

Causal Analysis based on Systems Theory (CAST) of an Adverse Event involving Laboratory Data

September 6, 2024

Authors:

Polly Harrington

Dr. John Thomas

Prof. Nancy Leveson

Massachusetts Institute of Technology

Partnership for Systems Approaches to Safety and Security

<https://mit.edu/psas>

and

Dr. Stephen Powell

Alana Keller, PMP

Synensys, LLC

Contact: Dr. Stephen Powell

spowell@synensysglobal.com

DISCLAIMER

The opinions expressed in this presentation are those of the research team members. They do not purport to reflect the opinions or views of the U.S. Food and Drug Administration (FDA), its affiliates, or the organizations included in the research. This report was completed under FDA Contract #75F40122C0012.

Table of Contents

Table of Contents	1
Abstract.....	3
1. Research Goals	4
2. Adverse Event Summary	5
3. Initial Investigation and Response	5
4. Research Method: CAST.....	7
5. CAST Results.....	8
System Boundary.....	8
Relevant Hazards and Constraints.....	8
Proximal Events	8
Control Structure.....	12
Component Analysis.....	12
Physical Contributions	12
Presurgical Testing Advanced Practice Provider Description.....	14
Oncologist	20
Surgery Department	22
Urgent Care Center	24
Laboratory	26
Chemistry and Endocrine Services	30
Department Leadership.....	32
Hospital Administration.....	36
Regulatory Authorities	39
Accreditation Agencies.....	44
Systemic Factors and Additional Recommendations.....	46
Communication and Coordination.....	47
Safety Management System	49
Culture	51
Changes and Dynamics.....	51
Economic Pressure	52
Regulatory Ecosystem	53
Additional Recommendations	53
6. Conclusions.....	54

7. Reference List55

Acronyms57

Appendix A: Laboratory Testing Background Information58

 Cortisol58

 Immune Checkpoint Inhibitors (ICI)58

 Display of Laboratory Data58

 Laboratory Terms61

 Reference Range61

 Flagging range62

 Critical Value Range.....64

 Regulatory Requirements.....65

 Standards and Guidance69

Abstract

This report documents the results of a CAST (Causal Analysis based on Systems Theory) analysis of an adverse event involving laboratory diagnostic testing in a clinical setting. CAST uses systems theory to understand how the interactions within the broader system led to unanticipated and undesirable outcomes [1]. The goal of CAST is to understand and address the systemic causes of an incident in order to avoid other adverse events with similar causes from happening in the future. CAST avoids blaming individuals and instead tries to identify why actions that were ultimately unsafe appeared reasonable at the time.

The analysis resulted in the identification of several systemic factors that contributed to the adverse event, ranging from the communication and coordination between departments and individuals to economic pressures and insufficient regulatory controls regarding laboratory testing. CAST provided dozens of recommendations in addition to the changes that were made after the initial incident investigation.

1. Research Goals

Preventable medical errors are a significant source of injury and death in the United States. Since the publication of the 2000 National Institutes of Health report, *To Err is Human*, research and resources have been dedicated to identifying and reducing the impact of preventable medical errors on patients. However, there is limited evidence of substantial progress. One source of medical errors is in diagnostics. Estimates suggest that up to 70% of adverse events in hospitals are a result of diagnostic errors [2].

This report builds on a systems-theory-based safety review of the United States Diagnostic Laboratory Data System [3], [4]. While the previous reports focused on potential safety hazards in the system based on current design, this research focuses on the causes of one adverse event to identify some of the systemic factors involving laboratory data. The goal is to demonstrate a state-of-the-art root cause analysis technique applied to real adverse events.

While this particular adverse event took place in one hospital, the underlying causal factors and system vulnerabilities are present across the United States healthcare infrastructure.

This report is based on interviews with 22 key personnel, some from within the medical center as well as outside expert contributors. Interviewees included physicians, leadership, advanced practice providers, laboratory leaders, Information Technology (IT) experts, regulators, and others.

2. Adverse Event Summary

The healthcare facility where the adverse event occurred is a highly respected academic medical center in the United States with ancillary clinics. Like all medical centers, adverse events can and do occasionally happen, and it is important to learn from them so that changes can be made to prevent future losses.

The adverse event involved a patient who was undergoing treatment at the medical center for breast cancer. After initial chemotherapy and immunotherapy, the patient was scheduled for surgery and was seen for Presurgical Testing (PST).

During the PST visit, the patient had a battery of tests done, including a serum cortisol level test. The cortisol level was low, indicating adrenal insufficiency. There was no phone call from the lab to the ordering provider to notify of a critically low value, but the results were available via the Electronic Health Record (EHR). The PST provider did not identify the low cortisol level.

The patient was referred by PST to the Urgent Care Center (UCC) for evaluation of a possible pulmonary embolism. The UCC staff were notified of the PST visit and received the lab results. The patient's low cortisol level, indicating adrenal insufficiency, was not identified by UCC staff and the patient was not diagnosed. UCC referred the patient for an echocardiogram the next day, which did not lead to a diagnosis.

The oncologist and surgeon were both notified of the patient's lab results, symptoms, and referral from PST to UCC. However, they did not identify the low cortisol level indicating adrenal insufficiency.

Ten days after the initial cortisol lab test, the patient called an after-hours nursing service to report their worsening symptoms. The patient was told to follow up with their local (external to the hospital system) cardiologist the next day. This local cardiologist identified the low cortisol value from the presurgical testing ten days earlier, made a note of it, and referred the patient to a local (non-affiliated) Emergency Department (ED) for treatment.

The next day the patient called their oncologist to report being in the ED with a diagnosis of adrenal insufficiency. The following day, 12 days after the original PST visit and cortisol test, the medical oncologist learned of the diagnosis and treatment of adrenal insufficiency and spoke to the patient.

The patient recovered and successfully underwent the planned surgery about two weeks later.

3. Initial Investigation and Response

The adverse event was originally investigated by other groups using standard approaches (not CAST). The investigation produced the following findings:

1. The PST provider reviewed the lab results quickly and missed the low cortisol level.
2. Cortisol has a diurnal reference range. For a.m. samples, the range is 5-25, and for p.m. samples, the range is 3-12. As a result, the EHR had not been set up to flag cortisol values that are abnormal (outside the reference range).

3. Cortisol does not meet the definition of a critical value, which is why the laboratory did not notify the provider of the extremely low cortisol level.
4. The event primarily involved an oversight by the PST Advanced Practice Provider (APP). No Medical Doctors (MDs) were involved in the event in any way that is substantial.

As a result of the findings, the following corrective actions were taken:

1. PST providers were reminded to carefully review lab results like cortisol and better document reasons for the test.
2. The EHR and Laboratory Information System (LIS) were changed to flag some abnormal cortisol values. The cortisol test (which may be performed at any time of day) was updated to flag values that exceed the a.m. cortisol reference range (5-25). A text note was added to explain that clinicians should check the reference range and interpret the result differently if it is not an a.m. test.
3. The critical value policy was reviewed. Changes were not found to be justified as cortisol does not meet the definition of a critical value that requires a phone call from the lab.
4. The Department of Medicine Quality Assurance group declined to investigate the role of MDs as none appeared to have been involved in the event.

Several months later, the event was selected as part of this CAST research project to explore systemic factors behind adverse events that involve diagnostic laboratory data. CAST identified that the five clinicians (including Nurse Practitioners [NPs] and MDs) were unable to diagnose the patient because of systemic factors, including the laboratory data that is transmitted, the way it is displayed, and regulatory requirements. The CAST analysis found that this is not simply a case of random, isolated, human error, but a set of systemic factors that enabled this adverse event to occur.

4. Research Method: CAST

CAST is an accident investigation methodology that goes beyond individual contributions and identifies the systemic, organizational, and regulatory contributions to an accident, adverse event, or incident so they can be addressed. CAST is based on systems theory, which recognizes that individual behaviors cannot be understood in isolation. For more information about systems theory, refer to the Systems Theory section of the *Laboratory Data Exchanges Report* [3].

The goal of the CAST investigation is not to identify who should be blamed for an adverse event. Instead, the goal is to identify why the actions that led to unsafe outcomes made sense to the individuals at the time. If the contextual factors that made the actions seem reasonable are not addressed, similar unsafe actions may recur in the future. Blame hinders the ability to prevent similar future adverse events that may have similar systemic causes.

To understand the complex causes behind adverse events, CAST models the interactions between various system components, such as technology, medical practitioners, management, organizational components, and regulators. The role of each component in the adverse event is analyzed to understand the reasons for their actions. The reasons can include potential beliefs (mental models) that made the actions appear reasonable, potential information (feedback) that was available or not available at the time, and other contextual factors. Feedback includes relevant information, such as laboratory data, the EHR screens and alerts, phone or other verbal communication, reports, alarms, or other information.

CAST (and systems theory) assumes that adverse events are caused not by individual errors but by a lack of control over the behavior of the system components, that is, inadequate control over the unsafe interactions among the system components. By analyzing the interactions, CAST identifies what unsafe behaviors are not effectively controlled by the overall system design. Examples of ineffective control can include communication deficiencies, like inadequate information to make safe decisions, gaps in responsibilities among the components, lack of authority to enforce certain constraints, and other factors. Understanding these deeper systemic factors enables the identification of more effective recommendations and opportunities for improvement.

More information about CAST can be found in the *CAST Handbook* [1] and the *CAST Handbook for Healthcare* [5]. CAST is based on the same principles as STPA (System-Theoretic Process Analysis), which is a forward-looking analysis to proactively anticipate vulnerabilities, weaknesses, and future accidents before they happen.

5. CAST Results

System Boundary

The analysis focuses on the healthcare facility, including the hospital, its administration, employees, patients, and information systems, like the EHR and LIS. The behaviors of these components are, in turn, partly shaped by the regulatory context, including coding and standards organizations, federal and state regulatory bodies, and communications with peer institutions. This analysis aims to identify the relevant interactions between these components that facilitated the adverse event and to identify opportunities to improve system safety moving forward.

Relevant Hazards and Constraints

CAST analyzes the causes of a loss, which is any outcome that is unacceptable to the system stakeholders, such as death or injury to patients or personnel, equipment or property damage, environmental pollution, mission loss, negative business impact (damage to reputation, etc.), loss or disclosure of protected information, financial losses, etc. In the selected adverse event, the loss relates to the patient's quality of life as diagnosis and treatment were delayed.

Loss: Deterioration of the patient's health and quality of life

Next, the hazard is identified. Hazards are states or conditions of the system that, along with a particular set of worst-case environmental conditions, will lead to a loss or adverse event. Two examples of hazards in healthcare are 1) patients receiving less than acceptable standard of care and 2) HIPAA (Health Insurance Portability and Accountability Act) privacy rule violations.

Hazard 1: Patient receives less than acceptable standard of care

Steps should be taken to prevent hazards from occurring and to deal with them if, despite efforts to eliminate them, they do occur. Once the hazards are identified, constraints are identified at a high level to define the overall safety-related responsibilities of the healthcare system relative to the hazards.

Constraints:

1. Patients must receive the accepted standard of care
2. It must be possible to identify when standard of care has not been administered
3. Patients with information identifying life-threatening health conditions must be identified
4. Those who make care decisions must have the necessary information to make those decisions

Proximal Events

Table 1 provides a timeline of the proximal events that led to the adverse event. The purpose of the timeline is to provide a starting point for the analysis, not a conclusion. At the start of the analysis, some basic events may be known but not the reasons or the context behind the events. More distant events in the past that contributed to the current event may not be known yet. The purpose of the timeline is to establish some of the basic facts of the event and generate questions that will be analyzed in later steps to uncover deeper systemic factors.

Table 1: Proximal Event Timeline

ID	Event	Questions
1	Patient is diagnosed with triple negative invasive ductal carcinoma on the right breast.	
2	Patient treatment plan is determined, including chemotherapy and a partial mastectomy.	
3	Day 1 Patient completes neoadjuvant chemotherapy / pembrolizumab).	
4	Day 6, 1:30 PM Patient has presurgery anesthesia evaluation appointment with a PST Advanced Practice Provider (APP). Labs were drawn, including: <ul style="list-style-type: none"> • Complete Blood Count • Comprehensive Metabolic Panel • Hemoglobin A1C • Thyroid Function Tests • Cortisol Patient's heart rate was measured at 122 beats per minute. Patient described symptoms, including shortness of breath. Patient was sent to the UCC at the medical center for evaluation in order to rule out a pulmonary embolism ¹ based on elevated heart rate and shortness of breath. Patient's surgeon and oncologist were notified and consulted. Patient was referred to cardiologist to get clearance for surgery.	How was the PST APP trained to evaluate cortisol results? How do patient symptoms get followed up on after PST appointments? Was the patient's case reviewed by a cardiologist at the medical center or an outside cardiologist?
5	Day 6, 4:14 PM Lab results are returned to PST APP. All lab values were benign, except cortisol which was 0.5 mcg/dL.	Who were the lab results reviewed by?
6	Day 6, 7:05PM Patient arrives in UCC. No issues are found in a respiratory panel A CT scan of patient's chest revealed no central pulmonary embolus.	How many patients are referred to UCC after a PST visit? Did the UCC review the patient's labs?

¹ Pulmonary embolisms refer to blood clots that prevent blood from reaching a lung artery [6]

	A subcentimeter perifissural ² nodule in the upper right lobe was identified. The patient was sent home but was scheduled for an outpatient echocardiogram the next day.	
7	Day 7 Outpatient echocardiogram revealed a normal result with an ejection fraction of 59%. Patient was given medication to address elevated heart rate. ³	Did the outpatient clinic have access to the patient lab data?
8	Day 16 Patient visited their external cardiologist. Cardiologist viewed the patient's labs from the patient's PST visit on Day 6 and identifies a low cortisol value.	Was this cardiologist the same or different from the cardiologist who gave the echocardiogram?
9	Day 16 * Patient calls the medical center's after-hours number to report low blood pressure, lightheadedness, and fatigue. Patient notes that the cardiologist is concerned with adrenal insufficiency.	What actions did the after-hours clinician take? Did they review the patient's lab results? Did they refer the patient to an ED? Events 8, 9, and 10 were reported in different orders by different sources. What was the process for gathering information after the missed diagnosis was caught? When was the appointment with the local cardiologist scheduled? Was the appointment scheduled based on emerging symptoms or after the PST appointment? Did the cardiologist refer the patient to the ED for treatment, or was the patient referred by the after-hours clinician?
10	Day 17 * Patient calls the medical oncologist at the medical center to notify about being in external ED for possible adrenal insufficiency.	
11	Day 18 Medical oncologist spoke to patient again and determined patient was being treated with cortisone replacements but was still experiencing issues with tachycardia on	

² Lung Nodules are abnormal areas of lung tissue found in CT scans [7] and can indicate a variety of conditions including scarring and cancerous growths. Perifissural Nodules are a subcategory and usually refer to nodules that are “a noncalcified solid nodule with sharp margins and a regular shape abutting or near the pleural or fissural margin.” [8] These nodules are considered benign in most cases [8].

³ Ejection Fractions are the percentage of blood that leaves the heart during a contraction. The regular range is between 50% and 70% [9].

	ambulation ⁴ and would stay in the external care facility for another day. Oncologist planned to get an endocrinology appointment for the patient before determining surgery date.	
12	Day 32 The patient was cleared for surgery and underwent planned right partial mastectomy.	

⁴ Postural Tachycardia Syndrome (PoTS) describes when a patient's heart rate elevates rapidly when moving from a sitting or lying down position to standing [10].

Control Structure

A control structure models the interactions between decision-makers and the controlled processes they manage. For example, Figure 1 shows that clinicians can receive lab results from a Clinical Information System (CIS), such as an EHR, and they can take actions, like providing referrals and prescribing treatment. The clinicians are decision-makers who collect feedback (like lab results), form mental models (like a belief that a patient may have adrenal insufficiency or not), and provide control actions (like issuing a referral). This relationship (feedback, decision, control action) is called a *control loop*.

