[3] because he was unable to get out of bed[4] when he woke after his midday nap because of right side paralysis.[5] The patient shows right sided facial droop[6] and states numbness in right arm[7] and leg.[8] Patient states he's a moderate smoker[9] and social alcohol[10] use and has a family history of cardiovascular disease[11] and stroke.[12] The initial plain CAT scan was fine.[13] Waiting on the CT angiogram of head and neck and [14] and a final read from radiology.[15] Neurology to determine if they want to do a TPA[16]. His chest x-ray[17] and labs were all normal.[18]

Control Case 5: Abdominal pain – DKA + UTI – MIDDLE

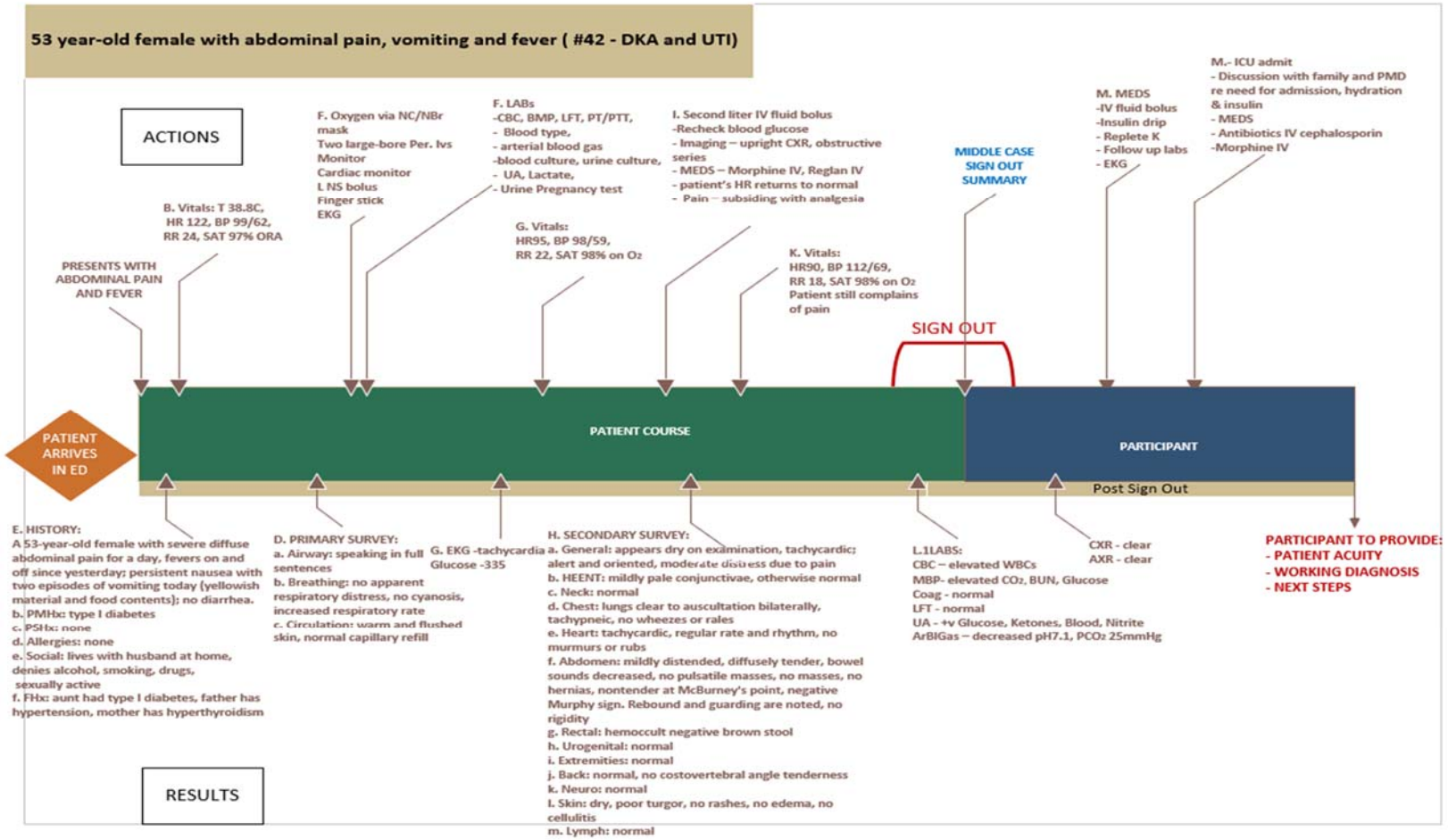


Figure 20: Control Case 5: DKA & UTI (Middle) – diabetic ketoacidosis and urinary tract infection

Control Case 5: Abdominal pain – DKA + UTI – MIDDLE

SIGN OUT NARRATIVE

In room 8 we have 53-year old[1] female[2] who came in with diffuse[3] abdominal pain[4] and vomiting[5] since yesterday[6]. She has type 1 diabetes[7] and denies alcohol[8] or smoking[9]. Her fingerstick was 435,[10] her anion gap was 29 [11] and her EKG shows sinus tachycardia[12]. She's received IV fluids[13]. She's been given IV morphine[14] Zofran[15] and fluids[16]. She was sent for CT scans of abdomen[17] and pelvis[18]. The urinalysis is positive for infection.[19]

Control Case 6: AMS – Hyponatremia -Alcohol intoxication – LATE

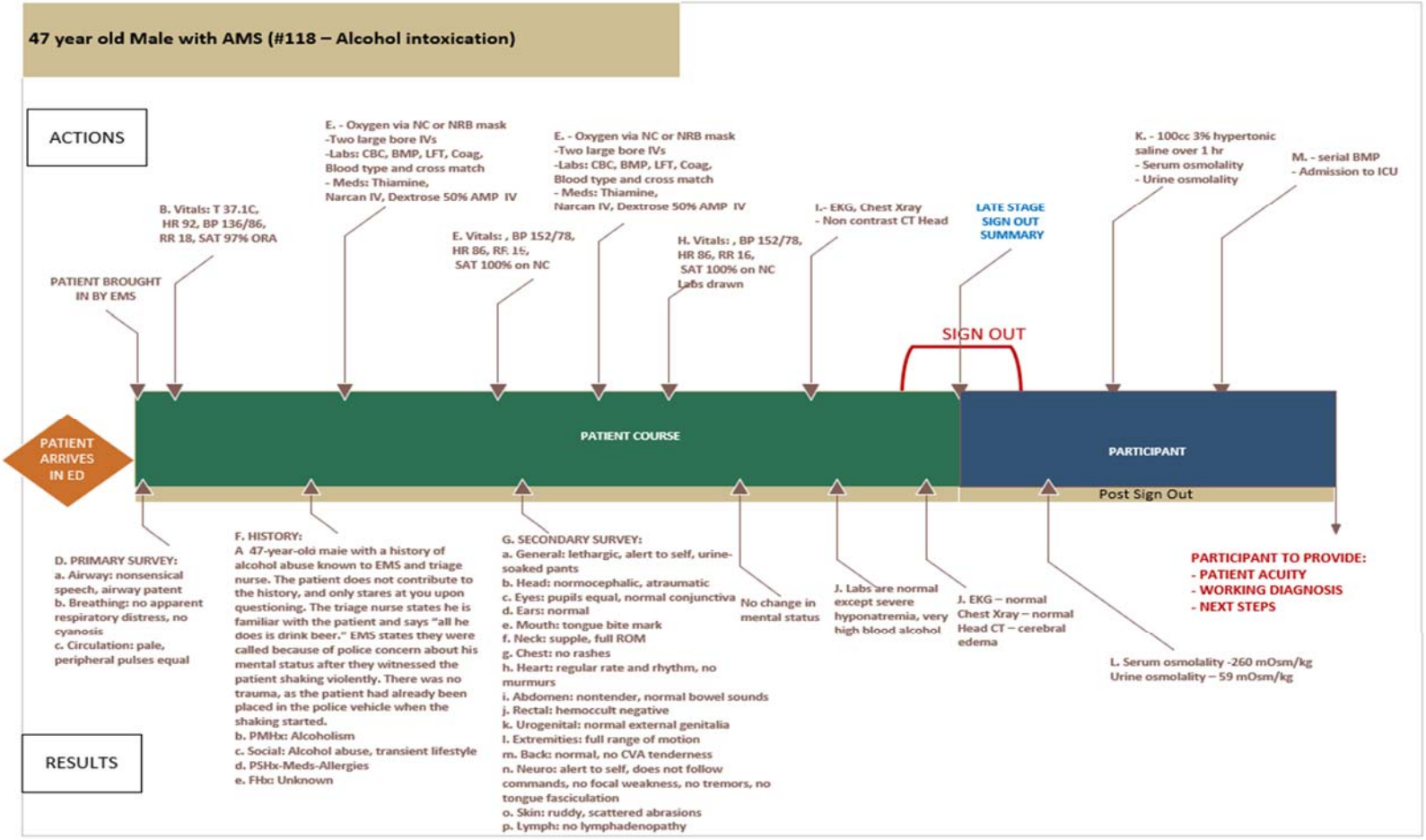


Figure 21: Control Case 6: AMS from Hyponatremia (Late) – hyponatremia from extreme alcohol intoxication