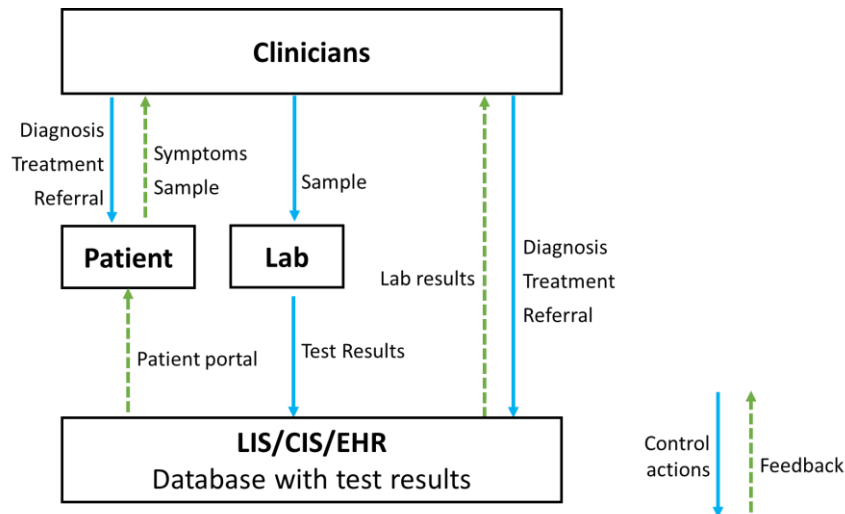


Figure 1: Initial Control Structure

Gaps and disruptions at any point along a control loop can lead to adverse events. For example, if lab results provide missing or misleading information, then they can contribute to an incorrect clinical diagnosis. CAST analyzes these control loops to understand the context behind the decisions and the actions that lead to an adverse event.

Several control loops are modeled in Figure 1, and more are added as the analysis uncovers additional interactions that played a role in the adverse event. At the end of the analysis, the final control structure will represent the overall Safety Management System (SMS) with respect to the adverse event. Recommendations will be developed not only to improve individual elements of the control structure, but to strengthen the SMS as a whole.

Component Analysis

Physical Contributions

The analysis begins with the bottom layer of the control structure, which in this case includes the data and information systems used by clinicians.

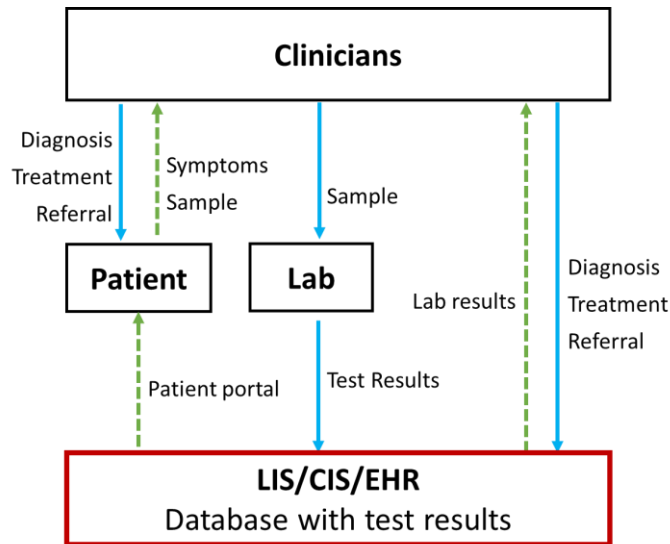


Figure 2: Control Structure with Data and Information Systems

The relevant technical/physical components were the:

- Laboratory Information System (LIS)
- Electronic Health Record (EHR)

Failures

No piece of software failed. Everything behaved exactly as it was designed to, including the EHR.

Unsafe interactions

The flagging range field in the LIS was left blank when the cortisol test was set up. Instead, the notes field was used to describe the diurnal reference range via text.

There were no requirements that flags be applied to all tests. Regulatory authorities require that reference ranges accompany all lab results. However, text notes fulfill this requirement.

Contextual factors

The cortisol test results that were outside of their reference range did not have any flags applied.

- Cortisol did not have critical values set in the LIS/EHR because critical values can be time and context dependent
- The LIS/EHR had limited ability to have flags based on contextual parameters like time of day.
- The LIS and EHR were different systems and therefore developed separately.
- The current medical center EHR is a legacy system. The medical center is moving to a new LIS/EHR system in the next year. The new system will have an integrated LIS and EHR.

Questions raised so far

- Do clinicians know when tests have a missing flagging range (flags disabled)?
- Who is responsible for determining which tests are flagged?

- Who is responsible for identifying missing flags?

Summary of role of technical components

The technical and data components of this system functioned as designed. All pieces of data were transmitted accurately and in a reasonable amount of time. However, the information was not displayed on some screens (in accordance with the intended design), and it was displayed inconsistently or counter-intuitively on other screens. Although the display matched the intended design, it did not effectively communicate the required information to clinicians and contributed to the errors and the delayed diagnosis.

Presurgical Testing Advanced Practice Provider Description

The first clinician involved in the event was a Presurgical Testing Advanced Practice Provider (PST APP). PST APPs are nurse practitioners or physician assistants who meet with patients before surgeries to ensure the patient is prepared for surgery and is sufficiently healthy to undergo the procedure with anesthesia. During the presurgery appointment, the PST APP takes the patient's vitals and orders lab tests for the patient. The lab tests are ordered based on the planned surgery and the patient's history. Figure 3 models the control structure specific to PST APPs.

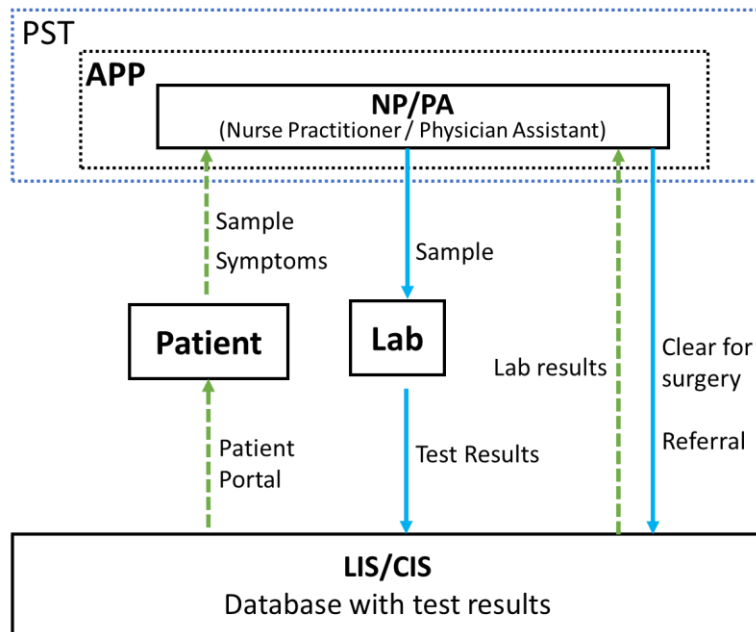


Figure 3: Control Structure Including PST APP Clinicians

Relevant responsibilities

- Order presurgical testing
- Determine patient's fitness for surgery
- Refer patient to other providers for presenting issues

Role in Adverse Event

- PST APP did not recognize the low cortisol lab result and thus did not refer patient to Endocrinology or begin treatment for adrenal insufficiency.

- PST APP did not document the reason for the tests (i.e., that cortisol test was ordered to rule out possible adrenal insufficiency for this patient) or the plan to follow up regarding the tests and interventions.

Mental model

- PST APP believed the cortisol value in the lab results was within the reference range.
- PST APP believed that abnormal cortisol values would be flagged, just like most tests. See Figure 12, Figure 13, and Figure 14 in Appendix A for examples of EHR flagging.
- PST APP believed that lab values displayed in black font with no flag are within the reference range. See Figure 12, Figure 13, and Figure 14 in Appendix A for examples of how normal lab results are displayed.
- PST APP believed the typical low reference range for cortisol was 5 (which is correct for a.m. sample, not p.m.).
- PST APP may have believed the tested cortisol value was 5.0 rather than 0.5.
- PST APP was not fully aware of the contextual factors that affect cortisol interpretation.
- PST APP believed that it was not necessary to reach out to an endocrinologist for consultation to interpret cortisol.
- PST APP may have expected any abnormal test results to be managed by the UCC after referring the patient to the UCC.
- PST APP may have expected the surgeon and oncologist, who were notified of the patient symptoms and test results, would identify any abnormal results that were not obvious to PST APP.

Context

Behavior can only be understood completely by looking at the context in which it occurs. In order to understand why the PST APP took these actions and formed these beliefs, one needs to understand the context in which the work was conducted. A number of contextual factors influenced the PST APP actions, including the EHR, cortisol tests, PST scheduling, and laboratory communication.

EHR Context

The EHR was the final destination of laboratory test data. A number of deficiencies caused the laboratory data to be transmitted and communicated in misleading ways or hidden from view during typical clinical workflows. Laboratory data⁵ relevant to this event includes the reference ranges, flags, flagging range, time of the blood sample, indicators about additional information that is available, and text notes about the test.

- Most test results are displayed in red when abnormal and in black when normal. However, cortisol values are always displayed in the same way as a normal test result (in black).
- The test results screen used by clinicians does not automatically show the reference ranges for each test result.

⁵ See Appendix A for more information about the types of laboratory data relevant to this case.

- The test results screen used by clinicians does not show which test results are missing a flagging range, meaning the flags will not work and the color is meaningless. In other words, there is no easy way to tell when the red flags are working.
- There is no simple way for the APP to distinguish between normal test results (black font) vs. abnormal test results that are missing a flagging range (also black font).
- The various test result screens communicate different times, including the “time of the order”, “time of report”, “time performed by lab”, and “time result received”. The EHR screens do not appear to clearly display the time the blood sample was actually drawn. This information is needed by clinicians to interpret diurnal values, like cortisol.
- There are no defined critical values for cortisol, even extremely low cortisol, such as 0.5 (reference range 3-12). Therefore, there was no call from the lab to notify the PST APP of the low cortisol result.
- The EHR/LIS had the cortisol reference ranges for a.m. and p.m. listed in a note on the screen. The note is not normally displayed on one of the screens used by clinicians unless you rest the cursor over the cortisol result or click to see the notes. Another screen displays the notes on the left side, away from the lab value and away from the field that normally contains the reference range. On each screen, the abnormal cortisol value is displayed in the same way as a typical normal result that is within the reference range (black value with no flags).
- The EHR shows an asterisk (*) to indicate when there is a hidden notes field with additional information. However, there is no legend, and most clinicians do not know what the asterisk means. Clinicians do not regularly open the notes field for every lab result, as it is often unnecessary. Typically, the notes field is only opened if the clinician has a doubt or question about a value.
- If a test result exceeds defined critical values, then it is typically indicated with a double arrow. However, there is no indication on any screen to see what the critical values are or if they exist for that test.
- A common practice when under time pressure is to scan for flagged values and then read any notes that correspond to the flagged values. When flags are disabled, there may not be a perceived reason to check the notes for every unflagged test that appears normal.
- One of the screens used by clinicians does display the flagging range for each test result (See Figure 14 in Appendix A). However, the field is not blank when the flagging range is missing. Instead, the field would list the units in the brackets that are normally used for the range (i.e., “[mcg/dL]”). As a result, a missing flagging range is not prominent on the screen and can be missed.

Cortisol Test Context

The nature of cortisol testing introduces unique challenges unlike other common PST tests. Training, scheduling, interpretation, and documentation can require special consideration due to the complexity of accurately assessing and responding to the results.

- Cortisol was added two years ago as a standard test in PST for patients with previous immunotherapy. When cortisol was added, clinicians received training on the effects of low cortisol and the reason for testing but not detailed guidance about how to interpret cortisol or when specialist consultation is needed.
- Cortisol levels vary throughout the day—highest in the morning and lower throughout the day. The most accurate and well-understood cortisol reference range is the range at 8 a.m., which is

ideally when the sample would be drawn. The patient's blood sample was taken at around 2 p.m. Presurgical testing is done whenever the patient comes in for the appointment and it is rarely at 8 a.m.

- Most test results are easy to interpret. Cortisol is one of the more difficult tests to interpret because of the many contextual factors that must be considered, including the time of day the blood was drawn. In some cases, a specialist is needed to properly interpret cortisol levels.
- Cortisol levels are influenced by a multitude of factors beyond time, including steroid use. Endocrinologists often need multiple data points to decide whether a value is high or low. Endocrinologists consulted during this CAST analysis confirmed that non-specialists would have difficulty fully evaluating a patient's cortisol lab result.
- The Endocrinology department did not provide a memo to PST regarding how to interpret cortisol or when specialist consultation is necessary, but they would be happy to do so if asked.
- Clinicians in practice don't know what critical values exist or don't exist. Their understanding of what tests have critical values is often based on their past experience and what tests they remember getting calls about.
- PST cortisol tests are not routinely ordered with accompanying documentation about the reason for the test or a follow-up plan.

Scheduling and Workload Context

PST faces challenges related to scheduling and workflow that compound the difficulties in reviewing and interpreting the patient's lab tests. These challenges involve limitations in scheduling visits, resource allocation, high workload, and support systems available to clinicians. Together, these issues increase the difficulty of interpreting tests, especially those like cortisol, and contributed to the oversight in this event.

- PST has no control over their patient scheduling, including which patients are scheduled for morning or afternoon visits. PST scheduling is handled by surgery.
- PST patients requiring cortisol testing are not routinely scheduled for a morning visit (as recommended for cortisol testing).
- PST scheduling templates allow limited time for each visit. The template includes about 20 minutes in the room with the patient and 40 minutes for documentation, follow up, review, and any other new tasks or interruptions that arise. There is no dedicated time to review a patient's lab results, to prevent interruptions when doing so, or to recover when other tasks or interruptions cause delays.
- The PST APP was the unit lead in addition to the normal casework. This means the APP had more duties between patients, but no extra time is allocated to accomplish the extra workload.
- Other departments provide clinicians with administrative days as a dedicated time to review lab results and complete other tasks. PST APPs are not given administrative days and must complete all tasks in between the scheduled patient visits.
- The nature of PST means that most clinicians are meeting the patients for the first time during this visit.
- Test order sets are developed by specialists. The individuals who are responsible for reviewing the results may not have complete insight as to why a test was ordered, added, or removed.

- APPs may not have quiet workspaces that are practical to use for uninterrupted reviews of lab results.
- APPs may not have time to get to a protected workspace to look at labs between patients, even if such a place exists.
- There is no ability for APPs to review results using portable devices in a way that is effective and meaningful.
- A regular PST APP shift is ten hours long.
- Order sets are particularly relied upon in presurgical testing where PST APPs are reviewing case files for many sorts of patients with very different concerns.
- The APP may need to use multiple EHR systems throughout a patient’s visit. There is a different system for anesthesia, labs, and scheduling. They need to sign in each time they switch systems, which can present distractions and reduces short term memory.
- PST APPs may view presurgical appointments with a checklist mindset. The way the job is performed can feel like checking steps off of a list to approve/not approve a patient as opposed to taking a holistic view of the patient.
- Most APPs are not assisting physicians anymore—they have their own workload now.
- APPs do not have access to assistance from nurses. Normally, nurses can act as a secondary review of lab results, but this has not been an option for APPs.
- The historical philosophy of the medical center is that nurses are not there to support APPs. There have been efforts to change this, but many APPs do not have assistance with any tasks.
- Many labs need clinical context, including patient history, background, and current and planned treatment. PST APPs have rarely seen the patient before and may not have time to review the patient’s file in order to generate a holistic view of the patient or understand which lab tests should be ordered and why.
- The values 0.5 and 5.0 may look similar to an APP who is trying to review results while also accomplishing other tasks. This may be especially likely if the 0.5 is very unexpected but a 5.0 is very typical, and the patient was not presenting with symptoms that the APP associates with adrenal insufficiency.
- PST APPs are present at all medical center facilities, resulting in no shared overhead duties, which can spread them thin at each location.

Laboratory communication

The lack of special communication from the laboratory regarding the extremely low cortisol was considered normal, not unusual. The contextual factors behind the lack of communication involve the critical value policy, the high volume of laboratory tests, laboratory automation, and information flow between clinicians and lab personnel.

- The low cortisol test did not trigger a required call to the APP because there was no established critical value for cortisol at the medical center. Low cortisol was not deemed by the medical center to meet the definition of a critical value that requires a call.⁶
- The laboratory runs thousands of tests per day. Laboratorians cannot call for every lab test and physicians will not be available all day.

⁶ For more information about critical values and how they are defined, see Appendix A.