Case 6: AMS – Hyponatremia -Alcohol intoxication – LATE

SIGN OUT NARRATIVE

In room 12 we have a 57-year old[1] male[2] brought in by EMS[3]. This patient has a history of alcohol use[4] but brought in after police concern for his mental status[5] and witnessing him shaking violently[6]. His exam shows no trauma[7] or tenderness[8] and his labs for alcohol[9] and urine toxicology[10] show he was intoxicated[11] and sodium of 110.[12] His chest x-ray[13] and EKG [14] were normal and have a head CT ordered[15]. He's on 2litres of oxygen,[16] he's got thiamine[17] folic acid [18] and a multivitamin[19] but no significant change in mental status so far.[20] He needs to continue with the hypertonic saline[21] and needs to be monitored with follow up labs[22] and admitted to the hospital.[23] The CT head needs to be followed up[24] and after that he can be admitted to ICU.[25]

Control Case 7: Jones Fracture Left Foot – END

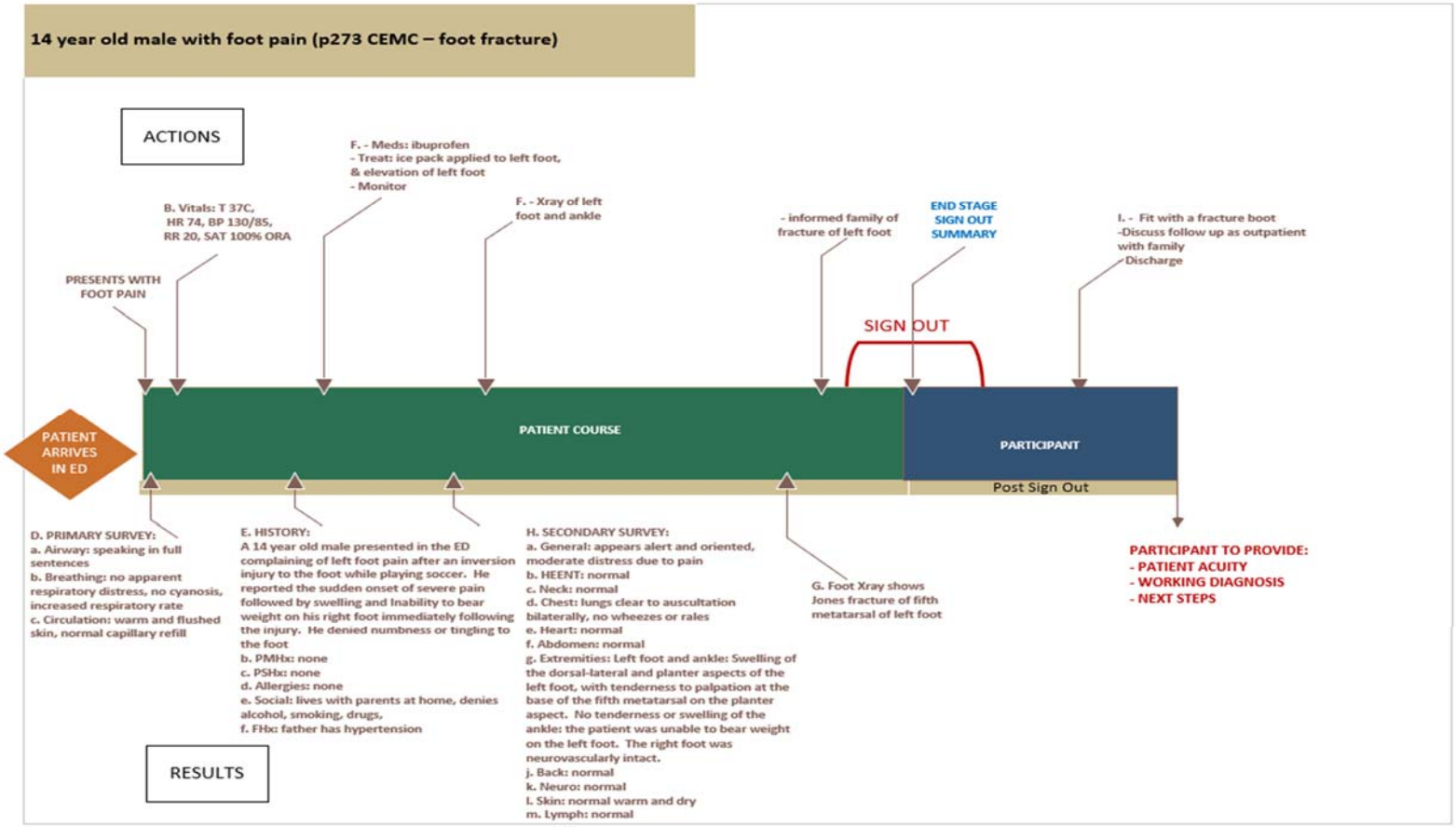


Figure 22: Control Case 7: Jones Fracture – (End) - facture of the metatarsal in the left foot

Control Case 7: Jones Fracture Left Foot – END

SIGN OUT NARRATIVE

In room 10 the case is a young male[1], he's 14[2] and he's got an inversion injury[3] to the left foot[4] while playing soccer. Has pain[5], swelling[6] and he's unable to bear weight[7]. His physical exam is otherwise healthy[8] and no significant medical history[9]. The patient has received Tylenol for pain[10] and foot has been elevated and iced[11] and x-ray shows a Jones facture[12]. Talked with ortho[13]. He's going to be put into a fracture boot[14] and will follow up with them as an outpatient[15].

Stimuli Head Case – Short Sign Out

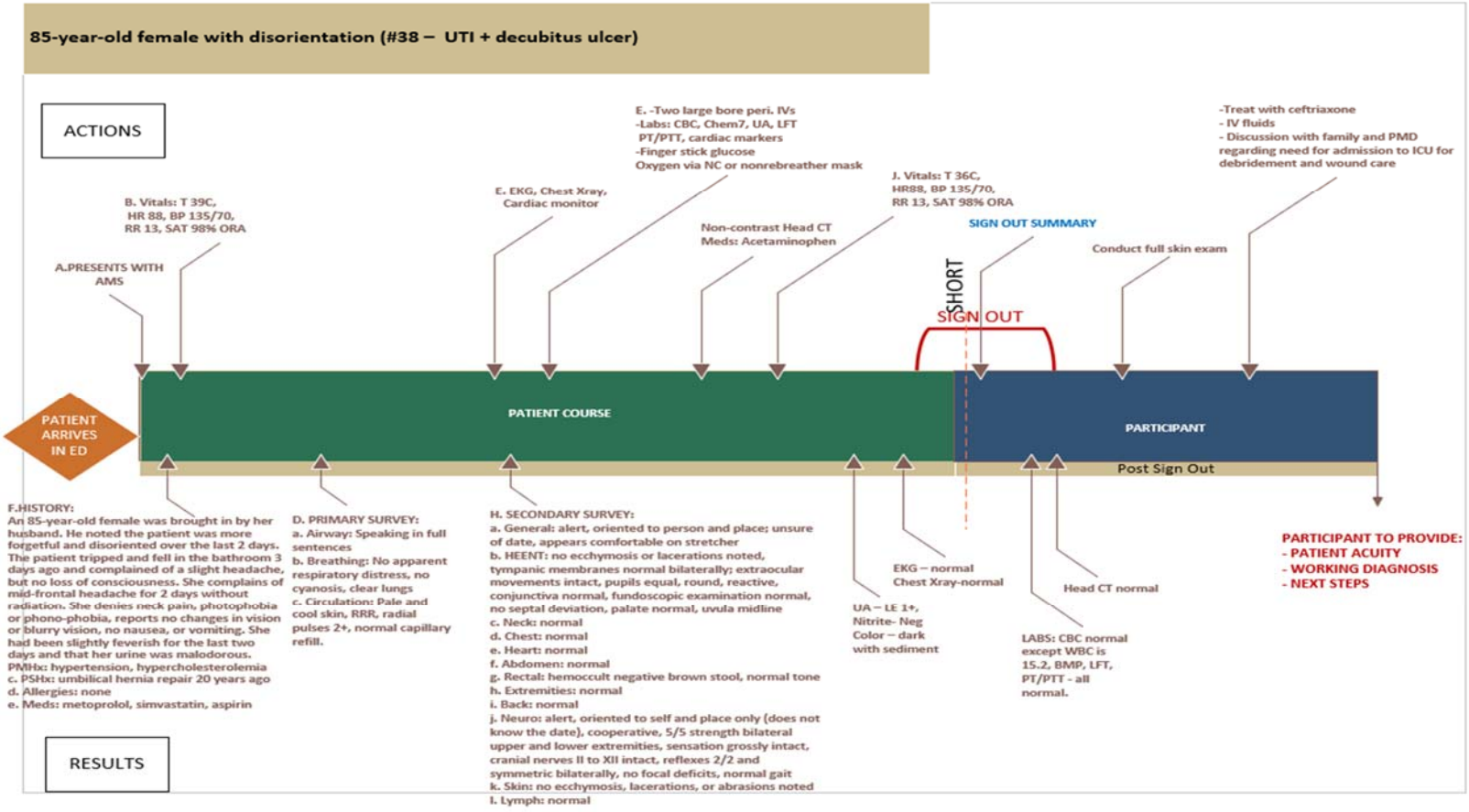


Figure 23: Stimuli Head Case – Short Sign Out

Stimuli Head Case – Long Sign Out

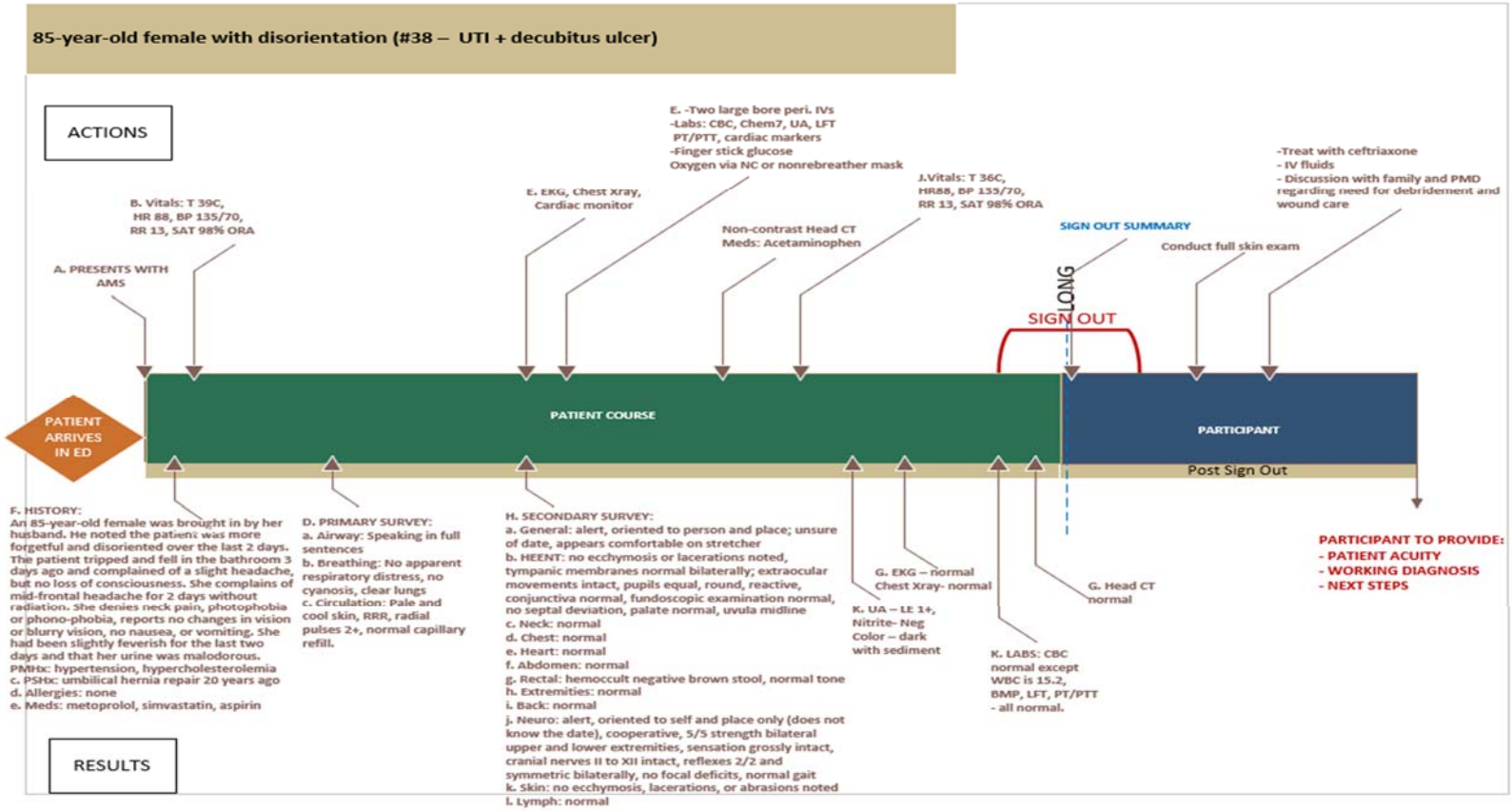


Figure 24: Stimuli Head Case – Long Sign Out

SIGN OUT NARRATIVES FOR STIMULI HEAD CASES

SHORT & IMPLICIT

In room 1 we have an elderly woman[1], in her mid-eighties[2]. She was brought in by her husband[3] because she's been disorientated[4] and forgetful[5] with a slight fever[6] the last two days[7]. She did have a fall[8] in the bathroom 3 days prior[9] but no loss of consciousness[10]. She complains of headache[11] but no neck pain[12] or changes in vision[13] or any nausea or vomiting.[14] She has a history of hypertension[15] and high cholesterol[16] and she's controlled with medication including aspirin.[17] On exam, she was alert and oriented[18] but didn't know the date.[19] She had no focal neuro deficits,[20] including 5/5 strength[21] and normal sensation.[22] Her evaluation showed normal EKG[23] and chest xray[24] was normal. Urine was dark,[25] negative for nitrite[26] and 1+ for leukocyte esterase.[27]

SHORT & EXPLICIT

In room 1 we have an elderly woman[1], in her mid-eighties[2]. She was brought in by her husband[3] because she's been disorientated[4] and forgetful[5] with a slight fever[6] the last two days[7]. She did have a fall[8] in the bathroom 3 days prior[9] but no loss of consciousness[10]. She complains of headache[11] but no neck pain[12] or changes in vision[13] or any nausea or vomiting.[14] She has a history of hypertension[15] and high cholesterol[16] and she's controlled with medication including aspirin.[17] On exam, she was alert and oriented[18] but didn't know the date.[19] She had no focal neuro deficits,[20] including 5/5 strength[21] and normal sensation.[22] Her evaluation showed normal EKG[23] and chest xray[24] was normal. Urine was dark,[25] negative for nitrite[26] and 1+ for leukocyte esterase.[27] Still waiting on her other labs[29] and the head CT.[30]

LONG & IMPLICIT

In room 1 we have an elderly woman[1], in her mid-eighties[2]. She was brought in by her husband[3] because she's been disorientated[4] and forgetful[5] with a slight fever[6] the last two days[7]. She did have a fall[8] in the bathroom 3 days prior[9] but no loss of consciousness[10]. She complains of headache[11] but no neck pain[12] or changes in vision[13] or any nausea or vomiting.[14] She has a history of hypertension[15] and high cholesterol[16] and she's controlled with medication including aspirin.[17] On exam, she was alert and oriented[18] but didn't know the date.[19] She had no focal neuro deficits,[20] including 5/5 strength[21] and normal sensation. [22] Her evaluation showed normal EKG[23] and chest xray[24] and her labs for CBC[25], BMP[26], LFTs[27] and coags[28] came back normal.[29] Urine was dark,[30] negative for nitrite[31] and 1+ for leukocyte esterase.[32] Head CT is normal.[33]

LONG & EXPLICIT

In room 1 we have an elderly woman[1], in her mid-eighties[2]. She was brought in by her husband[3] because she's been disorientated[4] and forgetful[5] with a slight fever[6] the last two days[7]. She did have a fall[8] in the bathroom 3 days prior[9] but no loss of consciousness[10]. She complains of headache[11] but no neck pain[12] or changes in vision[13] or any nausea or vomiting.[14] She has a history of hypertension[15] and high cholesterol[16] and she's controlled with medication including aspirin.[17] On exam, she was alert and oriented[18] but didn't know the date.[19] She had no focal neuro deficits,[20] including 5/5 strength[21] and normal sensation.[22] Her evaluation showed normal EKG[23] and chest xray[24] and her labs for CBC[25], BMP[26], LFTs[27] and coags[28] came back normal.[29] Urine was dark,[30] negative for nitrite[31] and 1+ for leukocyte esterase.[32] Head CT is normal.[33] The patient has a mild UTI [34] so she needs to be checked[35], treated[36] and admitted.[37]