- The lab test may have been fully automated so the laboratorian may not have been aware of this test's low value.
- It is not common for a laboratory to contact an APP with a clarification/context question after an order. It is more common with other orders, like radiology (though the order volume for clinical laboratories is usually much higher). Ordering clinicians may not include context with an order, as they may not believe that the laboratorians have enough clinical knowledge to recognize the implications.
- Laboratorians often lack the patient's full medical context to determine whether a value is indicative of a harmful condition. This is especially true for non-automated, specialized tests which require an understanding of what is being tested and which results are being sought to determine which tests are best. For example, if the clinician is concerned about how low a value is, then the lab tech may prefer a test method that has higher sensitivity for detecting low values. For some tests, the lab techs review the patient's medical record to deduce what the clinician is looking for, to fill in the context that is missing in the order, and to select the best test. However, that is not practical for most tests.

Initial Recommendations

At this point, some initial recommendations can be developed related to the laboratory data and how it is transmitted and provided to clinicians.

- Recommendation:** Create two different orders for Cortisol a.m. and Cortisol p.m. to accommodate the different reference ranges and enable flagging for the appropriate range.
- Recommendation:** Develop the capacity for clinicians to easily understand which test values do not have a flagging range defined (meaning flagging is disabled). For example, colors and symbols could be used: red for values in the flagging range, green for values not in the flagging range, and black for values that do not have a defined flagging range and for which flags are disabled.
- Recommendation:** Identify and review the cases where reference ranges are only described in a notes field, which creates a human error trap and increases the chance that an abnormal value will be missed. Develop processes and guidelines to reduce the number of lab tests in this category.
- Recommendation:** When lab results are displayed, separate the field that displays the reference range and the units to make it more prominent when the reference range is missing. Currently, the field normally used to communicate the reference range (the flagging range field) will never appear blank even if the range is missing. Instead, the field will show the units (e.g., “[mcg / dL]”) in the same brackets used to communicate the range, which makes a missing reference/flagging range less prominent and increases the human error of overlooking the missing range.

Role in Adverse Event

- The oncologist did not identify the adrenal insufficiency following the PST lab tests, PST APP consultation request, and notification by PST of ongoing symptoms.

Mental model:

- The oncologist may have believed that the symptoms of fatigue/high heart rate were caused by a pulmonary or cardiovascular problem.
- The oncologist may have believed they did not need to monitor the patient's cortisol levels.
- The oncologist may have believed that unsafe values would be flagged, like other tests.
- The oncologist may have believed there is no need to check the PST lab results when notified of a problem and consultation request, as that is the PST APP's responsibility.

Context

- The lab values provided to the oncologist were not flagged as abnormal by the system.
- The oncologist may not have had time to review the lab results for every presurgical patient.
- Cortisol tests may not have been ordered prior to the PST appointment following immunotherapy.
- Cortisol values may not have been tracked longitudinally over the patient's treatment.
- The oncologist may not have been responsible for evaluating the patient before surgery or ordering tests.
- Usually, when a lab test is ordered, the responsibility for recognizing and responding to abnormal values lies with the ordering clinician. However, in this case the ordering clinician was a PST APP. PST provides a unique role in that their primary purpose is not to treat patients; they are clearing patients for surgery, and they rely on others to oversee patient treatment. In this case, the oncologist, whose role is similar to a patient's primary care physician, had a crucial responsibility to review the lab results from other providers, recognize abnormal values that indicate an underlying concern, and initiate treatment. While there are mechanisms to ensure an ordering provider receives and reviews the results that they ordered, there are not strong mechanisms to ensure the oncologist receives and reviews results ordered by others, like PST. This factor also impacts investigations after an adverse event. In this case, prior to CAST, most of the focus was on the PST APP who ordered the test. There was no exploration of why the oncologist did not identify the abnormal results, did not follow up with the patient after notified by PST and UCC of patient symptoms, and did not follow up after the echocardiogram and other efforts did not return a positive diagnosis. The groups that typically investigate medical errors chose not to investigate this event, assuming that the mistakes were made by the ordering clinician, and that no MDs were involved.
- The responsibility for identifying and acting on laboratory values is typically with the ordering physician (regardless of who treats the patient).
- The oncologist may have seen the test result in a view that made it difficult to notice the abnormal value.
- The oncologist may not have suspected adrenal insufficiency and therefore may not have paid extra attention to the cortisol values in the lab results.
- Oncologists receive emails with lab results of any patient visit to the UCC. However, in this case the labs were taken during PST and not as part of the UCC visit. The PST labs would be available

to the oncologist through separate means, and it may not have been obvious they were related to the UCC visit.

- There are different quality groups for APPs and for MDs that cross departments.

Initial Recommendations

Recommendation: **The Quality Assurance group should investigate why the oncologist was not able to identify the low cortisol level and diagnose adrenal insufficiency.**

Recommendation: **Tests for anticipated side effects should be ordered by the team that is in charge of the patient’s care and treatment, not PST. PST may repeat tests to verify the patient’s current condition is stable.**

Recommendation: **Oncologists should assume the role of the primary care physician and care coordinator for oncology patients. All tests, consults, and other healthcare visits should be sent to and reviewed by the oncologist. If these policies already exist, then investigate why they were not effective and address the gaps.**

Remaining Questions

- How are patient's symptoms tracked between patient visits?
- What is the procedure for monitoring patients for adrenal insufficiency after immune therapy treatment? Is there anyone other than PST clinicians with this responsibility?

Questions for Other Controllers

- How do the oncologist and surgeon coordinate to ensure the patient is ready for surgery?

Surgery Department

Description

The surgery department schedules PST visits according to the templates defined by department leadership. During a PST visit, PST will notify surgeons of potential health concerns and request recommendations to potentially assist or alter the treatment plan. The surgeons have access to all PST lab results to assist when requested.

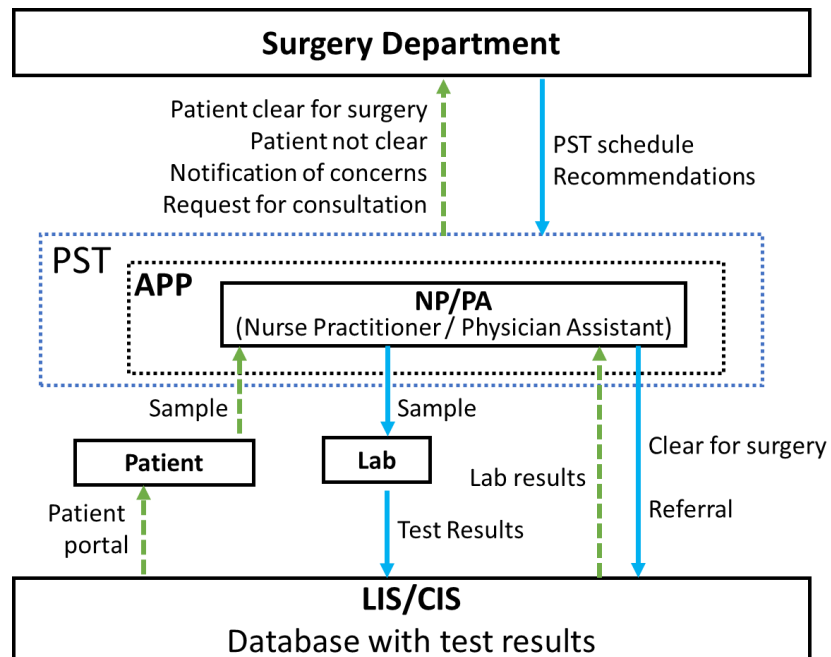


Figure 5: Control Structure Including Surgery Department

Relevant responsibilities

- Manage the treatment plan for patients undergoing surgeries
- Schedule PST patient visits
- Receive PST notifications about patient concerns and support development of corrective actions as appropriate
- Ensure the patients are fit for surgery

Role in Adverse Event

- The surgery department schedules PST patients, including those who require cortisol testing, at various times of day. However, best practices for cortisol testing suggest that the tests should be done at approximately 8 a.m. to better enable accurate interpretation.
- The surgery department scheduled PST patients in 30-minute blocks without dedicated time to review lab results.
- The surgery department did not recognize the low cortisol lab result nor diagnose the adrenal insufficiency following the PST lab tests, PST APP notification, and consultation request.

Mental model

- Surgeon may not have been aware that cortisol flags were never defined, increasing the interpretative burden.
- Surgeon may not have been aware that abnormal cortisol results would be displayed in the same way as a normal test result.
- Surgeon may not have been aware that PST APPs did not have adequate time or workspace to review lab results between patients, increasing their risk of missing abnormal labs.
- Surgeon may have believed that PST notifications and consultation requests do not require review of the lab results to assist with diagnosis.

- Surgeon may have believed that the responsibility for responding to PST concerns and tests lies with PST, not Surgery.

Context

- The surgery department schedules PST visits according to a template that is defined by hospital administration. The template required 30-minute blocks without dedicated time to review lab results between patient visits. PST attempts to assign one patient per hour to a clinician, meaning two different clinicians would be assigned to handle two consecutive 30-minute patient blocks.
- The surgery department may not have been informed of concerns regarding the PST scheduling templates, and therefore may not have a reason to communicate those concerns to the authors of the templates.
- Most groups are able to manage their own schedules and have the ability to adjust schedules to meet their needs. PST is one of the few groups that do not have the ability to manage their own schedules.
- The surgery department might not be aware that PST would prefer to schedule patients differently than how it is done today. For example, elderly patients would ideally be scheduled for morning visits due to certain challenges PST has encountered.
- The surgery department may not have a process in place to routinely review every patient and their PST labs upon PST notification and request for consultation.
- The surgery department may schedule same-day visits for patients for PST to reduce patient burden, but this can unintentionally reduce the time PST APPs are expected to use for administrative tasks.

Initial Recommendations

Recommendation: **Increase coordination between PST and Surgery regarding PST scheduling. For example, Surgery must ensure that patient context is factored into scheduling so that PST APPs are better prepared to accurately interpret patient lab data.**

Questions for Other Controllers

- The surgeon agreed with PST's decision to refer the patient to the UCC for further assessment. Why was UCC unable to diagnose adrenal insufficiency?
- Who determines how PST schedules are made and what information Surgery uses to schedule appointments. How did they make this decision?

Urgent Care Center

Description

The Urgent Care Center sees patients for a variety of reasons. UCC has access to previous lab results and the patient's medical records. UCC evaluates if the medical patient is medically fit to be discharged, and can provide diagnosis, treatment, referrals, or order new tests.

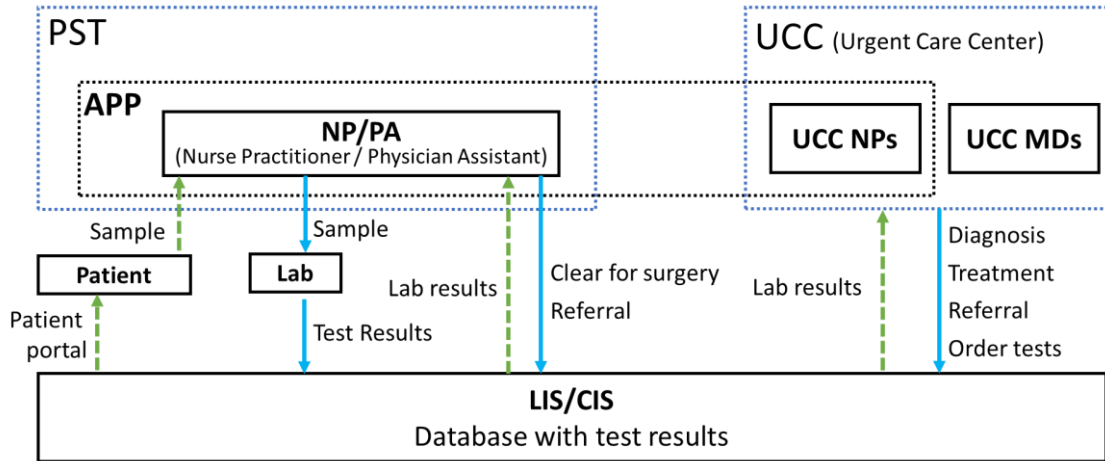


Figure 6: Control Structure Including PST and UCC Interactions

Relevant responsibilities

- Diagnose patients experiencing acute medical concerns based on symptoms, history, and test results.
- Prescribe appropriate treatment or refer to appropriate specialists as needed.

Role in Adverse Event

- UCC did not recognize the patient's low cortisol level and diagnose the patient with adrenal insufficiency.
- UCC did not refer the patient to Endocrinology or begin treatment for adrenal insufficiency.

Mental model

- UCC provider may have believed the patient's symptoms were primarily cardiovascular in nature.
- UCC provider may have believed the PST labs were already evaluated by the PST APP and ruled out non-cardiopulmonary diagnoses.
- UCC provider may have believed the PST labs would be flagged automatically if they were abnormal, similar to other tests.

Context

- The primary role of the UCC is to address acute complaints. The UCC has a high patient volume and APPs and physicians do not have time to review the full medical record for all patients.
- UCC may have been focused too narrowly on PST's concern for pulmonary embolism without reviewing earlier lab tests or considering other potential problems.
- The UCC did refer the patient for a follow-up echocardiogram. The UCC may not have had the resources and time to identify rationale for symptoms that did not appear to be acute in nature.
- After the patient is discharged from the UCC, they are typically not followed by UCC staff unless they return. The UCC wouldn't revisit their original diagnosis unless the patient elected to return. The UCC staff may not be familiar with cortisol labs as it is not something they would regularly evaluate. Without flags, and without any indication that flags are disabled for cortisol, it would not have stood out. The notes with text describing the reference ranges, which

sometimes require additional actions to view, may not have been sufficient for UCC to identify a non-flagged cortisol value as potentially harmful.

Initial Recommendations

Recommendation: When labs are displayed, especially to non-ordering clinicians, the clinical context describing the rationale for the order should be clearly displayed. Communicating rationale to future viewers of the test results is particularly useful for clinicians in contexts where they need to make quick decisions and may not have a history of treating that patient (such as in the UCC).

Questions for Other Controllers

- How does the laboratory determine what information will accompany the test result, such as notes and enabling/disabling flags?

Laboratory

Description

The lab's goal is to provide the most accurate results possible in a timely manner. They receive millions of test orders per year from across the medical center, and these results are returned through the EHR system. Physicians are contacted immediately if a result crosses a critical value threshold. The lab configures the LIS and EHR, including entering the reference ranges for a test, entering the flagging ranges, and entering notes as required by regulation. The lab consults with the Endocrinology department to identify appropriate reference ranges.

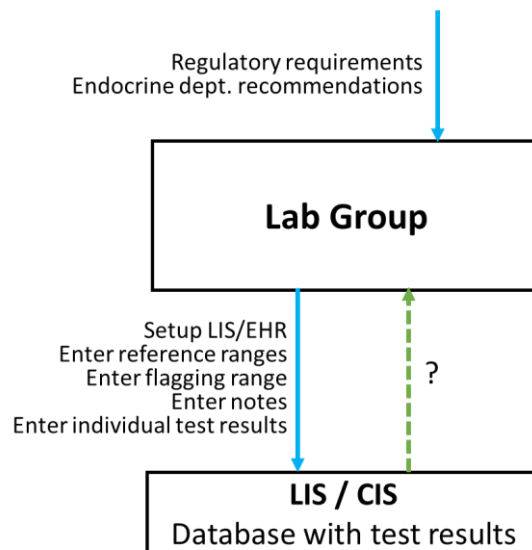


Figure 7: Control Structure of the Lab Group

Relevant responsibilities

- Run ordered lab tests.
- Enter lab results into LIS.
- Establish institutional reference ranges.

- Setup LIS/EHR to flag abnormal values outside the reference range, as appropriate.
- Call/notify ordering physician with any critical results.
- Comply with relevant regulatory requirements.

Role in Adverse Event

- Lab group did not program the flagging range for cortisol to enable automatic flagging of abnormal results.

Mental model

- The lab group believed that text-based notes were an effective alternative to flagging and programmed flagging ranges.
- The lab group may believe that clinicians always read the notes when assessing lab results.
- The lab group may believe that clinicians are familiar with symbols (like “*”) that indicate a note exists.
- The lab group may not be aware of which result views clinicians are using, such as the view that does not automatically show reference ranges or notes.
- The lab group may have assumed that clinicians know which tests have flagging enabled versus those tests that need additional clinical interpretation.