Stimuli Lung Case – Short Sign Out

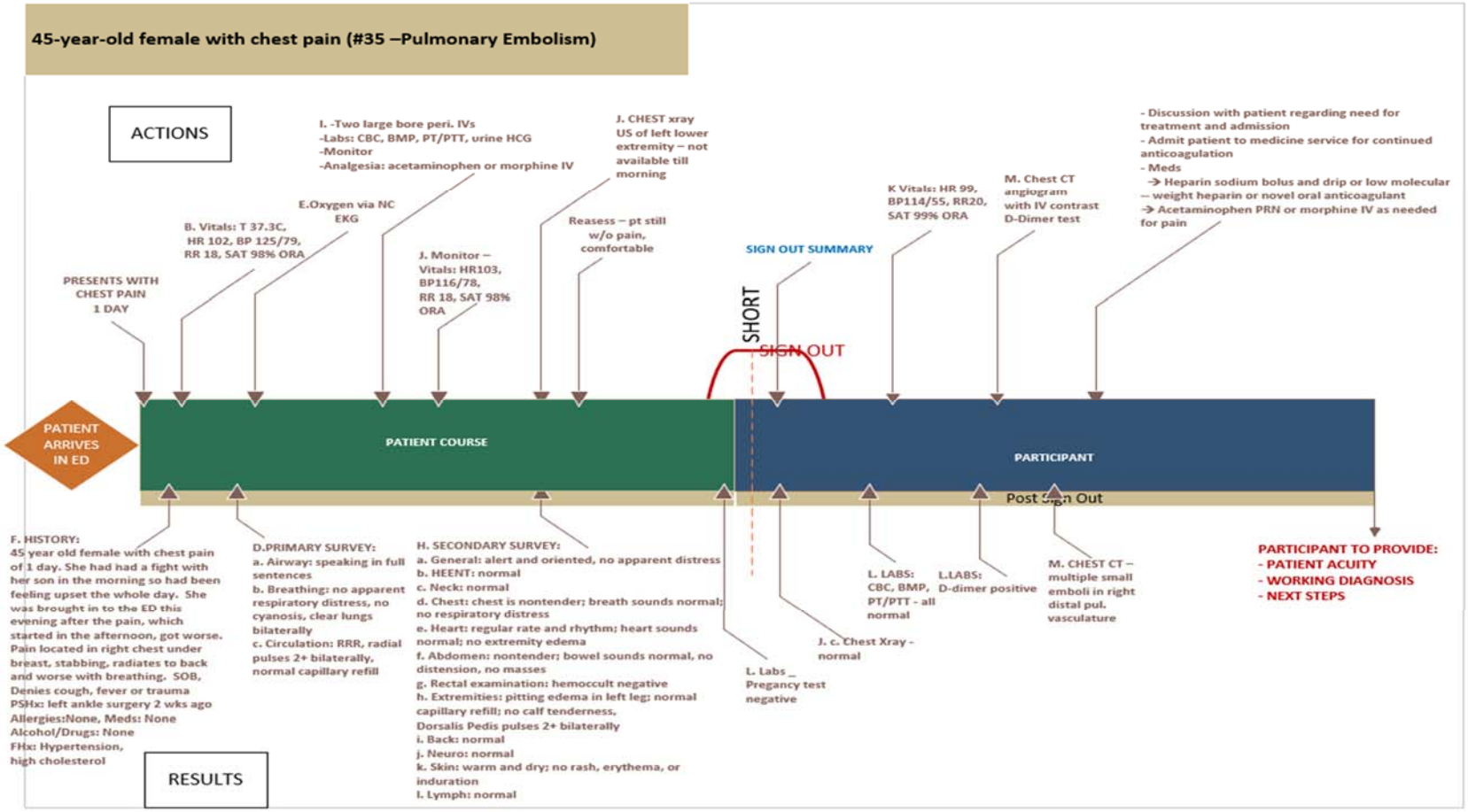


Figure 25: Stimuli Lung Case – Short Sign Out

Stimuli Lung Case – Long Sign Out

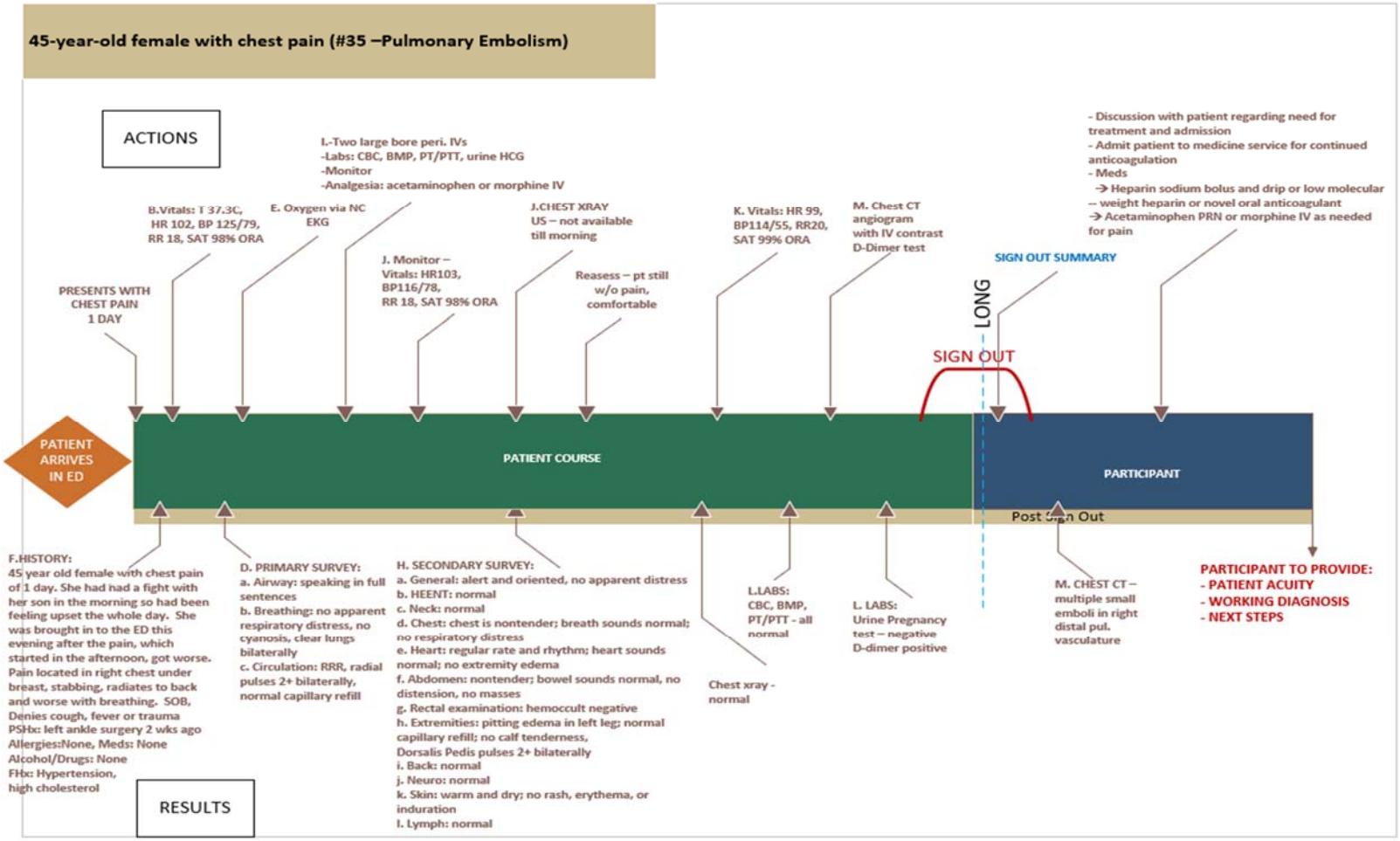


Figure 26: Stimuli Lung Case – Long Sign Out

SIGN OUT NARRATIVES FOR STIMULI LUNG CASES

SHORT & IMPLICIT

In room 3, is a 45-year-old[1] female[2] symptomatic of chest pain[3] under the right breast[4] accompanied by dyspnea[5] for duration of one day.[6] She states the pain started in the afternoon[7] after she had an argument with her son earlier[8] and describes it as stabbing[9] and radiating to her back.[10] Past medical history is not significant.[11] No recent trauma[12] or fever[13], she had ankle surgery 2 weeks ago.[14] Her pain has resolved since arrival, so she thinks it was just stress from the argument. So, in her evaluation, her chest x-ray was clear[15] and EKG showed sinus tachycardia[16] and bedside echo was unremarkable[17] with a grossly normal EF.[18] Her pregnancy test was negative[19] but her other labs aren't back yet[20].

SHORT & EXPLICIT

In room 3, is a 45-year-old[1] female[2] symptomatic of chest pain[3] under the right breast[4] accompanied by dyspnea[5] for duration of one day.[6] She states the pain started in the afternoon[7] after she had an argument with her son earlier[8] and describes it as stabbing[9] and radiating to her back.[10] Past medical history is not significant.[11] No recent trauma[12] or fever,[13] she had ankle surgery 2 weeks ago.[14] Her pain has resolved since arrival, so she thinks it was just stress from the argument. So, in her evaluation, her chest x-ray was clear[15] and EKG showed sinus tachycardia[16] and bedside echo did not show any right heart strain[17] and showed a grossly normal EF.[18] Her pregnancy test was negative.[19] We ordered labs[21] for her including a D-dimer[22]but they are not back yet, so you need to follow them up.[23]

LONG & IMPLICIT

In room 3, is a 45-year-old[1] female[2] symptomatic of chest pain[3] under the right breast[4] accompanied by dyspnea[5] for duration of one day.[6] She states the pain started in the afternoon[7] after she had an argument with her son earlier[8] and describes it as stabbing[9] and radiating to her back.[10] Past medical history is not significant.[11] No recent trauma[12] or fever,[13] she had ankle surgery 2 weeks ago.[14] Her pain has resolved since arrival, so she thinks it was just stress from the argument. So, in her evaluation, her chest x-ray was clear[15] and EKG showed sinus tachycardia[16] and bedside echo was unremarkable[17] with a grossly normal EF.[18] We ordered labs[19] for her including a D-dimer.[20] Her CBC,[21] BMP[22] and coags[23] were all negative and her urinary pregnancy test was negative[24]but her D-dimer was positive.[25]

LONG EXPLICIT

In room 3, is a 45-year-old[1] female[2] symptomatic of chest pain[3] under the right breast[4] accompanied by dyspnea[5] for duration of one day.[6] She states the pain started in the afternoon[7] after she had an argument with her son earlier[8] and describes it as stabbing[9] and radiating to her back.[10] Past medical history is not significant.[11] No recent trauma[12] or fever[13], she had ankle surgery 2 weeks ago.[14] Her pain has resolved since arrival, so she thinks it was just stress from the argument. So, in her evaluation, her chest x-ray was clear[15] and EKG showed sinus tachycardia[16] and bedside echo did not show any right heart strain[17] and showed a grossly normal EF.[18] We ordered labs[19] for her including a D-dimer.[20] Her CBC,[21] BMP[22] and coags[23] were all negative and her urinary pregnancy test was negative[24]but her D-dimer was positive.[25] The plan is to do a chest CT[25] and then pending result treat accordingly[26] and she may need admission to medicine depending on the results.[27]