Context

Lab IT System Context

- There is not an easy way to implement two different flagging ranges that consider the time of lab draw for diurnal values, like cortisol.
- Currently there is only one cortisol test that can be ordered and result. Due to technical limitations of the LIS/CIS, the only way to add diurnal flags may be to create different orders.
- Other institutions have solved the diurnal reference range problem by splitting cortisol orders into a.m./p.m. orders. On paper, that would increase the total menu of orders, although the actual cortisol test method is identical.
- One drawback of defining separate a.m./p.m. cortisol tests is that the time of a patient’s visit may not always be known when the test is ordered. Procedures exist in both clinical and laboratory environments to detect and correct errors if the type of test ordered does not match the time of the blood sample.
- There are LOINC (Logical Observation Identifiers Names and Codes) codes that allow cortisol tests to be coded by a.m. (57824-5) or p.m. (57825-2) [14], in addition to codes for each hour of the day.
- The medical center has currently placed a moratorium on making changes like splitting a test into two orders until the new information system is in place next year. Any changes would have to be added to the new system, which could delay the EHR implementation schedule.
- The lab generally does not know what time the test was drawn unless they search for the information. They do not immediately have this context when reviewing a result. Therefore, it may not have been easy to interpret the result or select the appropriate reference range.
- There are several tests that do not have standard reference ranges, as the ranges can depend on contextual factors. The information system does not have a way to encode context-dependent flags, so many of these tests do not have flags entered at all.

- There are no national standards for how tests should be flagged. Hospitals set their own guidelines, and it can be difficult to compare institutions to identify best practices or common deficiencies. One factor making comparisons difficult is that different patient populations may have different reference ranges (cancer institute vs. pediatrician for example).

Clinician-Lab Communication Context

- The lab might not have known that a different group (PST) with less knowledge of cortisol values is now ordering the cortisol tests as a standard practice.
- Endocrine specialists usually schedule patients for a morning cortisol test, ideally at 8 a.m. when the reference range is well understood. The lab may not have had a process to track that a substantially higher percentage of the tests were now being drawn in the afternoon.
- Reference ranges are developed in coordination with endocrinologists.
- The laboratorians do not receive detailed information about who ordered the test and whether they are an expert (endocrinologist) versus a generalist, like a PST APP.
- Cortisol has not had flags enabled for 15 years at this medical center. No reviews have been conducted to identify previous reports or instances of low cortisol values that were overlooked.
- Clinicians may send optional contextual information about a patient if they believe it could impact the lab test performance (e.g., if they expect a very high/low result). However, this is not a common practice, and most clinicians do not have time to do this for all orders.
- There are limited options for alerting physicians of lab values. A new type of lab result alert called an “urgent value” was introduced at the medical center that requires an email to be sent. Urgent values fall between an abnormal value and a critical value in their urgency, and they may be useful for preventing similar adverse events. However, not all clinicians are aware of urgent values, urgent values have not yet been defined for cortisol, and there are no regulations that require urgent values to be defined.
- There is no way for clinicians to identify the types of automated alerts they want for different patients/tests.
- It is easy for laboratorians to see when flagging range fields are blank (thus no flags) when using the LIS. However, clinicians may not see or understand when the flagging range fields are missing.
- In the past, physicians at the medical center have occasionally caught missing flags for the few tests they are intimately familiar with. This voluntary reporting approach catches missing flags by chance after they were already missing (reactive) rather than proactively preventing missing flags from the start.

Specific Institutional Context

- There may not have been a defined process for identifying lab values (such as cortisol) that are missing flags, reviewing that decision, and adding flags proactively rather than after an adverse event.
- Clinical Chemistry has the greatest number of tests at the hospital with over 1 million tests per year (thousands per day). Many larger medical centers have a smaller test menu.
- Changing test reference ranges can be challenging for some hospitals, because patients may be routinely monitored for values longitudinally given that small changes can signify a significant health concern. Therefore, if a test report changes, and values are no longer comparable from the previous week, it could disrupt care.

- Most tests are not time-sensitive, so creating new processes for time-sensitive tests was not a high priority.
- It is difficult to ascertain standard practices across the industry, such as how many medical centers have different cortisol orders for different time ranges. If hospitals code their tests differently, then it could be difficult to share and compare test results.
- The lab focuses on ensuring the correct values are returned and that equipment is working correctly, but interpreting results is seen as outside their responsibility.
- There are insufficient human resources to manually review all flagging protocols on a regular basis. Because lab results are not often attributed to adverse events, it is difficult to justify allocating resources to reviews.
- There are no procedures currently in place in the laboratory to monitor test ordering trends and identify new parties who are ordering specific lab tests, like cortisol.
- Tests like cortisol may not trigger special reviews, as it is considered a standard test and has been available for a long time.
- Most labs do not have procedures in place to monitor test ordering trends or compare or review flagging protocols. In fact, most of these contextual factors are common throughout the industry.

Regulation

- Labs are regulated to ensure they have appropriate reference ranges. However, fewer regulations exist regarding the display of reference ranges on the screens that are used by clinicians and the flagging of test results outside the reference range. A comment stating, “See reference range” may be argued to comply with regulations, but there is not a clear consensus among labs, accreditation bodies, and inspectors regarding what is acceptable.
- Reference ranges may not always be equally relevant for all tests, but the regulations may require reference ranges anyway. For example, for certain drug tests, the reference range for a standard patient may be 0-0 (reflecting no trace of drugs in their system). However, this may not be helpful information. The lab test may return a value of <0.01, indicating no trace was detected, but it can be problematic when compared to a reference range 0-0.

Remaining Questions

- Is it possible for a generic test (such as cortisol) to be automatically returned to the physician as a more specific test (such as cortisol p.m.) with contextual reference ranges?

Questions for Other Controllers

- Who determines funding for and staffing for laboratory quality initiatives?
- How do regulatory authorities determine if and how to regulate flagging behavior?

Initial Recommendations

Recommendation: Minimize the number of tests that rely on comments to convey reference ranges. Identify which labs currently rely on comments fields to convey reference ranges, review them periodically, and identify possible alternatives.

Recommendation: Periodically audit the number of tests with reference ranges that do not have flagging ranges. Identify solutions to minimize the number

of tests with reference ranges that do not have a corresponding flagging range.

Recommendation: For tests that have a reference range but no flagging range, define a process to compare the reference and flagging ranges with other institutions. Identify improvements to the way the reference range and flagging range is entered into the LIS/CIS.

Recommendation: Identify solutions to provide feedback on whether and when test results are actually received and presented to clinicians. There should be a mechanism to detect when the lab group believes clinicians are reviewing certain test data (such as hidden notes) but in reality, the lab data is never presented to clinicians. Additionally, there should be a mechanism to detect when the final lab report (which the lab group assumes is received and reviewed) is never actually presented to clinicians, and instead, the clinicians only see a summary screen.

Although the lab group believes the comments field is an effective way to communicate reference ranges and other critical information, the field is hidden in commonly used EHR screens and has been shown to be ignored by most clinicians. While it is tempting to simply provide more training for clinicians, the field of human factors recognizes that training and education are the least effective strategies to improve safety. The most effective strategy is to fix the context and the systems that are used by humans.

Chemistry and Endocrine Services

Description

Chemistry and Endocrine services provide consultation or guidance on interpreting a test like cortisol when asked to do so. Chemistry and Endocrine services also provide expertise regarding appropriate reference ranges for lab tests, like cortisol.

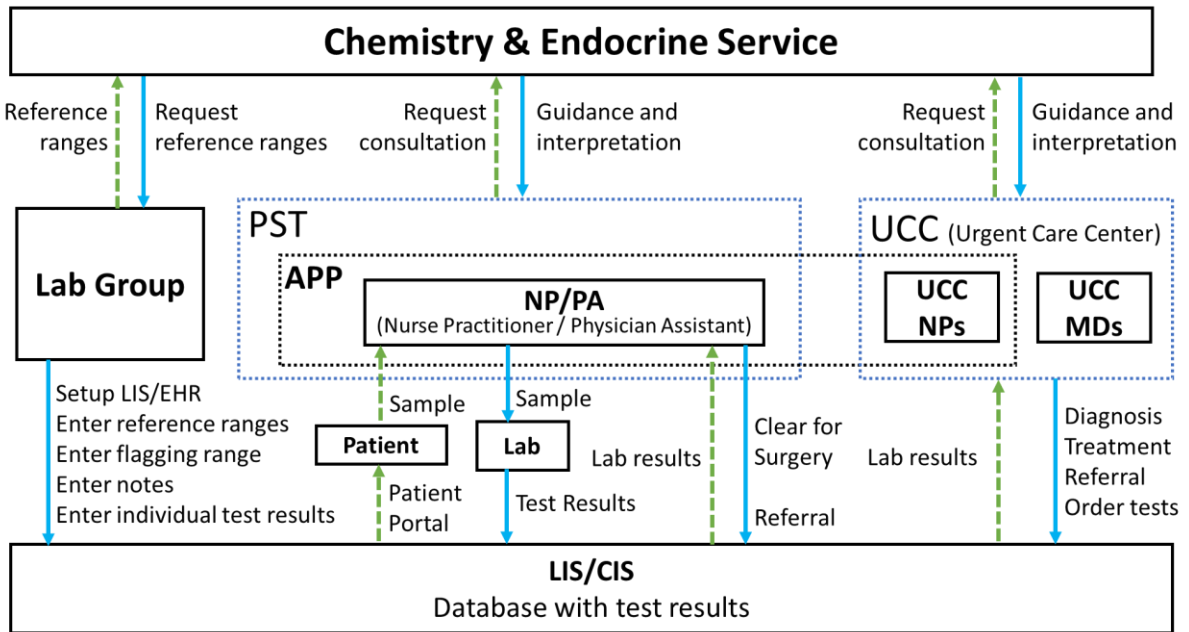


Figure 8: Control Structure Including Chemistry and Endocrine Service

Relevant responsibilities

- Collaborate with the laboratory to determine reference ranges and flagging for chemistry and endocrine lab tests.
- Provide consultation for clinicians when requested.
- Diagnose and treat patients with adrenal insufficiency.

Role in Adverse Event

- Did not request cortisol flags when it began to be used by groups without the necessary specialization to fully appreciate cortisol contextual factors.

Mental model:

- May have believed that clinicians ordering cortisol tests (e.g., PST APPs) would be able to identify unsafe cortisol levels with reference range notes (as opposed to flagged values).

Context

- Cortisol had been ordered without a known adverse event for the past 15 years.
- Until recently, most cortisol tests were ordered by specialists and endocrinologists who could readily interpret cortisol values without flags and may have recognized from experience that the flags were not available for cortisol.
- Although endocrinologists are very familiar with cortisol interpretation, other departments have less familiarity with the test and its contextual dependencies.
- Immunotherapy is rising as a treatment for cancers, therefore adrenal insufficiency was a concern for more patients.
- Reports of missing flags usually come from practicing clinicians with specialized knowledge who happen to notice it (not typically identified by a quality group or audit).
- The lack of flagging may have been reported at some point but not acted upon due to the technical difficulties of applying a different reference range in the EHR / LIS.

Initial Recommendations

Recommendation: Request that Endocrinology develop a memo to aid PST APP in interpreting cortisol levels, including clarifying when a specialist consultation may be necessary.

Recommendation: When future PST tests are added, consult with Endocrinology to determine whether similar memos may be appropriate.

Questions for Other Controllers

- Does department leadership direct Endocrine services to train or otherwise support clinicians who may need to interpret endocrine tests?
- How many of the contextual factors identified in CAST are within the control of department leadership?

Department Leadership

Summary

Department leadership are the groups that are responsible for managing the day-to-day operation of departments, such as PST, Laboratory Medicine, Oncology, UCC, etc. Responsibilities vary across the organization but can include schedule management, training, safety reviews, quality assurance, and more.

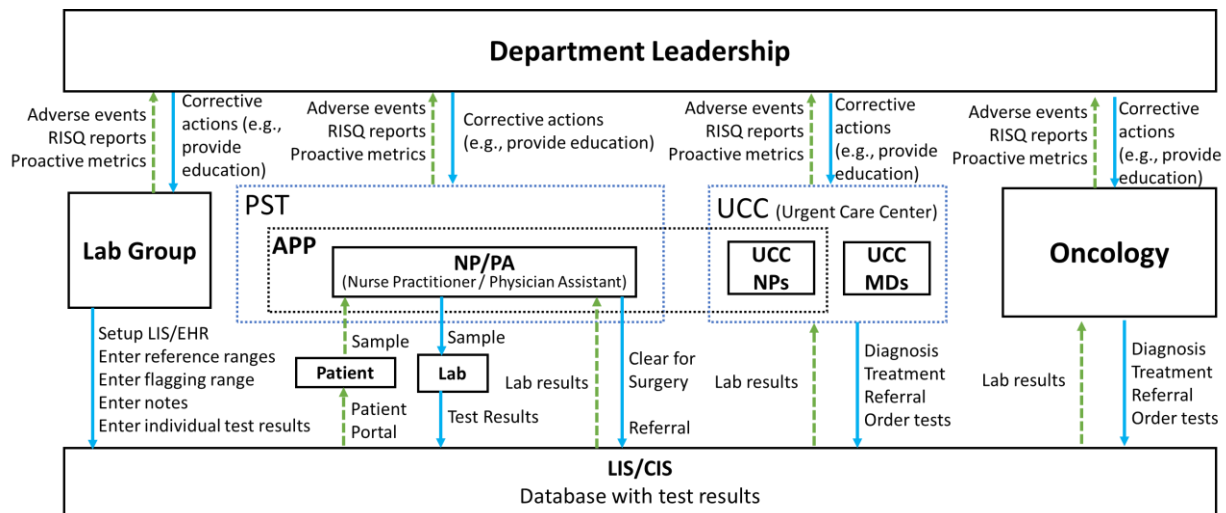


Figure 9: Control Structure Including Department Leadership

Relevant responsibilities

- Ensure clinicians have the information and training necessary to complete their tasks.
- Ensure all patients receive the appropriate standard of care.
 - Review and update procedures as standards of care change.
- Ensure labs are communicating accurate and complete information, as necessary, to support clinical interpretation.

Role in Adverse Event

- PST leadership did not ask the lab to add cortisol flags when cortisol was added as a standard test order for PST clinicians who do not have the expertise to fully interpret cortisol contextual factors.
- PST leadership did not ask Endocrinology to provide a cortisol interpretation guide when cortisol was added as a standard test order for clinicians without the expertise to fully interpret cortisol contextual factors.

Mental model:

- Believed that PST APPs understood how to interpret cortisol.
- Believed that PST APPs would be able to identify unsafe cortisol levels using a standard reference range.
- May not have been aware that cortisol flagging was missing / disabled.
- May have believed that reference range notes (rather than a programmed reference range or flagging range) would be sufficient.
- May have believed that flagged values were not needed to aid cortisol interpretation.

Context

- Cortisol had been a standard test without a known incident for 13 years prior to its use by PST and for 2 years following its use by PST.
- Cortisol is very familiar to endocrinologists, who had been the primary user of cortisol testing. Other departments may have less familiarity.
- Immunotherapy is rising as a treatment for cancers; therefore, adrenal insufficiency is a concern for an increasing number of patients.
- Reports of missing flags have usually come from practicing clinicians (not typically identified by a quality group or audit).
- The lack of cortisol flagging may have been reported at some point in the past, but not acted on due to technological limitations (may require defining separate a.m./p.m. cortisol tests, as in other institutions).

Role in Adverse Event

- Lab leadership did not compare flagging ranges (flagged values) to best practices at other institutions for tests like cortisol.
- Lab leadership did not compare test order definitions (e.g., a.m. vs. p.m.) to best practices at other institutions for tests like cortisol.

Mental model:

- Believed that flagging ranges are not regulated and associated requirements do not exist
- Believed that reference ranges primarily drive clinical decision-making rather than flagging ranges or flags. Therefore, believed it is sufficient to simply compare reference ranges with other institutions, not the flagging ranges or test order definitions. In reality, sometimes the flagging range is equally (if not more) relevant for clinical decision-making than the reference range.
- Believed the medical center was compliant with relevant requirements since they had never been cited by inspectors for the missing cortisol flagging ranges.
- Believed that distinct a.m./p.m. cortisol tests were not necessary.