Stimuli Abdomen Case - Short Sign Out

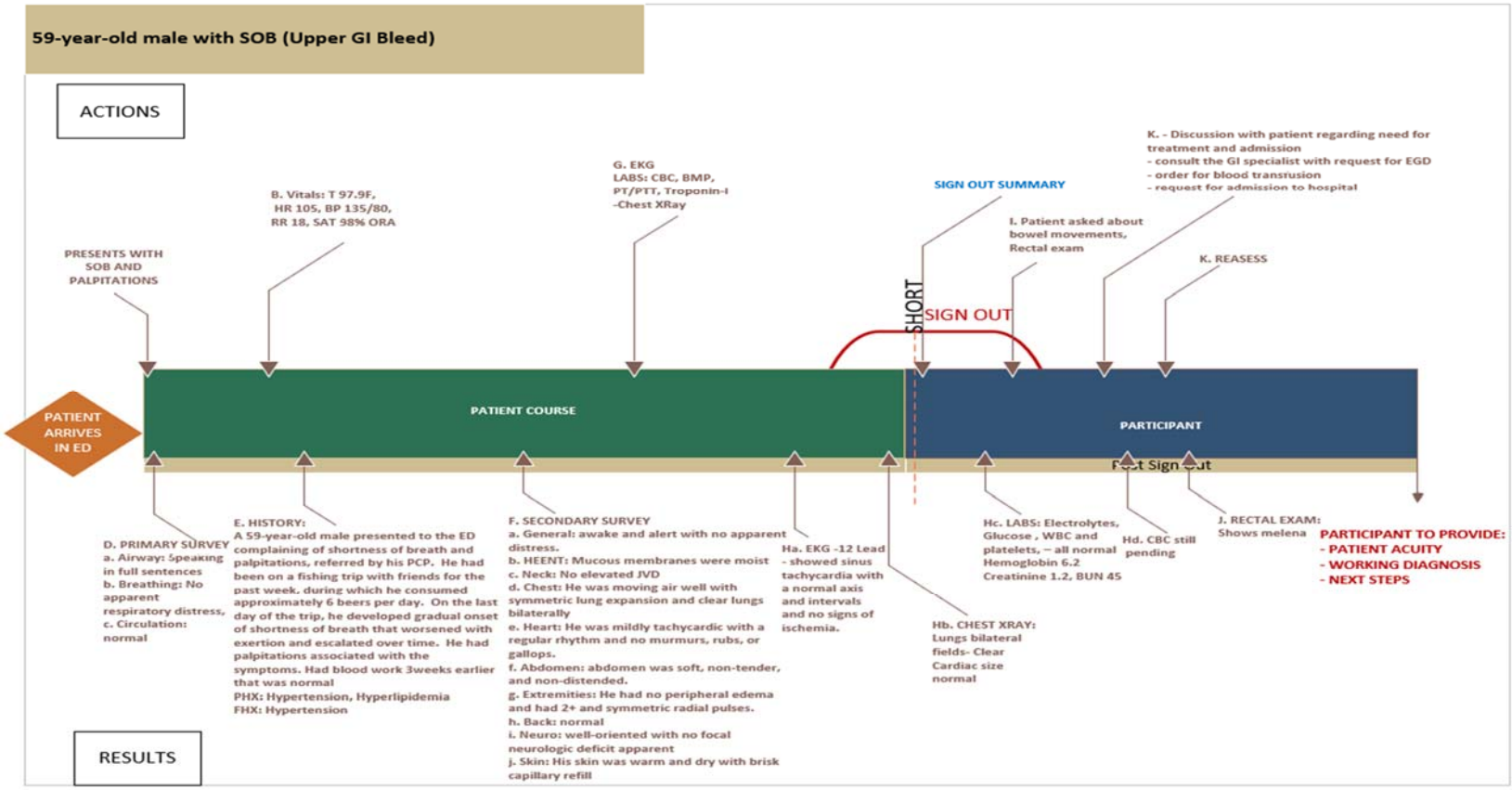


Figure 27: Stimuli Abdomen Case - Short Sign Out

Stimuli Abdomen Case - Long Sign Out

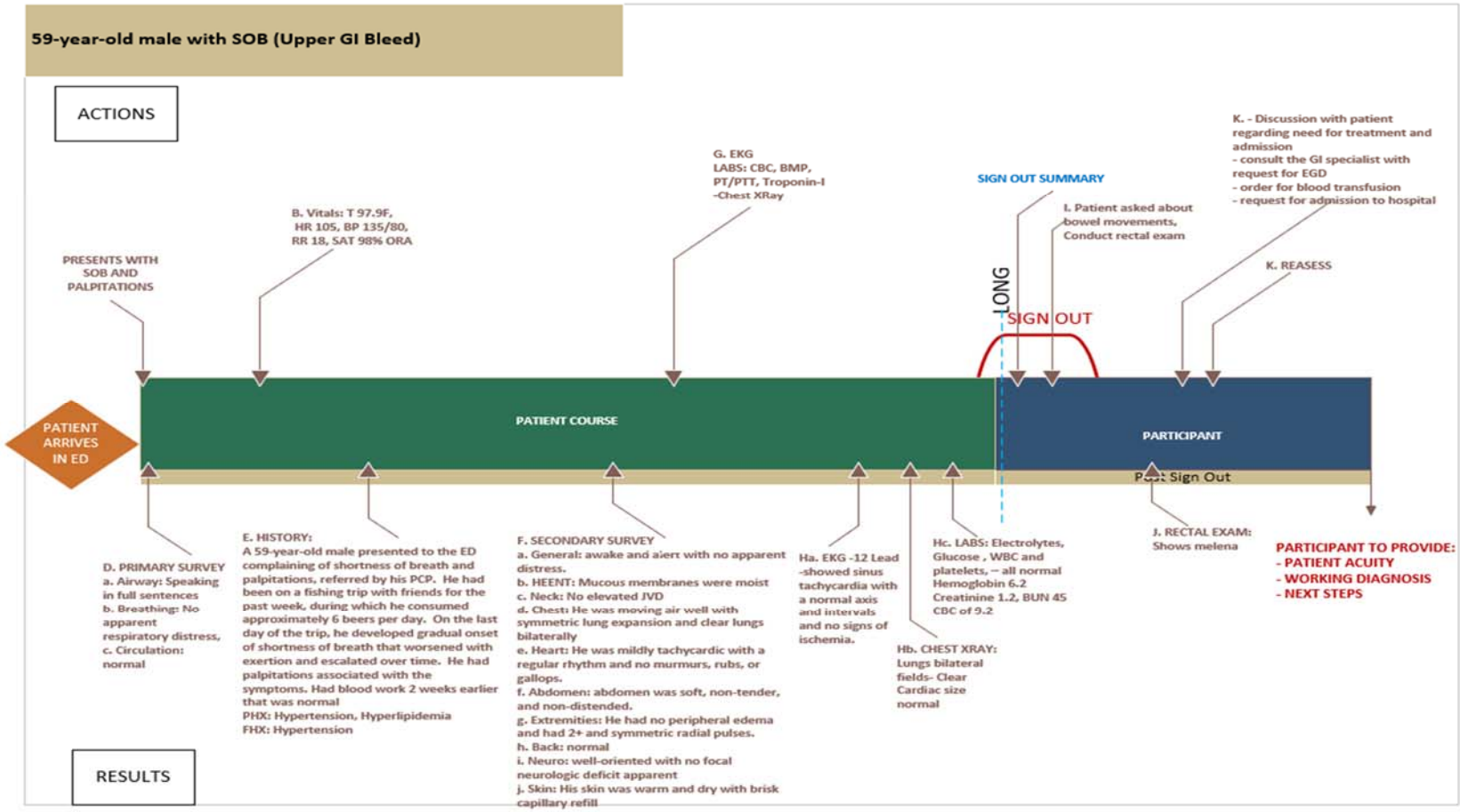


Figure 28: Stimuli Abdomen Case – Long Sign Out

SIGN OUT NARRATIVES FOR STIMULI ABDOMEN CASES

SHORT & IMPLICIT

In room 5 is a 59-year old[1] man[2] who came in with shortness of breath[3] and was sent in by his PCP[4] for evaluation of his new onset of SOB[5], which his PCP suspects may be due to holiday heart syndrome[6]. He has never had this type of shortness of breath before[7] and developed it after going on a fishing trip with some buddies[8] during which he consumed a six pack of beer per day[9]. He described it as gradual in onset[10] associated with dyspnea on exertion.[11] Has no edema[12] and clear lungs.[13] He has occasional palpitations with it.[14] He had routine blood work done around 2 weeks ago[15] and everything was normal[16] with a hemoglobin of 14.[17] So in his evaluation, his EKG just shows sinus tach,[18] doesn't show any evidence of dysrhythmia[19] and his chest x-ray was normal.[20]

SHORT & EXPLICIT

In room 5 is a 59-year old[1] man[2] who came in with shortness of breath[3] and was sent in by his PCP[4] for evaluation of his new onset of SOB[5], which his PCP suspects may be due to holiday heart syndrome. [6] He has never had this type of shortness of breath before[7] and developed it after going on a fishing trip with some buddies[8], during which he consumed a six pack of beer per day.[9] He described it as gradual in onset[10] associated with dyspnea on exertion.[11] Has no edema[12] and clear lungs.[13] He has occasional palpitations with it.[14] He had routine blood work done around 2 weeks ago[15] and everything was normal[16] with a hemoglobin of 14.[17] So in his evaluation, his EKG just shows sinus tach,[18] doesn't show any evidence of dysrhythmia[19] and his chest x-ray was normal.[20] His labs aren't back yet,[21] so you need to follow those up.[22]

LONG & IMPLICIT

In room 5 is a 59-year old[1] man[2] who came in with shortness of breath[3] and was sent in by his PCP[4] for evaluation of his new onset of SOB,[5] which his PCP suspects may be due to holiday heart syndrome.[6] He has never had this type of shortness of breath before[7] and developed it, after going on a fishing trip with some buddies[8], during which he consumed a six pack of beer per day.[9] He described it as gradual in onset[10] associated with dyspnea on exertion.[11] Has no edema[12] and clear lungs.[13] He has occasional palpitations with it.[14] He had routine blood work done around 2 weeks ago[15] and everything was normal[16] with a hemoglobin of 14.[17] So in his evaluation, his EKG just shows sinus tach,[18] doesn't show any evidence of dysrhythmia[19] and his chest x-ray was normal.[20] His troponin was negative[21] times two,[22] normal white count[23] and a creatinine of 1.2[24] and BUN 45[25] but the CBC showed a hemoglobin of 9.2[26]

LONG & EXPLICIT

In room 5 is a 59-year old[1] man[2] who came in with shortness of breath[3] and was sent in by his PCP[4] for evaluation of his new onset of SOB,[5] which his PCP suspects may be due to holiday heart syndrome.[6] He has never had this type of shortness of breath before[7] and developed it, after going on a fishing trip with some buddies,[8] during which he consumed a six pack of beer per day.[9] He described it as gradual in onset[10] associated with dyspnea on exertion.[11] Has no edema[12] and clear lungs.[13] He has occasional palpitations with it.[14] He had routine blood work done around 2 weeks ago[15] and everything was normal[16] with a hemoglobin of 14.[17] So in his evaluation, his EKG just shows sinus tach,[18] doesn't show any evidence of dysrhythmia[19] and his chest x-ray was normal.[20] His troponin were negative[21] times two,[22] normal white count[23] and a creatinine of 1.2[24] and BUN 45[25] but the CBC showed a hemoglobin of 9.2,[26] which has dropped some.[27] Might need to check back with patient about any recent changes to his habits,[28] illnesses[29] or bleeds.[29]