- May not have been aware that other institutions define separate a.m./p.m. cortisol tests to enable flagging and to minimize clinical errors.
- Believed that if there were an issue, clinicians would have recognized it and reported it.

Context

- There is no clear indication to clinicians when flagging is disabled. Clinicians who don't know that cortisol flagging is missing may be more likely to misinterpret cortisol. In other words, clinicians cannot report what they don't know is missing.
- There are committees and reporting mechanisms that allow clinicians to raise concerns they may have. However, there is no specific feedback sent to laboratory leadership to validate assumptions about which screens the clinicians actually use to review test results or how often the test notes are actually displayed to the clinicians in practice.
- There is no feedback available to laboratory leadership about how prevalent missing flagging ranges are within the test menu.
- It is common practice (and a regulatory requirement) to periodically review reference ranges and critical values, which can include comparisons to best practices at other institutions.
- There is no standard practice or regulatory requirement to compare flagging ranges or test order definitions to best practices at other institutions.
- There are regulatory requirements related to flagging and displaying reference ranges, but they are phrased in a way that could be subject to interpretation.
- Inspectors have limited time to review compliance, and they typically only review a sample of all lab tests. The lack of a citation does not mean that every test complies with all requirements.
- The medical center has been cited in the past for missing flagging ranges (disabled flags) on other tests, but the inspections are not always consistent because the regulations can be interpreted in different ways by different inspectors.
- There is more emphasis on reactive changes than on proactive changes. In practice, a RISQ (Report Incidents of Safety and Quality) report or an adverse event would be needed to trigger a change in a flagging range. There are few triggers in place to proactively identify this type of vulnerability and make necessary changes.
- There is no industrywide knowledge about how widespread flagging problems are. This may partly be due to adverse events not being investigated systematically or deeply. When investigations stop at blaming individuals immediately involved in an adverse event, they do not go far enough to identify CLIA gaps. Also, adverse events are usually investigated by the hospitals themselves—not an independent group. Therefore, they have little interest in examining causal factors that originated outside the hospital and therefore are not controlled by the hospital.

Role in Adverse Event

- APP leadership provided training on cortisol testing that explained why it was added and potential outcomes but did not offer detailed guidance about how to interpret cortisol levels or when to consult a specialist.

Mental Model

- May have believed that the EHR would flag abnormal values.
- May have believed that the reference range would be visible in the EHR in a similar way as other test results.
- May have believed the EHR notes were sufficient to explain the multiple reference ranges and ensure abnormal cortisol values would be identified.
- May have believed that the PST APPs would have the time and ability to review all lab results carefully and with full attention.

Context

- Most PST tests have flagging enabled.
- Most PST tests are not difficult to interpret and do not require a specialist.
- PST APP leadership has repeatedly reported concerns about time constraints and productivity measures that have resulted in an increasing number of mistakes over the past two years, but these factors have not been linked to specific adverse events and no specific actions were identified to address these concerns.

Questions for Other Controllers

- How does the hospital administration manage their information systems to ensure that clinicians are able to understand and utilize data appropriately?
- How does the hospital make decisions on departmental leadership responsibilities?
- How does the hospital administration respond to requests for resources for safety and quality projects?
- How does the hospital administration set targets for departments?

Recommendations

Recommendation: **Review other tests with diurnal reference ranges and evaluate adopting LOINC or other standard codes that allow a.m./p.m. or similar order designations.**

The planned transition to a new IT system, which will cause some continuity disruptions, may be an ideal opportunity to implement this recommendation. Adopting a new test definition during this period can prevent a second or tertiary disruption later. Note that this is a clinical perspective focused on minimizing the number of clinical disruptions. In contrast, the existing decision to place a moratorium on test definition changes until after the new IT system is in place takes a technical perspective, aiming to simplify the technical migration process. These trade-offs should be considered to balance the clinical needs with the technical requirements of the system transition.

Recommendation: **Create standard metrics to track the portion of common lab tests that have defined reference ranges but no defined flagging ranges. Periodically report these metrics to quality assurance groups and to hospital administration. Review potential ways to improve.**

Recommendation: **Establish regular audits to determine the extent to which PST APPs can identify the common PST labs that will not flag abnormal results. The results of the audit will enable more effective feedback to leadership related to safety and risk at the medical center. If the**

audits reveal a knowledge gap, then determine if it is possible to fix the information system to enable flagging. If a technical limitation prevents flagging, then it should be reported to the vendor, regulators, and industry peers. Training and education are the least effective mitigations, but these may be necessary if more effective measures cannot be implemented.

Recommendation: Define a process to compare test order definitions with other institutions for those test order definitions that do not match LOINC or other codes (e.g., with a.m./p.m. designations).

Recommendation: Assign a clear responsibility to appropriate groups to identify when critical values, urgent values, and flagging ranges are missing and should be added.

Hospital Administration

Description

The administration is responsible for organizational function and operation. The administration includes executives and high-level directors. The administration oversees the environment and selects the digital tools used in the medical center’s healthcare system. The administration has controls over the appointment lengths, staffing levels, and the availability and kinds of space for work among others. The administration also plays a role in what adverse events are investigated and how.

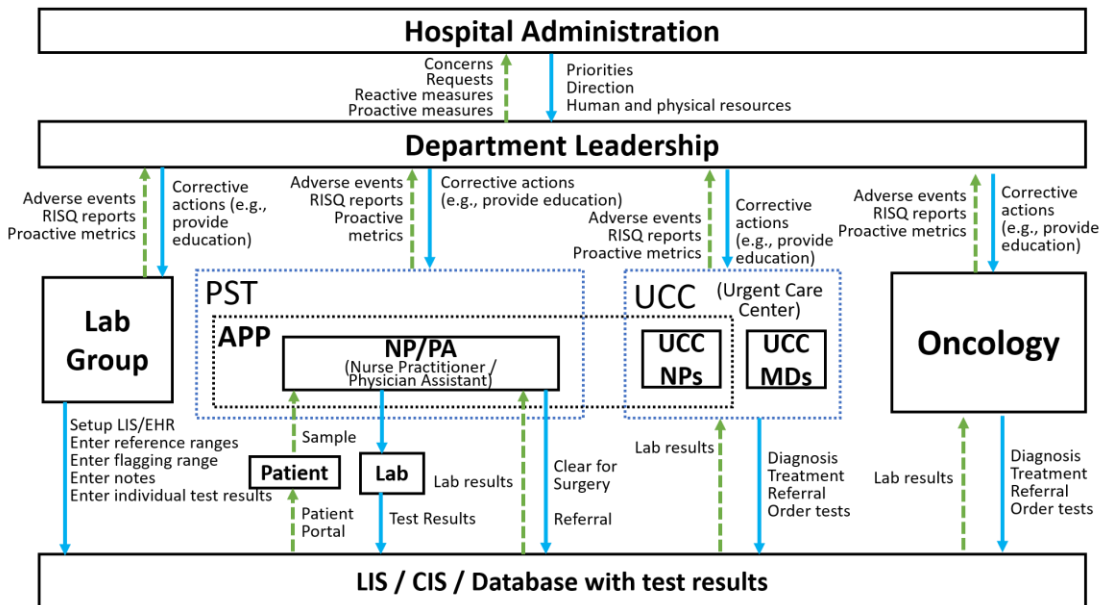


Figure 10: Control Structure Including Hospital Administration

Relevant responsibilities

- Purchase EHR.

- Set targets for patient volume.
- Set work schedules for APPs.

Role in Adverse Event

- Defined templates and scheduling constraints for PST APPs that did not allow sufficient time to complete all tasks.
- Defined 30-minute PST templates that did not allow sufficient time to review patient records, including test results.
- Defined and created workspace for PST APPs that may not have been sufficient for completing quiet, uninterrupted work.

Mental model

- Believed PST APPs did not need as much administrative time as other departments.
- May believe that APPs do not need administrative time and that when administrative time is given, it is not used to actually review patient files.
- May not have been aware of perceived time pressures that resulted when PST cannot maintain their own schedule.

Context

- The medical center has been increasing patient throughput steadily over the last decade and a half. The patient volume was not high enough to maintain a free-standing medical center and the goal was to raise the volume to a level comparable to peer institutions.
- There is a shortage of healthcare practitioners in the United States. There are not enough nurses and APPs to staff at optimal levels.
- Many staff members are experiencing burnout or stress from low staff levels, pandemic experiences, and organizational changes (like the switch to a different EHR system).
- Appointment lengths have been variable in the past (15-60 minutes). There has been an effort to standardize appointment lengths to 30 minutes across the board.
- 30 minutes is typical for other groups. However, PST involves additional challenges, as the clinicians are usually meeting patients for the first time and must review each medical record without the benefit of past interactions.
- The administrators may have a different view than the APPs as to what a quiet space looks like. For example, a common office space may be viewed as a quiet space to an administrator, but an APP might find that area chaotic if people are walking in/out, patients are coming through, or other interruptions are frequent.
- Surveys of clinicians have not been followed up on adequately in the past. There may be a lack of trust that reporting will yield changes, even if improvements are being made.
- The administration was aware of many of the generic APP contextual factors influencing the accident, such as time pressures. However, it may be difficult to justify resources to address potential problems absent of any identified and associated incidents.
- PST has reported concerns in the past about excessive time pressures that could lead to mistakes. However, the concerns have usually been reported without a specific adverse event as an example.
- Different groups may have assumptions about the time pressures experienced by other groups. For example, there may have been a belief that the PST APP had a day for admin

work, like reviewing test results, because other APPs have one. However, this is not the case.

Role in Adverse Event

- Assigned the task of PST appointment scheduling to surgery department (PST does not control their schedule or when patients are scheduled).

Mental model

- May not have been aware of advantages of PST maintaining their own schedule.
- May not have been aware of adverse events linked to scheduling.

Context

- Most groups outside PST have control over their schedule.
- Without control over their schedule, constraints (like ensuring morning visits for patients requiring a cortisol test) are not met. This can create inconsistencies, inaccuracies, and additional burden on interpreting context-dependent values like cortisol.

Role in Adverse Event

- Was not able to direct department leadership to address the gaps and deficiencies related to lab results and clinical interpretation.

Mental model

- Was not aware of the lab-related contextual factors identified above that increase the difficulty of cortisol interpretation and flagging.
- Was not aware that lab-related best practices at other institutions have resolved the difficulties related to cortisol interpretation and flagging.
- Was not aware of the PST APP-related contextual factors identified above that increase the difficulty of cortisol interpretation.
- Was not aware of best practices at other institutions that address the PST APP-related contextual factors.

Context

- There is no strong feedback available to administration to highlight and summarize the lab-related contextual factors identified throughout this report.
- Feedback existed about some of the generic PST APP-related contextual factors, such as perceived time pressures and lack of support staff. However, the generic feedback was not tied to specific adverse events or specific risks like cortisol interpretation.
- There is no strong feedback available to administration tied to specific risks, like the lack of indication that flags were disabled together and lack of dedicated time or space for clinicians to review results.
- There is no formal process or requirement that would proactively detect the specific lab-related and PST APP-related contextual factors to create an accurate indication of risk at the administration level. Most of the measures in place are reactive not proactive, requiring an adverse event before problems can be recognized, reported, and fixed at the administrative level. For example, there was not a process in place capable of catching the fact that 15% of common PST labs are missing a flagging range to enable flagging. Clinicians do not know (and have no clear indication to tell them) which labs are missing the flagging capability.

Clinicians tend to review labs quickly at a glance as they have no dedicated time to review labs (no administrative day and no buffer time in the scheduling templates), and PST notifications to Oncology and Surgery of patient concerns are not always followed up to assist.

- There is no standard policy about the type of education and training to be rolled out when new tests are added for PST APP. For example, one philosophy would focus education on how to interpret cortisol values, while another would focus on the reason for cortisol testing and the potential dangers.
- No one has a clear and specific responsibility to identify and verify test results that may not be clearly communicated or flagged to clinicians.

Initial Recommendations

Recommendation: Determine the extent to which PST APPs feel that leadership actions have prioritized quantity over quality and safety. If so, then determine why and address the issue.

Recommendation: Provide additional time for PST APP patient visits to enable careful review of lab results and to reflect the fact that PST does not have administrative days. PST must review each new patient without the benefit of a history of interactions with the patient. The additional time could be accomplished by increasing the scheduling template time, or by including a buffer time between patient visits to review lab results and other administrative work.

Regulatory Authorities

Description

There are numerous regulatory authorities that were involved in different aspects of this adverse event. The FDA requires that approved IVD devices have set reference ranges that are accurate. The Centers for Medicare and Medicaid Services (CMS) and state-level authorities define lab regulations (including policies) that dictate how information must be reported and to whom. CMS also creates incentives to ensure that care facilities use EHRs with specific functionality. The Office of the National Coordinator for Health Information Technology (ONC) sets standards and certifies that EHR systems are able to perform specific functions. Regulatory authorities can also authorize and designate inspectors to perform audits and ensure the relevant regulations are followed.

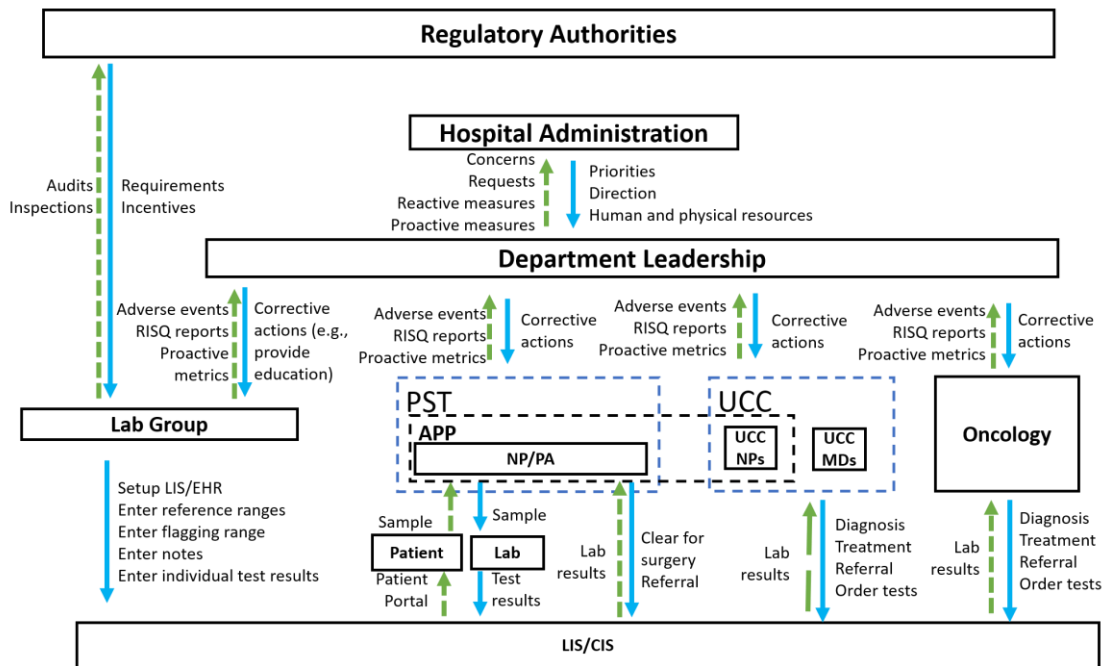


Figure 11: Control Structure Including Regulatory Authorities

Relevant responsibilities

- Set standards and approve EHRs for use in patient care.
- Ensure laboratory data is consistent, safe, and usable.
- Require and audit that laboratory data necessary for successful clinical interpretation is sent and appears in the final destination.

Role in Adverse Event

- Did not require processes for medical centers to periodically review their flagging ranges and flags, to perform comparisons with other institutions, or to identify best practices.
- Did not provide effective regulations or guidance regarding the laboratory data that must be transmitted, including flagging ranges, flags, and other information that was not transmitted by the lab.
- Did not define effective requirements about laboratory data that is presented or hidden from clinicians. For example, abnormal cortisol values could be presented in the same way as a confirmed normal value without violating any existing requirements.

Mental model

- Believed laboratories would be capable of determining appropriate flagging mechanisms.
- Believed laboratories would recognize that CLIA is incomplete and would voluntarily develop and implement additional requirements and audits related to flagging to ensure safety.
- Believed that administrators' intrinsic safety goals alone would result in medical centers recognizing and funding work to fill CLIA gaps with additional requirements.
- Believed that hospitals do not rely on regulatory requirements and guidance to determine which tests need flags and to establish processes to compare flagging ranges (in addition to reference ranges).

- May believe that other entities, such as accreditation agencies, have more stringent standards, such as checking for missing flags. They may have believed that the accreditation agencies had standards that covered this case.
- Regulatory groups may believe that labs will do more than required for CLIA. Regulatory groups believe that safety goals alone are enough to incentivize hospitals to implement the necessary requirements beyond CLIA and fill the known CLIA gaps.

Context

- The existing regulations have been effective in other cases. For example, the medical center was previously cited during inspections for similar instances of missing flags and flagging ranges on other tests. However, the language of the relevant regulations often includes phrases like, “if appropriate”, without any process or guidance to determine when it is appropriate. As a result, each inspector and medical center has a different understanding of when flags and flagging ranges are required.
- Reference ranges and flagging can depend on various contextual factors, and the understanding of which factors are relevant may change over time. Careful consideration is needed to create regulatory policies that are effective but broad enough to cover all cases meaningfully.
- One of the challenges of adequately regulating the use of reference ranges and flagging is that not all tests use reference ranges in a standard way. Some tests are interpreted only through the change in a value over time. In these cases, flagging a result that is low or high in an absolute way may not make sense. Instead, these flags would benefit patient care if they were able to identify when the change in value was low or high.
- There are no regulatory requirements to periodically review flagging ranges and flags or to perform comparisons with other institutions to identify best practices.
- United States Core Data for Interoperability (USCDI) V4 and USCDI V5 [11] both contain data fields for reference range and interpretation. However, the ONC final rule for certified health information technology only requires the use of USCDI V3, which contains neither [12].
- Standards can be ambiguous, such that even when healthcare facilities are required to use certain standards, the mapping decisions across institutions may differ widely [13], [14].
- Adding regulatory requirements may have significant financial implications for care facilities. CMS must weigh additional regulations against the cost of implementation.
- Labs may not feel the need to do more than the minimum required to maintain certification and hedging language, such as, “if appropriate”, around flagging requirements and the lack of consistency between inspectors makes checking for consistent and safe flagging a lower priority.
- Regulatory groups are hesitant to alter existing regulatory requirements, even if there is a recognized need, due to concerns about financial burdens. The process to add or change requirements is extensive and involves assessing the cost it would impose on labs.
- There is no effective feedback mechanism to regulatory groups, like CMS, to identify and prioritize CLIA gaps, such as missing abnormal value flags. The full extent of data issues, like missing flags, is neither known nor measured by regulatory groups, and these issues are sometimes viewed as the responsibility of the laboratory director or administrator rather than a CLIA problem.

- The regulatory scope and responsibilities are not clearly defined in a way that addresses whether missing laboratory data, such as flags, fall within the scope of CLIA.
- Adverse events are not investigated systematically or thoroughly, so investigations typically do not go deep enough to identify CLIA gaps. Additionally, adverse events are investigated by hospitals rather than independent groups, which means there is little incentive to examine causal factors that originate outside the hospital and are beyond its control. As a result, CLIA gaps that contribute to adverse events are often not recognized or monitored by any body.

Role in Adverse Event

- Did not perform effective audits capable of discovering that cortisol flagging and flagging ranges were not defined.
- Did not perform effective audits to discover that cortisol reference ranges could be hidden in a notes field that is not visible to clinicians following their normal workflow.

Mental model

- Believed it would be too onerous on the part of the regulatory body and the facility to conduct such audits.
- Believed it is not the responsibility of the regulatory body to uncover every possible deficiency in every test.
- May not believe that these types of laboratory data pose significant risk to patient health and is not worth the cost of extra oversight [3].

Context

- Labs offer thousands of tests. Auditors only have the time to conduct spot checks.
- Minimal processes and authority exist to discover or investigate deviations outside of regulatory requirements.
- Most laboratory regulations focus on the test performance and accuracy rather than the safety implications of the data generated.
- Reports of safety concerns regarding clinical laboratory data from the lab or care facility may not reach regulatory agencies with the authority to amend regulations. Typically, data collection focuses on whether a product meets existing regulatory requirements. Consequently, due to insufficient regulations on the display of laboratory data, issues with how this data is presented are neither collected nor monitored by any regulatory agency.
- Because there is no effective feedback to CMS to identify and prioritize CLIA gaps like missing flags, gaps are not recognized, and when they are recognized, they are easily argued as hospital administrators' responsibility, not CLIA. In other words, the scope and responsibilities of CLIA are not clearly defined in a way that would resolve the question of whether missing laboratory data (like flags) is in scope for CLIA or not. Scope questions increase the difficulty of adding flagging checks to audits—which again decreases the feedback about the true state of the system.

Role in Adverse Event

- Inconsistently issued citations when flags were undefined or not visible.

Mental model

- Believed that the existing regulations are sufficient for consistent interpretation by different inspectors and facilities.
- Believed that regulations must be written in a way that is simultaneously applicable to both computer and paper-based systems.

Context

- Many of the regulations related to laboratory data were created at a time when lab results were not electronically saved, stored, and communicated. Therefore, the regulations do not distinguish between the laboratory data transmitted in a final report and the actual screens that are used by clinicians. Policies that address abnormal result flagging are often vague and are interpreted differently by regulators and health care facilities.
- Clinicians frequently use data summary screens to review lab values. These summary screens often do not display all contextual information about a test, including reference ranges or interpretive guides. Flagging may be the only cognitive aide to help clinicians identify abnormal lab values.

Role in Adverse Event

- Defined ambiguous requirements related to ensuring that laboratory data successfully reaches the “final destination”, resulting in different interpretations and lack of scrutiny of the actual screens used by clinicians.

Mental model

- Believed that the agency regulating laboratories did not have regulatory authority over the receiving EHR system.

Context

- There are different regulatory bodies that regulate healthcare facility EHRs and laboratories.
- Some labs send data to hundreds of different healthcare facilities. It would require significant resources for labs and regulatory agencies to review/audit all possible connections regularly.
- Clinical Laboratory Improvement Amendments (CLIA) issued by the Centers for Medicare and Medicaid Services (CMS) does not have the authority to regulate lab display screens used by clinicians; it only has the ability to regulate the contents of the “official lab report”. Accessing this document usually requires several clicks, and it often contains more detailed information than the clinician may need. Most clinicians do not take the time to open the official lab report page to view test results, particularly if they believe the test result is normal.
- CLIA guidelines were established before EHR systems became widespread. Before EHRs, clinicians received and reviewed the full paper report for each lab test. However, modern EHRs provide summary screens that eliminate the need for clinicians to manually review the full report every time. In other words, the “final destination” of the laboratory data has changed from a report to a summary screen. There are minimal regulations from the ONC or other regulatory agencies that apply to EHR screens.

- Generally, interactions between different groups in the healthcare system are not regulated because it is often unclear who has the jurisdiction to do so effectively. This is particularly true when the interacting entities are regulated by different bodies, such as hospital EHRs, which are partially regulated by the ONC, and laboratory LISs, which must adhere to regulations set by CMS.
- Furthermore, there is limited oversight of how each hospital implements its EHR. The ONC evaluates the base product, but “how a certified EHR is implemented or used at individual care facilities is not [3].” The regulations enforced by the ONC have different goals compared to those of CMS's CLIA. ONC regulations are intended to ensure that certain functionality criteria are met. They do not explicitly evaluate safety. CLIA has a more direct focus on safety.
- The guidelines for reporting laboratory test results were established before EHRs became widespread. Although the guidelines are generally technology-agnostic (e.g., paper versus electronic), the shift to modern technologies like EHRs has introduced new hazards that were not anticipated decades ago. There have been limited policy changes to enhance safety since the existing guidelines, most of which predate modern EHRs, were created. Additionally, some healthcare facilities still use paper records, so regulations must cover all types of facilities.

Initial recommendations

Recommendation: **Regulations addressing hospital EHR systems (such as the ONC certification program) must be updated to require that lab results distinguish between the following cases:**

- **Lab value that has no associated flags and requires manual interpretation**
- **Lab value is within reference range**
- **Lab value is within flagging range**
- **Lab value is in critical range**

Recommendation: **Improve the ability for auditors to identify missing reference ranges and flags quickly.**

Recommendation: **Investigate how many laboratory tests are not interpretable through traditional reference ranges. Tests that are exceptions to the traditional reference range should be identified in a standard way, and other guidelines for flagging behavior should be developed.**

Accreditation Agencies

Description

Accreditation agencies are organizations that provide certifications to care facilities or products to demonstrate that the facility or product meets certain standards. At minimum, these organizations often certify that care facilities meet minimum regulatory requirements. These organizations may also have additional standards that the facilities must meet, though many do not. Insurers and other payors may require a certification from one of these bodies before they will

pay for patient care. Regulatory authorities may also accept certification as proof of meeting regulatory requirements. Accreditation agencies exist throughout the healthcare system. For example, some certify laboratories, some certify EHR technology, others certify care facilities.

Relevant responsibilities

- Determine standards for certification.
- Assess performance.
- Monitor for performance deviations.
- Update standards.

Role in Adverse Event

- The accreditation agencies that certify laboratories did not enforce standards that clearly indicated that lab results like cortisol must be accompanied by appropriate flags.

Mental model

- Believed that care facilities would be able to determine safe flagging practices on their own.
- Believed that standards were crafted strictly enough to enforce appropriate flagging practices.

Context

- Accreditation agencies cover thousands of care facilities across the U.S. Similar to regulatory agencies, it may be difficult for them to craft policy that is flexible enough to cover all cases but strict enough to catch all errors.
- Accreditation agencies may not have the resources to go over every lab test at every facility. Random spot checks may not be enough to catch all safety hazards.
- Accreditation agencies may not look at data points like, “What percentage of test results have flags associated with them?”

Recommendation: **Accreditation agencies should develop quality assessment research alongside safety audits. These may include questions like, “How many tests do not have flagging ranges programmed?” Inquiries like this can help the accreditation agencies identify areas where their current standards are insufficient.**

Systemic Factors and Additional Recommendations

A number of causes involve the LIS/EHR system and will require mitigations. In addition to the cortisol-specific LIS/EHR recommendations identified earlier, solutions should be identified for similar tests that do not have a flagging range defined and are not automatically flagged. During this study, other common PST tests at the facility were reviewed. It was found that 15% of the tests lack a defined flagging range, which is likely to contribute to similar adverse events in the future. Other institutions were examined to determine if similar issues exist elsewhere, and several tests at these institutions were also found to have missing flagging ranges that clinicians rely on. However, the exact proportion remains unknown, as it is not routinely tracked or measured. There is no indication that this 15% figure would differ significantly at other facilities.

Many of the systemic causes behind these missing flagging ranges are shared and not specific to cortisol. It is unclear what proportion of common tests outside of PST experience a similar issue, as this data is not typically monitored.

Often, the lack of a flagging range is because the range depends on patient-specific contextual factors and may not be universal for all patients in all contexts. Solutions to enable flagging or otherwise aid nonspecialists should be considered. For example: for the 15% of tests commonly used by clinicians that do not have a flagging range defined, the provider could be asked to enter a custom flagging range when ordering the test. A provider should not be ordering a test that they cannot interpret, especially for tests that don't have automatic flags. Since providers are already expected to know what values would be abnormal for their patient—especially when the EHR is not programmed with any flagging range—this would not be a significant workload. The benefits of this recommendation include:

- Would enable the results to be flagged, ensuring abnormal results are not missed by the ordering provider or other clinicians in the future.
- Would provide feedback to clinicians to refresh their own mental model about whether flagging is enabled for the test or not (e.g., when a prompt automatically appears asking for a flagging range, it is clear that no pre-defined flagging range exists).
- Would prompt clinicians to update their own mental model about how the test should be interpreted and what an abnormal result would look like.
- Provide a proactive leading indicator of potential tests that are being misunderstood by clinicians because unreasonable ranges are being entered. This could be used to proactively determine if and where additional education is warranted.
- Minimize the various cognitive biases that cause interpretive errors when clinicians receive results first and then try to decide what thresholds to mentally apply to interpret the results. For example, if, as in this case, the physician already has a suspected diagnosis, confirmation bias may lead them to focus more on signals that support that theory. Factors that do not support the theory may need to be even more prominent to overcome this inherent bias.
- Provide a record of individual flagging ranges at the medical center for those tests without a defined flagging range. This could be used to inform future decisions to review or update reference ranges, flagging ranges, and critical values.
- Improve trending capability by enabling flags when these tests are compared over time, especially when many different tests are being trended.

- Reduce the complexity and potential for human error when interpreting a suite of test results that may have mixed flagging capabilities. Otherwise, some values can be flagged and others cannot, which increases interpretive mistakes.
- Reduce the number of potentially abnormal test results that would otherwise be displayed in the same way as a normal result (e.g., black).
- Provide a means to identify tests where it was previously assumed that flagging ranges cannot be defined while in practice there may be a common flagging range used by most physicians.
- Provide additional information for committees responsible for critical values and urgent values to support their decision-making.
- Provide additional contextual information for lab technicians, such as what values the ordering provider is primarily interested in. This information can inform the lab technician's selection of test method. For example, different tests may be used if the ordering provider is primarily interested in a low value.
- Could optionally allow providers to mark the custom range as a critical value in cases where the patient's condition may require the provider to be notified immediately to take action.

Recommendation: Require providers to enter a custom flagging range when ordering tests that are not automatically flagged.

One potential drawback of this recommendation is the possibility that an incorrect flagging range may be entered by the ordering clinician, potentially due to a lack of understanding of the test or the patient-specific factors that affect its interpretation. However, consider the outcome with versus without this recommendation. Current technology, without this recommendation, does not address this concern—it only obscures the error. In the current system, if a clinician lacks the ability to interpret a test and there is no flagging range, they would still overlook the unflagged abnormal result. Providers should not be ordering tests they cannot interpret. Essentially, today's common implementation forces clinicians to create their own undocumented mental flagging range on-the-spot when interpreting a result without a built-in flagging range, which increases the possibility of a clinical error. The proposed recommendation would shift today's error of omission (overlooking an abnormal result that isn't flagged) to an error of commission (potentially entering the wrong flagging range). Errors of commission are widely recognized in the human factors community as easier to detect and prevent compared to errors of omission.

However, this recommendation has not been widely studied to formally identify advantages and disadvantages in practice. Further research is warranted to explore this topic.

However, in addition to immediate fixes to the LIS/EHR that would prevent this specific accident from happening, analysis of the control structure revealed numerous systemic factors that must be addressed to prevent other similar (and more serious) adverse events in the future.

Communication and Coordination

One of the most prominent systemic factors was the need to improve communication and coordination between individuals and departments at the medical center.

For example, there was insufficient coordination and transfer of responsibilities between the different providers that the patient saw between the PST visit and their surgery. As the patient

moved between departments of the hospital, various groups made assumptions about where their responsibilities ended and the next group's responsibilities started. The APP referred the patient to the UCC and believed that they would recognize irregular test results if they appeared. The UCC could view their role as focused on the specific concern identified by the APP and did not carefully scrutinize the lab results for other potential causes. The oncologist was notified that the patient was receiving follow-up care and did not review the lab results critically. However, because the patient was being seen by the UCC (and later an outside cardiologist) the oncologist may have assumed that any unsafe lab results would be viewed by someone else.

Recommendation: **The roles and responsibilities of clinicians who see the patient before surgery must be defined more clearly. If the role of a PST APP is determined to only involve clearing patients to undergo anesthesia for surgery, they should not be responsible for tracking a patient throughout the medical center and external providers. The responsibility for following up and reviewing case notes and test results for patients who are referred out of PST should be given to clinicians who understand the patient's history and may be better able to interpret a patient's individual case. If PST APPs are determined to be responsible for tracking a patient after referring them to other providers, they should be given the time and resources to adequately understand that patient's medical history at the medical center and to review the results from the referral appointments and follow up if necessary.**

Furthermore, there are not many options for clinicians to be alerted to lab test results that do not qualify under the critical values policy but are more severe than a small abnormality. Cortisol, for example, does not usually have a critical value defined because low cortisol will not typically harm patients immediately. However, when values are very low, clinicians should be alerted.