Stimuli Cardiac Case – Short Sign Out

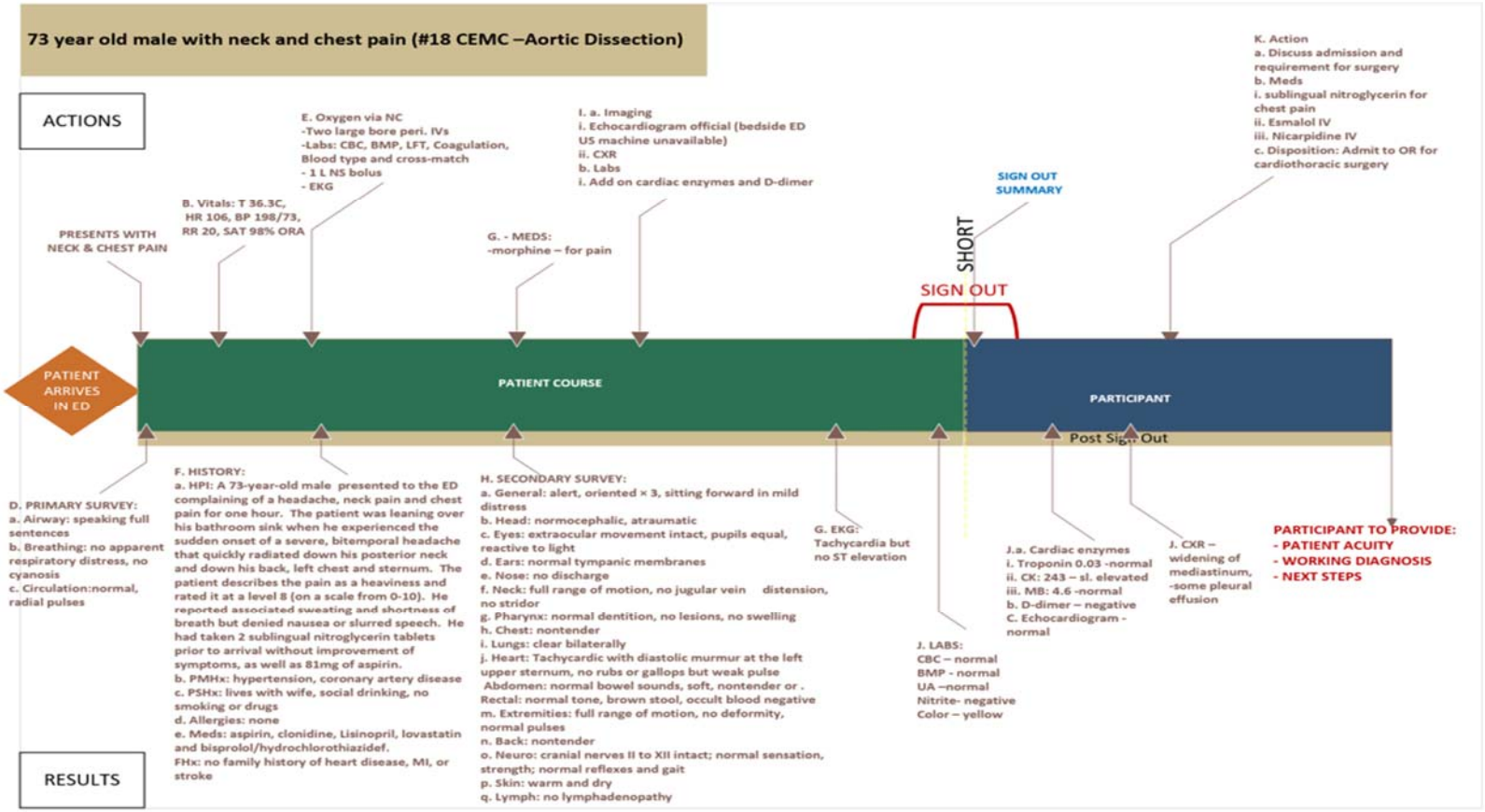


Figure 29: Stimuli Cardiac Case – Short Sign Out

Stimuli Cardiac Case – Long Sign Out

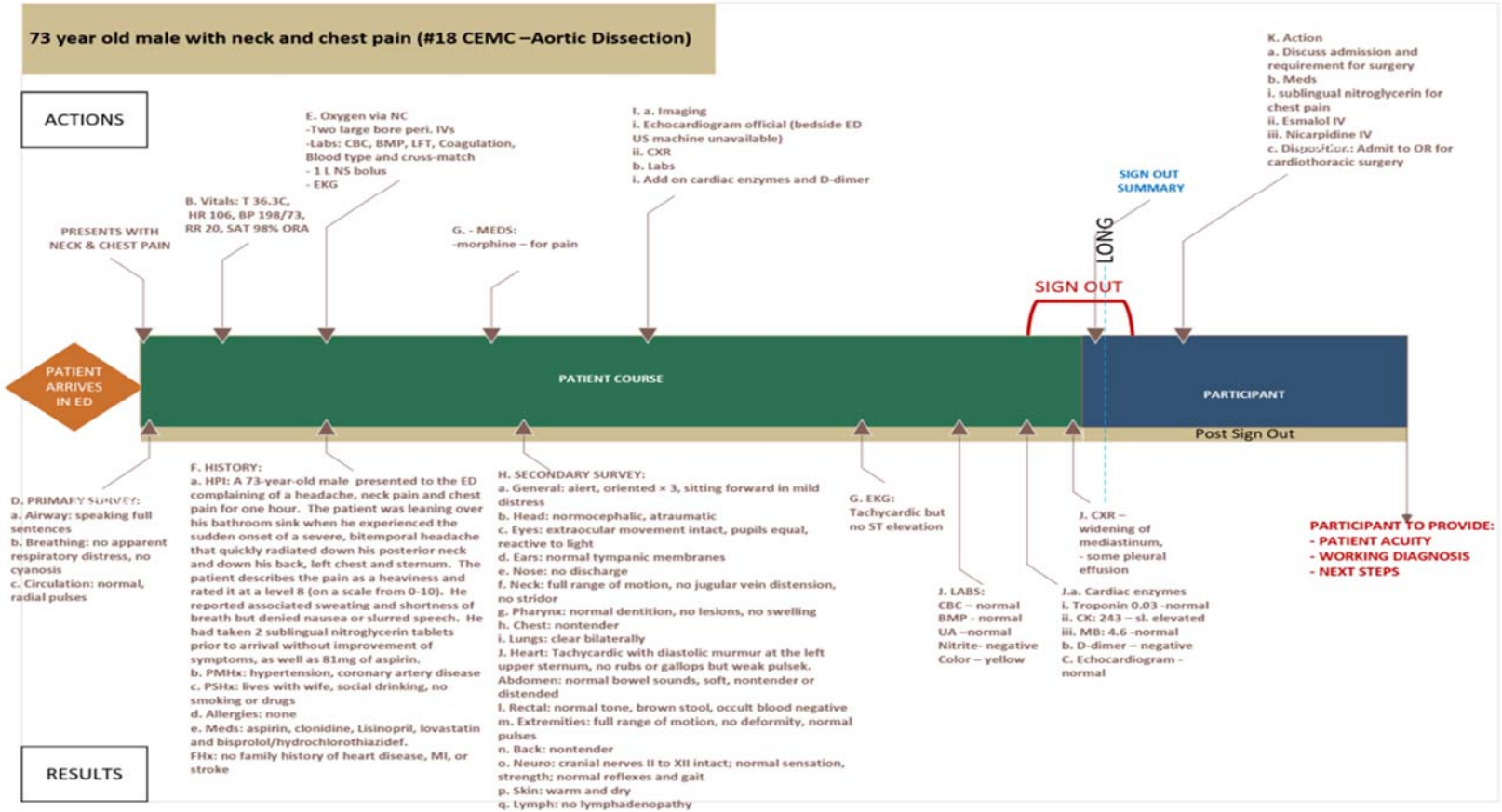


Figure 30: Stimuli Cardiac Case – Long Sign Out

SIGN OUT NARRATIVES FOR STIMULI CARDIAC CASES

SHORT & IMPLICIT

In room 9 is a 73-year old[1] male[2] who came in complaining of a sudden onset[3] headache[4], neck[5] and chest pain[6] that started about an hour ago[7]. He stated pain is a level 8 [8] and is radiating[9] down his posterior neck[10], down his back[11] on the left sternum[12]. On exam he was sweating[13] with shortness of breath[14] but without nausea[15] and slurred speech[16]. He has a history of hypertension[17] and CAD[18] which is controlled with medication[19] including aspirin[20]. During his examination he was alert[21] and oriented[22] with no focal neuro deficits[23] but there was a diastolic murmur[24] at the left upper sternum[25] and his pulse was weak[26]. His EKG shows tachycardia[27] and no ST elevation[28]. We ordered labs including cardiac enzymes[29]. His CBC[30], BMP[31], LFT[32], UA[33] and coags[34] were all normal. He needs to be re-assessed for cardiac risk stratification.[35]

SHORT & EXPLICIT

In room 9 is a 73-year old[1] male[2] who came in complaining of a sudden onset[3] headache[4], neck[5] and chest pain[6] that started about an hour ago[7]. He stated pain is a level 8 [8] and is radiating[9] down his posterior neck[10], down his back[11] on the left sternum[12]. On exam he was sweating[13] with shortness of breath[14] but without nausea[15] and slurred speech[16]. He has a history of hypertension[17] and CAD[18] which is controlled with medication[19] including aspirin[20]. During his examination he was alert[21] and oriented[22] with no focal neuro deficits[23] but there was a diastolic murmur[24] at the left upper sternum[25] and his pulse was weak[26]. His EKG shows tachycardia[27] and no ST elevation[28]. We ordered labs including cardiac enzymes[29]. His CBC[30], BMP[31], LFT[32], UA[33] and coags[34] were all normal. We are waiting on his cardiac enzymes[35] and the chest x-ray[36] and then he needs to re-assessed for cardiac risk stratification.[37]

LONG & IMPLICIT

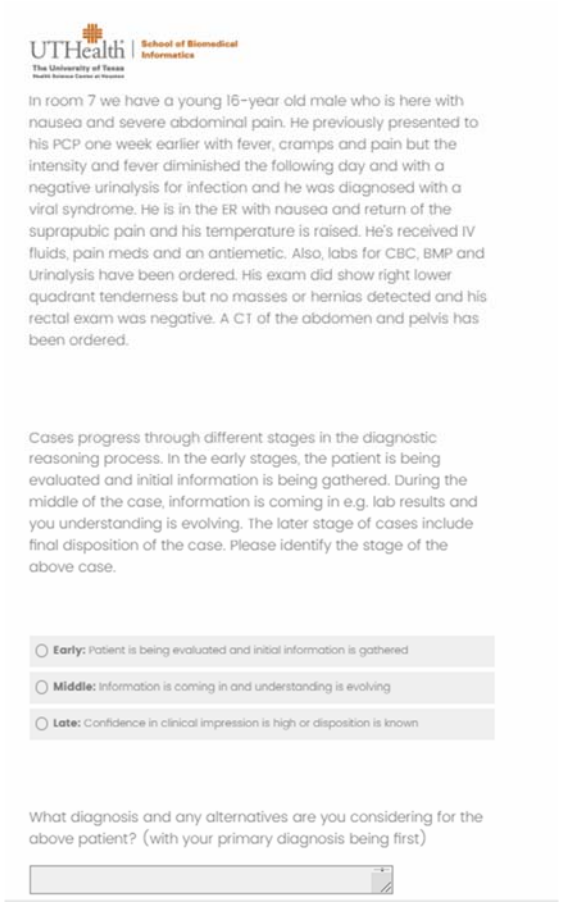
In room 9 is a 73-year old[1] male[2] who came in complaining of a sudden onset[3] headache[4], neck[5] and chest pain[6] that started about an hour ago[7]. He stated pain is a level 8[8] and is radiating[9] down his posterior neck[10], down his back[11] on the left sternum[12]. On exam he was sweating[13] with shortness of breath[14] but without nausea[15] and slurred speech[16]. He has a history of hypertension[17] and CAD[18], which is controlled with medication[19] including aspirin[20]. During his examination he was alert[21] and oriented[22] with no focal neuro deficits[23] but there was a diastolic murmur[24] at the left upper sternum[25] and his pulse was weak[26]. His EKG shows tachycardia[27] and no ST elevation[28]. We ordered labs including cardiac enzymes[29]. His CBC[30], BMP[31], LFT[32], UA[33] and coags[34] were all normal. His cardiac enzymes were normal[35] and his D-dimer was negative[36]. The chest x-ray shows widening mediastinum[37] and there are small pleural effusions[38] so we need confirmatory testing.[39]