The medical center recently added the option for "urgent values" to be defined. Urgent values are automatically emailed to the ordering physician. However, many clinicians are unaware of this alert option, and very few tests have had urgent values added.

Recommendation: **Define a process to compare urgent values (which trigger an immediate email) with other institutions to identify best practices and potential gaps.**

Another concern raised in the control structure over coordination is the degree to which patient scheduling can be coordinated by the providers. The PST APPs are unable to schedule patients based on contextual factors. For example, the APPs cannot influence the scheduling of patients who need cortisol tests, even though cortisol tests are best performed at specific times of day. Some departments, like Endocrinology, are able to ensure that their patients get tested at the appropriate time of day. However, the APPs have no control over the scheduling.

Recommendation: **Improve the ability for patients to be scheduled based on contextual factors, such as the tests needed.**

Additionally, there was a lack of coordination in terms of monitoring the side effects of treatment. The APP was the first clinician to order tests to check the patient for adrenal

insufficiency. Neither the oncologist nor the endocrinologists appeared to review these tests for the specific results related to adrenal insufficiency. PST, therefore, was effectively put into a position to serve as both a presurgery health check but also as a post-treatment follow-up. In fact, the patient in question reported feeling ill in the weeks leading up to the PST visit but did not see the need to report it to the main oncologist because of an upcoming PST appointment. The patient may believe that any provider at the medical center will be able to understand and diagnose concerns equivalently. The oncologist may have been able to catch the adrenal insufficiency before the PST appointment if there had been a screening available at the medical center or a local lab.

Recommendation: PST test order guides should include details about the specific patient information being sought. For example: Patients who undergo chemo are at a higher risk of adrenal insufficiency which is a risk for surgery success.

Recommendation: Identify whether patients feel comfortable raising symptom concerns to their primary oncologist outside of set appointments and whether it is easy to do so. If patients are not proactively reaching out, identify ways to enable patients to more easily raise concerns.

Recommendation: Educate clinicians on the “urgent value” feature that has been recently implemented. Develop a clear definition of urgent values and evaluate whether cortisol could meet the definition of an urgent value. Request clinicians to review and provide feedback on candidates to be included in the definition of urgent values.

Safety Management System

This adverse event uncovered several concerns (including those known and unknown prior to the event) that had not been addressed by the medical center. There is a common pattern of not investing resources to make improvements until adverse events occur to justify the change. Making changes proactively based on hazard analyses or clinician input before events occur could save time, resources, and greatly improve patient outcomes in the long run.

In this case, the lack of cortisol flags was only raised as an issue after the incident. There is a limited understanding of which group has responsibility to identify a lack of flagging range and provide oversight to ensure it is corrected.

One of the ways that prospective hazard identification could be improved is by ensuring that the EHR is being interpreted correctly by clinicians across the medical center. There were broad misinterpretations of the LIS/EHR by clinicians and leadership. For example, most clinicians did not know what the asterisk symbol signified on the EHR test results page. Clinicians were unable to identify which results have flagging disabled, and they may not always know when to consult a specialist for context-dependent tests like cortisol.

Recommendation: Put in place a means to identify gaps and misunderstandings involving clinicians and their information system. For example,

ensure that symbols are interpreted correctly and corrections are implemented when gaps are found.

Additionally, there were minimal processes in place to proactively identify and improve EHR flagging behavior. Many interviewees drastically underestimated the percentage of tests that had no flags, meaning the reference range was entered only as a note and not directly entered into the reference range data field that would automatically flag results. Processes should be in place to review labs regularly to improve flagging across the system.

Recommendation: Put in place a process to identify the list of labs for which the flagging ranges and reference ranges do not match. For example, the mitigations implemented following this adverse event resulted in a cortisol flagging range that does not match the reference range if the blood is drawn in the afternoon. This list of labs should be periodically audited and reviewed as any discrepancies will cause false positive and false negative flagging that will affect clinical decision-making in the future.

Some individuals in the medical center were aware that cortisol was not a test that would flag. However, there was also a belief that having the reference range in a note would be sufficient to ensure clinicians recognize values that are abnormal. In hindsight, this assumption was not valid. More importantly, there were no efforts to confirm the assumption and provide a leading indicator before the event. There were no practices or processes by any group to confirm that all clinicians were aware of which tests had flags, or aware that some tests did not have any flags. This was particularly problematic for PST nurses who are handling a large patient load and may not encounter unflagged abnormal values frequently.

Recommendations: If comments fields continue to be used to convey reference ranges, then put in place a means to monitor how often the comments field is actually opened by clinicians and how often values that are outside the reference range are overlooked. One solution may be to audit a sample of cases where the lab result was outside the reference range that was only described in a comment field. Use the results of this audit to confirm or reject the decision to rely on comments fields as an effective way to communicate reference ranges.

Recommendation: Update the lab “tracer” process (or a similar process) to include the clinician in the loop. In other words, confirming that information is available in a potentially hidden note is not satisfactory when many clinicians do not actually review the notes. In fact, many clinicians do not recognize when a note exists, since there is no legend to communicate which symbol indicates that a note exists. The tracer process (or another similar process) needs to cover end-to-end transmission of the data to the “final destination”—meaning successful receipt and review by the clinician or ordering provider on their screen.

Recommendation: For tests that have a reference range but no flagging range, define a process to compare the reference and flagging ranges with other institutions and revisit the way the reference range and flagging range is entered into the LIS/CIS.

Recommendation: Complaints and RISQ reports are not an effective way to ensure all abnormal flags and urgent values are appropriately defined. Establish review processes to ensure these values are safe and appropriate, similar to the critical values committee with literature reviews, comparisons to peers, and periodic reviews as appropriate. The process should not rely on one person making the correct flagging range decision years ago.

Culture

The culture surrounding nurse and APP relationships played a role in the adverse event. There is a shortage of nursing staff at the medical center and in the United States broadly. Assistance is spread thin. However, there is also a culture at the medical center wherein nurses are not expected to assist APPs. As APPs take on jobs that used to be solely done by physicians, their need for assistance will continue to increase.

Recommendation: Provide designated nurses to support PST APPs, including minimizing APP interruptions, assisting in reviews of patient lab results, and managing multiple information systems as required.

Recommendation: Determine whether all PST APPs agree that they have a practical quiet space in which to review lab results between patients. If not, then determine why and address the issue. Evaluate if this space is actually used. If not, then determine why and address the issue.

Another cultural factor that impacted this adverse event was the lack of information sharing between clinicians and laboratorians. Unlike other diagnostic fields, laboratorians rarely receive information about what the clinician is testing for. In some cases, this limits their ability to choose the optimal test. Therefore, the laboratorian's expertise is underutilized. There is a belief among some clinicians that the laboratorians would not be able to understand any clinical context and therefore, additional information would be a waste of time. However, simple changes, like enabling the laboratorians to see what time the sample was collected, can improve their ability interpret results.

Recommendation: Provide a means for the lab and clinicians to easily know what time the blood sample was drawn for diurnal tests like cortisol. It should be prominent, not hidden.

Changes and Dynamics

Adverse events often follow system changes. In this case, one of the notable changes was the addition of the cortisol test to the PST order set. The facility initiated efforts to train or educate PST APPs on the value of cortisol testing for patients who had received immunotherapy. However, there

was no proactive analysis of how mistakes could be made, how the cortisol test should be viewed by the PST APPs (e.g., which screen), and whether changes were required to make the test results easily interpretable by clinicians who were not endocrinology specialists. In this case, the system that worked for reviewing cortisol in endocrinology settings did not work when APPs who cover broad swaths of patients were reviewing the same information.

Recommendation: **When tests are added to a test menu, ensure there is a protocol for reviewing how the test is displayed and if the flags are sufficient for the new users.**

Recommendation: **When new flags are entered for a test (or not entered), a service other than the lab group should review the decision and reach agreement before it is approved.**

Economic Pressure

Pressure to see more patients to ensure financial viability is a factor that applies to all kinds of hospitals in the United States. This medical center had been steadily growing its patient population. This growth increases the amount of patients per clinician. Hospitals need to stay financially solvent, and patients benefit from access to high quality care. However, growth requires additional care to prevent oversights and errors.

One of the economic concerns involved the scheduling of patients consecutively, without a break between appointments. This resulted in an inadequate amount time explicitly set aside for the clinician to review lab tests. This allowed more patients to be seen, but the APPs did not have enough time to review patient cases before or after patient visits. Unlike other APPs at the medical center, the PST APPs do not have an administrative day to work on case management.

Several productivity metrics were recently introduced, which has created a perception of time pressures, as most of the metrics can be improved by sacrificing quality in favor of quantity. There are important trade-offs between productivity and safety that are not reflected in the metrics. Workers tend to prioritize the aspects of their work that are measured. When productivity metrics are emphasized, people will tend to prioritize the productivity goals against which their performance is assessed, which can unintentionally compromise safety.

Recommendation: **Correct the imbalance of metrics focused on productivity over safety. Measures should give credit for extra time that may be warranted to prevent a costly health or safety issue in the future. Taking a few extra minutes to carefully review lab results or to make use of a designated quiet space for reviewing a medical record should not count against a provider.**

Metrics should never be adjusted to count against a provider for reporting adverse events or near misses. These lagging safety metrics can be problematic because pressures to minimize them will result in reduced reporting of mistakes and adverse events. Reduced reporting prevents an organization from learning and identifying sources of risk, thereby impeding safety.

Pressures to increase efficiency may also have contributed to PST APPs functioning both as supervisors and practicing clinicians simultaneously. During the patient's PST visit, the APP was

responsible not only for patient care, but also for doing administrative duties for the PST unit. Despite the extra workload, the APP had no additional time scheduled for administrative duties. Therefore, the amount of time and attention the APP had for patients was diminished.

Recommendation: **On days when APPs are assigned additional administrative duties, their patient load must decrease accordingly.**

Economic pressures can also contribute to the lack of a sufficient prospective safety management system. It can be difficult for healthcare leaders at this medical center, and other care facilities, to justify spending resources on problems they believe may not exist. However, responding to accidents on an ad hoc basis is costly. Some recommendations may seem resource-intensive until the costs of inaction, including future adverse events, are taken into account.

Regulatory Ecosystem

There are gaps in the regulatory ecosystem that allow for laboratory results to be presented to clinicians in misleading ways. Many of the gaps stem from the lack of regulations that cover data as it transitions from one authority's jurisdiction to another. For example, CMS through CLIA (or equivalent state agencies) has authority over clinical laboratories, including quality management and what data needs to be transmitted to the ordering clinician. However, once that data is sent to the care facility, the way that data is used, displayed, and aggregated is not well controlled under any existing regulatory framework.

Recommendation: **Ensure that the displays of clinical laboratory data to clinicians are evaluated for safety before being certified for programs, such as the CMS Promoting Interoperability program, or another similar program.**

Additionally, because clinical laboratory interpretation is highly context-dependent, there are few regulatory controls on how laboratory data should be presented. Care facilities and labs are required to set their own critical values policy and identify what tests should be flagged and at what thresholds. Without any guidance or regulatory oversight, care facilities may rely on adverse events to identify when their individual policies may be insufficient.

Recommendation: **Set up an industry-wide mechanism to compare practices across institutions, such as flagging ranges and defining separate cortisol a.m./p.m. tests. Make this information available to peer institutions as well as to regulators to understand the state of the healthcare system and the prevalence of various challenges.**

Additional Recommendations

Finally, the following recommendations are proposed to ensure that similar adverse events do not continue to occur:

- Proactive hazard analyses must be conducted on systems like EHRs that physicians rely on to carry out safety-critical duties, such as monitoring patient progress, coordinating care, and managing complex treatments. A proactive hazard analysis on the overall U.S. laboratory data system uncovered many of the same gaps and issues that were uncovered

through this adverse event investigation, including gaps in the regulatory infrastructure, missing coordination between the laboratory and clinicians, missing or misleading data presented to clinicians, and economic pressures influencing EHR functionality choices. Hazard analyses should be performed by the company purchasing a system (such as an EHR) as well as by the vendor providing the system to anticipate and prevent these safety issues that negatively impact patient care.

- If there are recommendations that cannot be achieved due to technical limitations, then document the recommendations, requests, and limitations in a report. Provide the report to the technical vendor for feedback and future improvements. Provide the report to state and federal regulators as feedback about technology gaps that are contributing to adverse events. Such reports and feedback are rarely submitted by medical facilities, which is one of the reasons the technical gaps still exist and have not been fixed.
- Feedback is currently provided to hospital administrators regarding concerns about time pressures, performance, and quality vs. quantity, but there is a perception that the feedback is not acted upon. Strengthen the feedback by identifying ways to connect the feedback to past, present, and future loss events. For example, ensure that future adverse event investigations are capable of confirming or rejecting which of these previously reported systemic factors contributed. Aggregate these findings to provide specific metrics for leadership, such as: __% of adverse events in the last year have been linked to perceived pressures to prioritize quantity over quality. The administration should clearly communicate which actions, if any, have been taken in response to the feedback received. Without clear communication about the specific actions taken in response, the feedback will continue to seem like it goes into a “black hole” from a practitioner’s perspective. This will hinder future feedback and prevent the administration from having an accurate understanding of the medical center’s true state, impacting their decision-making.

6. Conclusions

This CAST report has demonstrated how using a system-theoretic approach to accident investigation can identify systemic factors and recommendations that can prevent this adverse event and prevent similar adverse events in the future.

By modeling the interactions between controllers in the system, CAST identifies the unsafe interactions in the system rather than blaming individuals or components. In other words, CAST identifies flaws in the overall structure and design of the healthcare system rather than simply identifying individuals or components as the root cause. The recommendations generated by a CAST report can benefit the individual facility in question, but the results are also applicable to many healthcare facilities across the United States. Moreover, the results of this study go beyond individual healthcare facilities to identify causal factors related to regulatory authorities at the national, state, and local level.

While this adverse event was fortunately identified by an external medical specialist before more severe consequences could occur, medical centers should have robust frameworks in place to detect and prevent such issues internally. The CAST analysis revealed several risks and vulnerabilities that, if not addressed, would lead to more severe adverse events in the future.