LONG & EXPLICIT

In room 9 is a 73-year old[1] male[2] who came in complaining of a sudden onset[3] headache[4], neck[5] and chest pain[6] that started about an hour ago[7]. He s

tated pain is a level 8[8] and is radiating[9] down his posterior neck[10], down his back[11] on the left sternum[12]. On exam he was sweating[13] with shortness of breath[14] but without nausea[15] and slurred speech[16]. He has a history of hypertension[17] and CAD[18], which is controlled with medication[19] including aspirin[20]. During his examination he was alert[21] and oriented[22] with no focal neuro deficits[23] but there was a diastolic murmur[24] at the left upper sternum[25] and his pulse was weak[26]. His EKG shows tachycardia[27] and no ST elevation[28]. We ordered labs including cardiac enzymes[29]. His CBC[30], BMP[31], LFT[32], UA[33] and coags[34] were all normal. His cardiac enzymes were normal[35] and his D-dimer was negative[36]. The chest x-ray shows widening mediastinum[37] and there are small pleural effusions[38] so there is concern for AD[39]. So, we going to order a CTA[40] and start them on IV esmolol.[41]

Appendix E: Qualtrics EDSO Survey Screens



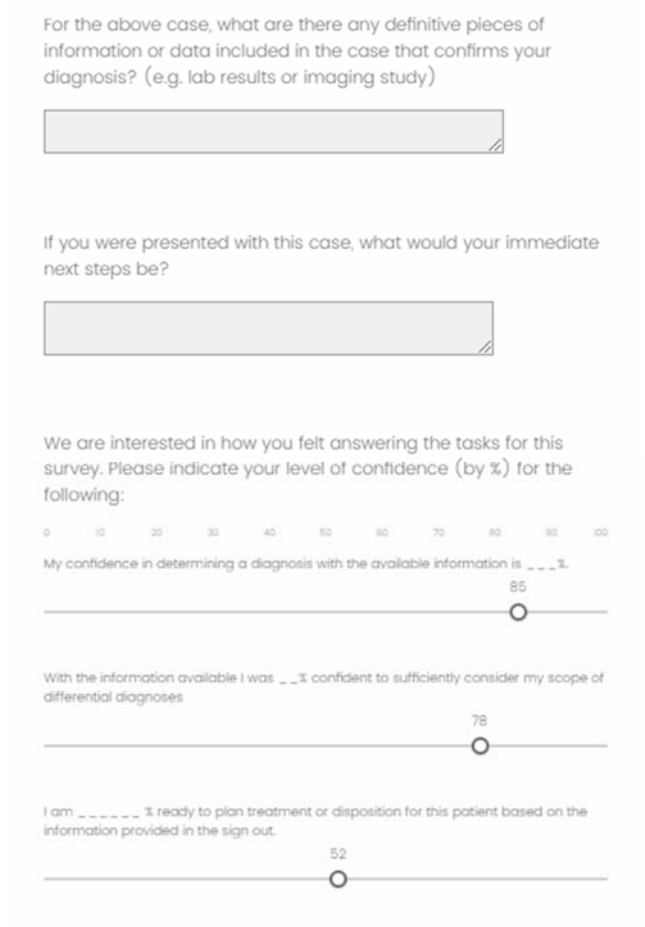
UTHealth | School of Biomedical Informatics
The University of Texas Health Science Center at Houston

In room 7 we have a young 16-year old male who is here with nausea and severe abdominal pain. He previously presented to his PCP one week earlier with fever, cramps and pain but the intensity and fever diminished the following day and with a negative urinalysis for infection and he was diagnosed with a viral syndrome. He is in the ER with nausea and return of the suprapubic pain and his temperature is raised. He's received IV fluids, pain meds and an antiemetic. Also, labs for CBC, BMP and Urinalysis have been ordered. His exam did show right lower quadrant tenderness but no masses or hernias detected and his rectal exam was negative. A CT of the abdomen and pelvis has been ordered.

Cases progress through different stages in the diagnostic reasoning process. In the early stages, the patient is being evaluated and initial information is being gathered. During the middle of the case, information is coming in e.g. lab results and you understanding is evolving. The later stage of cases include final disposition of the case. Please identify the stage of the above case.

- ○ **Early:** Patient is being evaluated and initial information is gathered
- ○ **Middle:** Information is coming in and understanding is evolving
- ○ **Late:** Confidence in clinical impression is high or disposition is known

What diagnosis and any alternatives are you considering for the above patient? (with your primary diagnosis being first)



For the above case, what are there any definitive pieces of information or data included in the case that confirms your diagnosis? (e.g. lab results or imaging study)

If you were presented with this case, what would your immediate next steps be?

We are interested in how you felt answering the tasks for this survey. Please indicate your level of confidence (by %) for the following:

0 10 20 30 40 50 60 70 80 90 100

My confidence in determining a diagnosis with the available information is ___%.

85

With the information available I was __% confident to sufficiently consider my scope of differential diagnoses

78

I am _____ % ready to plan treatment or disposition for this patient based on the information provided in the sign out.

52

Appendix F: Survey Build for Randomized Assignment

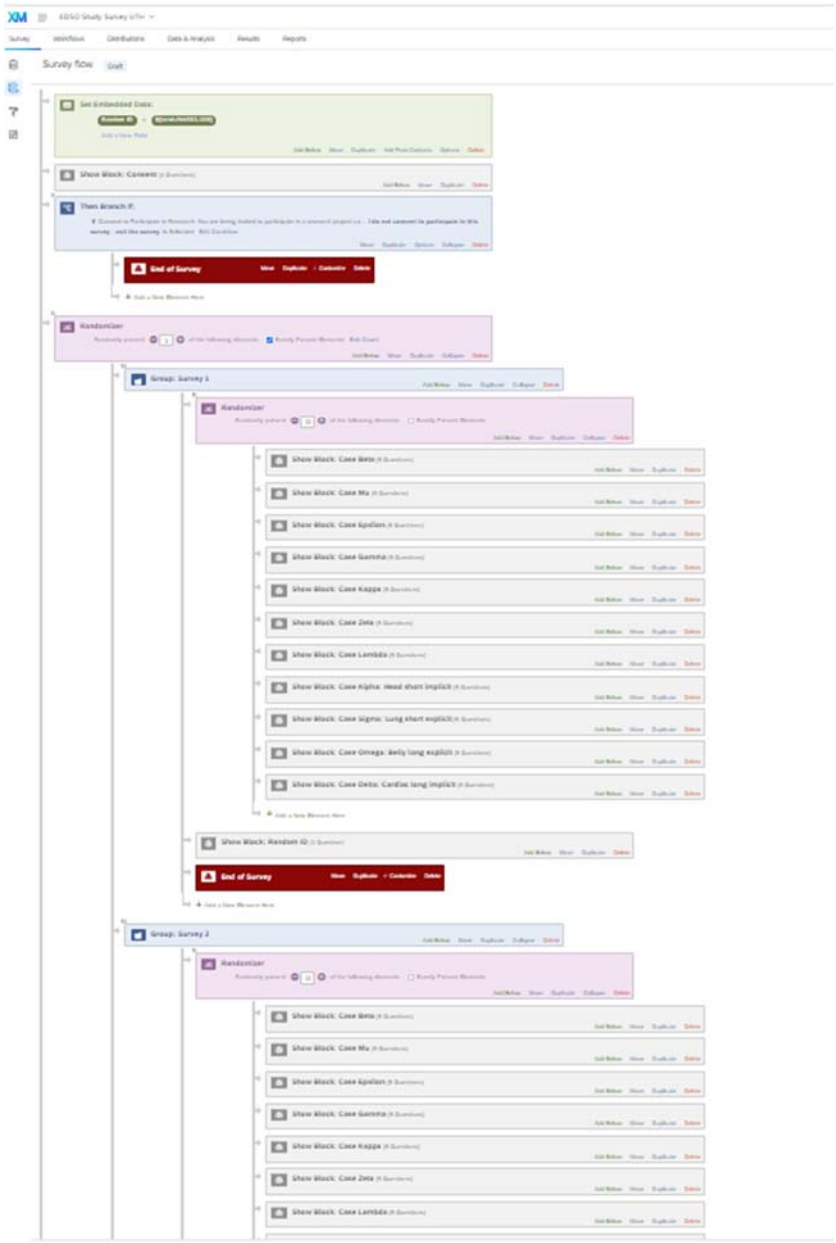


Figure 32: Survey Build for Randomized Assignment - within Qualtrics Survey Flow