7. Reference List

- [1] N. G. Leveson, “CAST Handbook: How to Learn More from Incidents and Accidents,” 2019. [Online]. Available: https://psas.scripts.mit.edu/home/get_file4.php?name=CAST_handbook.pdf
- [2] K. A. Sikaris, “Enhancing the Clinical Value of Medical Laboratory Testing,” *Clin Biochem Rev*, vol. 38, no. 3, pp. 107–114, Nov. 2017.
- [3] N. Leveson, J. Thomas, P. Harrington, R. Rodrigo, S. Powell, and A. Keller, “System Safety within Laboratory Data Exchanges Report,” Sep. 2023. [Online]. Available: <http://psas.scripts.mit.edu/home/wp-content/uploads/2023/10/System-Safety-within-Laboratory-Data-Exchanges-Report.pdf>
- [4] N. Leveson, J. Thomas, P. Harrington, R. Rodrigo, S. Powell, and A. Keller, “Addendum to the System Safety within Laboratory Data Exchanges Report: Over-the-Counter and Point-of-Care Testing,” 2024. [Online]. Available: <http://psas.scripts.mit.edu/home/wp-content/uploads/2023/10/System-Safety-within-Laboratory-Data-Exchanges-Report.pdf>
- [5] N. Leveson, “CAST Handbook; A ‘Systems Thinking’ Approach to the Investigation of Healthcare Adverse Events,” Jan. 2024. [Online]. Available: http://psas.scripts.mit.edu/home/get_file10.php?name=CAST_Handbook_Healthcare.pdf
- [6] “Pulmonary embolism - Symptoms and causes,” Mayo Clinic. Accessed: Jul. 06, 2024. [Online]. Available: <https://www.mayoclinic.org/diseases-conditions/pulmonary-embolism/symptoms-causes/syc-20354647>
- [7] “Lung Nodules | Is a Lung Nodule Cancer?” Accessed: Jul. 06, 2024. [Online]. Available: <https://www.cancer.org/cancer/types/lung-cancer/detection-diagnosis-staging/lung-nodules.html>
- [8] C. S. White and G. D. Rubin, “The Puzzle of the Perifissural Nodule,” *Radiology: Cardiothoracic Imaging*, vol. 2, no. 4, p. e200409, Aug. 2020, doi: 10.1148/ryct.2020200409.
- [9] “Ejection fraction: An important heart test,” Mayo Clinic. Accessed: Jul. 06, 2024. [Online]. Available: <https://www.mayoclinic.org/tests-procedures/ekg/expert-answers/ejection-fraction/faq-20058286>
- [10] “Postural tachycardia syndrome (PoTS),” nhs.uk. Accessed: Jul. 09, 2024. [Online]. Available: <https://www.nhs.uk/conditions/postural-tachycardia-syndrome/>
- [11] “USCDI Laboratory.” Accessed: Jul. 23, 2024. [Online]. Available: <http://www.healthit.gov/isp/uscdi-data-class/laboratory>
- [12] Health and Human Services Department, “21st Century Cures Act: Interoperability, Information Blocking, and the ONC Health IT Certification Program,” Federal Register. Accessed: Nov. 09, 2023. [Online]. Available: <https://www.federalregister.gov/documents/2020/08/04/C2-2020-07419/21st-century-cures-act-interoperability-information-blocking-and-the-onc-health-it-certification>
- [13] S. H. Brown *et al.*, “One-Way and Round-Trip Analysis Demonstrates Surprising Limitations of Standards-Based Terminology Maps,” *AMIA Annu Symp Proc*, vol. 2019, pp. 258–266, 2019.
- [14] A. B. Carter, M. E. De Baca, H. S. Luu, W. S. Campbell, and M. N. Stram, “Use of LOINC for interoperability between organisations poses a risk to safety,” *The Lancet Digital Health*, vol. 2, no. 11, p. e569, Nov. 2020, doi: 10.1016/S2589-7500(20)30244-2.
- [15] L. Thau, J. Gandhi, and S. Sharma, “Physiology, Cortisol,” in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2024. Accessed: Jun. 27, 2024. [Online]. Available: <http://www.ncbi.nlm.nih.gov/books/NBK538239/>
- [16] T. Woodcock *et al.*, “Guidelines for the management of glucocorticoids during the peri-operative period for patients with adrenal insufficiency: Guidelines from the Association of

- Anaesthetists, the Royal College of Physicians and the Society for Endocrinology UK,” *Anaesthesia*, vol. 75, no. 5, pp. 654–663, May 2020, doi: 10.1111/anae.14963.
- [17] “Immune Checkpoint Inhibitors - NCI.” Accessed: Jul. 08, 2024. [Online]. Available: <https://www.cancer.gov/about-cancer/treatment/types/immunotherapy/checkpoint-inhibitors>
- [18] K. Manaka *et al.*, “Immune checkpoint inhibitor combination therapies very frequently induce secondary adrenal insufficiency,” *Sci Rep*, vol. 11, no. 1, p. 11617, Jun. 2021, doi: 10.1038/s41598-021-91032-6.
- [19] K. Cui, Z. Wang, Q. Zhang, and X. Zhang, “Immune checkpoint inhibitors and adrenal insufficiency: a large-sample case series study,” *Ann Transl Med*, vol. 10, no. 5, p. 251, Mar. 2022, doi: 10.21037/atm-21-7006.
- [20] Y. Ozarda, “Reference intervals: current status, recent developments and future considerations,” *Biochem Med (Zagreb)*, vol. 26, no. 1, pp. 5–11, Feb. 2016, doi: 10.11613/BM.2016.001.
- [21] “CLIA Verification of Performance Specifications”.
- [22] *Clinical Laboratory Improvement Amendments (42 USC 263a)*. 1988. Accessed: Aug. 21, 2024. [Online]. Available: <https://www.govinfo.gov/content/pkg/USCODE-2011-title42/pdf/USCODE-2011-title42-chap6A-subchapII-partF-subpart2-sec263a.pdf>
- [23] New York State Department of Health Wadsworth Center Clinical Laboratory Evaluation Program, “Clinical Laboratory Standards of Practice General Systems Standards.” Aug. 01, 2020. Accessed: Aug. 21, 2024. [Online]. Available: https://www.wadsworth.org/sites/default/files/WebDoc/General%20Systems%20Standards%20-%20Effective%20August%202020_1_0.pdf
- [24] D. F. Sittig, J. S. Ash, and H. Singh, “The SAFER guides: empowering organizations to improve the safety and effectiveness of electronic health records,” *Am J Manag Care*, vol. 20, no. 5, pp. 418–423, May 2014.

Acronyms

APP:	Advanced Practice Provider
CAST:	Causal Analysis based on Systems Theory
CIS:	Clinical information system. Often used interchangeably with EHR
CMS:	Centers for Medicare and Medicaid Services
ED:	Emergency Department
EHR:	Electronic Health Record
HIPAA:	Health Insurance Portability and Accountability Act
ICI:	Immune Checkpoint Inhibitors
IT:	Information Technology
LIS:	Laboratory Information System
LOINC:	Logical Observation Identifiers Names and Codes
MD:	Medical Doctor
ONC	Office of the National Coordinator for health information technology
PST:	Presurgical Testing
RISQ:	Report Incidents of Safety and Quality
SMS:	Safety Management System
STPA:	System-theoretic Process Analysis
UCC:	Urgent Care Center
USCDI:	United States Core Data for Interoperability

Appendix A: Laboratory Testing Background Information

Cortisol

Cortisol is a hormone that plays a role in stress, metabolism, inflammation, and immunity functions. Cortisol affects systems throughout the body including the immune, nervous, and cardiovascular [15]. Testing for cortisol before surgery is meant to ensure that patients who are taking immunosuppressants have a hormone system that is sufficient to undergo the stress of surgery [16]. High cortisol levels can be caused by tumors or by use of steroids. Low cortisol levels can be caused by certain cancer treatments and autoimmune responses among others, and can cause fatigue, low blood pressure, and other symptoms [15]. Low levels are critical to diagnose and treat before surgery to “avoid adrenal crisis” [15].

Immune Checkpoint Inhibitors (ICI)

Immune Checkpoint Inhibitors are used in treatment against cancer. The human immune system employs immune checkpoints to ensure that the immune system does not destroy healthy cells. However, if immune checkpoints are present in cancer cells they may prevent cancer treatments from killing the cancerous cells. ICIs allow T-cells to destroy cancer cells by blocking the immune checkpoints [17]. However, using ICIs can have severe side effects. One of the risks is that patients can develop “Secondary Adrenal Insufficiency” [18] wherein the adrenal glands do not produce sufficient hormones, including cortisol. Some estimates suggest that 1-5% of patients on ICIs develop adrenal insufficiency [19]. Early warning symptoms of adrenal insufficiency include headache and fatigue [18], [19]. Although adrenal insufficiency is typically not immediately life threatening, prolonged adrenal insufficiency can result in significant harm and can be life threatening.

Display of Laboratory Data

Figure 12 shows a mockup representation of one of the screens used by clinicians to view laboratory results in the electronic health record. This is one of the “final destinations” of the laboratory data.

	[Day 1]	[Day 19]	[Day 36]
eGFR (CKD-EPI 2021)	* 75	* 75	* 75
Anion Gap (Calc)	10		8
Z-score			
Insulin-Like Growth Factor-1, LC-M...			
Cortisol Level	* 0.5		* 0.9
Estradiol			
Growth Hormone			
Hemoglobin A1c.	* 5.5		* 5.5
Thyroxine (T4), Free	0.81	0.79	0.96
Thyroid Stimulating Horm...	* 1.72	* 1.34	* 2.64

Figure 12: Representation of how the laboratory results were presented in the EHR

Although the cortisol level is accurate, hidden from view are the flags, reference range, and notes that provide important information about acceptable and abnormal levels of cortisol. There is no indication that the cortisol value, shown in black, is any different from the other values that are shown in black to indicate they are normal and within the reference range.

	[Day 1]	[Day 19]	[Day 36]
eGFR (CKD-EPI 2021)	* 75	* 75	* 75
Anion Gap (Calc)	10		8
Z-score			
Insulin-Like Growth Factor-1, LC-M...			
Cortisol Level	* ↓ 0.5		* ↓ 0.9
Estradiol			
Growth Hormone			
Hemoglobin A1c.	* 5.5		* 5.5
Thyroxine (T4), Free	0.81	0.79	0.96
Thyroid Stimulating Horm...	* 1.72	* 1.34	* 2.64

Figure 13: Expected and typical flagging of abnormal values in the EHR

Figure 13 shows the EHR representation of abnormal values with flagging, which is typical for most test results. Most clinicians, including those involved in the adverse event, believed this is what the display would look like if the cortisol values were too low. After the adverse event, it was discovered that cortisol tests are one of the few that did not have flagging capability enabled in the EHR and LIS.

Figure 14 shows an alternate screen that is also used by clinicians to view lab results.

Aspartate Aminotransferase (AST), Plasma	64	↑	[<=37 U/L]	Final
Alanine Aminotransferase (ALT), Plasma	89	↑	[<=55 U/L]	Final
Recommended healthy ALT level				
Males: 29-33 U/L				
Females: 19-25 U/L				
American College of Gastroenterology (ACG) Clinical Guideline: Evaluation of abnormal Liver Chemistries. 2017				
Alkaline Phosphatase (ALK), Plasma	65		[<=130 U/L]	Final
Bilirubin, Total Plasma	0.3		[<=1.2 mg/dL]	Final
eGFR (CKD-EPI 2021)	75		[>=60 mL/min/1.73 m2]	Final
The CKD-EPI (2021) is used to calculate the estimated GFR and is not adjusted to extreme body surface area. The CKD-EPRI (2021) equation has not been validated for children less than 18 years, pregnant women, or the elderly. It has been developed in 2021 by the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) and is recommended by the National Kidney Foundation and American Society of Nephrology Task Force Reassessing the Inclusion of Race in Diagnosing Kidney Disease. Inker LA, Eneanya ND et al. New Creatinine- and Cystatin C-Based Equations to Estimate GFR without Race. N Engl J Med 2021;385:1737-49.				
Hemoglobin A1c.	5.5		[4.0-6.0 %]	Final
Please note change in methodology as of 12/02/2021. The Hemoglobin A1c is performed on the Abbott Architect c8000. This assay measures the glycated fraction of total hemoglobin. It may not accurately reflect the glycemic control in patients with high levels of fetal hemoglobin (HgbF), recent transfusion, hemoglobinopathies or other conditions causing abnormal red blood cell turnover.				
Cortisol Level	0.5		[mcg/dL]	Final
Reference Range:				
A.M. 5.0-25.0				
P.M. 3.0-12.0				
Thyroid Stimulating Hormone (TSH)	2.30		[0.60-4.80 mIU/L]	Final
Includes reflex to Free Thyroxine for Thyroid Stimulating Hormone outside the reference range				
pH, Venous	6.950	↓↓	[7.320-7.430]	Final
Outside Reportable Range				
PCO2, Venous	56.0	↑	[42.0-51.0 mm Hg]	Final
PO2, Venous	53.0	↑	[<=50.0 mm Hg]	Final
Please note change in reference range 2/15/2024.				
Oxyhemoglobin, Venous	58.8	↓	[94.0-97.0 %]	Final
Please note change in reference range 2/15/2024.				
Chloride, Whole Blood	107		[98-109 mEq/L]	Final
O2 Sat., Venous	59.5		[50.0-85.0 %]	Final

Figure 14: Representation of the alternate view also used by clinicians

The screen in Figure 14 automatically displays any notes beside the name of the lab test. There is a column for the test result, red flags, and the flagging range (which typically matches the reference range). A single red arrow indicates the lab value is flagged (typically indicates it is abnormal and outside the reference range). A double red arrow indicates the lab value exceeds a critical value that requires an immediate phone call from the lab.

Another screen contains a final lab report with detailed information, including all of the laboratory data. However, this report is cumbersome to access and review, especially when multiple tests have been ordered. It is not typically used by clinicians unless there is a particular doubt or concern about the test. The final lab report screen was not reviewed by any of the clinicians or medical doctors (MDs) involved in the adverse event, which is consistent with normal practice.

Laboratory Terms

Table 2 summarizes laboratory terms that are relevant to understanding the adverse event.

Table 2: Summary of Laboratory Terms

Term	Relevance to adverse event
Reference Range: Defines the interval of normal values for 95% of the patient population. Values outside the reference range are considered abnormal.	The reference range for cortisol depends on time of day and was hidden in a note field that is not visible on screens used by clinicians.
Flagging Range: Defines the range that will trigger automatic flagging of the lab values. The flagging range thresholds usually, but not always, match the reference range thresholds.	The flagging range for cortisol was not defined, meaning that flags would never be displayed for abnormal cortisol values.
Flagged Values: Values in the flagging range are automatically flagged, meaning they are displayed differently than normal values. Flagging typically involves displaying the value in red with an accompanying red arrow. Flagging is intended to ensure any abnormal values are prominent and difficult to overlook among potentially many other normal test results. Usually, any flagged values are abnormal values, but not always.	The cortisol values were not flagged because the flagging range was not defined.
Non-flagged Values: Values that do not exceed the flagging range are usually displayed in regular black font with normal background. Usually, non-flagged values are normal values that are within the reference range.	The patient’s abnormal cortisol values were displayed in the same way as normal values—in regular black font, normal background, and no flags.
Urgent Values: Urgent values are those that require an email to be sent to the ordering provider because the values significantly exceed what is normal.	The patient’s extremely low cortisol values did not trigger an email because cortisol had no defined urgent value.
Critical Values: Critical values are those that require an immediate phone call from the lab to the ordering provider because the values are life threatening without immediate treatment and for which an immediate treatment is possible.	The patient’s extremely low cortisol values did not trigger a phone call because cortisol had no defined critical value

Reference Range

The International Federation of Clinical Chemistry sets standards for determining accurate reference ranges [20]. Reference ranges are defined as “the range of values expected for a population of individuals, e.g., 95 percent of individuals that are presumed to be healthy (or

normal)” [20]. Reference ranges can be those provided by the test device manufacturer (as long as it reflects the patient population). [21]

The reference range is standardized, reviewed, and regulated for each test. The cortisol reference range is one of the few presurgical tests that is diurnal, meaning the reference range depends on the time of day of the blood sample. At the time of the event, the cortisol reference range was defined as 5-25 mcg/dL for a typical morning blood sample or 3-12 mcg/dL for an afternoon blood sample. Many CIS/LIS systems do not natively support diurnal reference ranges for a single test order, so workarounds must be used, like adding a notes field to explain the reference range or defining multiple tests (e.g., “Cortisol a.m.” vs. “Cortisol p.m.”).

At this facility, a single “Cortisol Random” test order was defined for all cortisol tests, regardless of the time of day. The diurnal reference ranges were explained in a notes field, which was a decision made 15 years ago when cortisol testing was first introduced at this facility. The standard test results page used by clinicians in this case displays the test values without automatically showing the reference ranges or the notes field. Instead, an asterisk appears next to the cortisol value. There is no legend to explain what an asterisk means, and most clinicians did not know what that asterisk meant. The asterisk appears when a hidden note is available for further viewing if the field is selected. Otherwise, the field appears the same as any other normal test (same color, no flags). The facility is subject to state regulations that require the lab to transmit a reference range along with test results, if applicable. However, there is no requirement for the reference range to be displayed together with the test results, to be presented to clinicians in a uniform way for each test, or to prevent reference ranges from being hidden from view on the screens that clinicians use. Reference ranges were sometimes presented in brackets alongside the test result (e.g., “0.5 [5.0 – 25.0 mcg/dL]”), sometimes in a list as notes (e.g., “a.m. 5.0-25.0 p.m. 3.0-12.0”), sometimes embedded in sentences as notes (e.g., “0.5 [5.0 – 25.0 mcg/dL]. The reference interval provided above is appropriate for a.m. cortisol levels. Expected p.m. cortisol levels range from 5.0 to 12.0 mcg/dL. Note: Reference Range as of 3/20/2024”), and sometimes the reference range is not displayed when viewing the results screen (e.g., “0.5”).

Flagging range

In most laboratory tests, the flagging range is the range that is programmed into an Electronic Health Record (EHR) and Laboratory Information System (LIS) to automatically flag the result when it is outside the range. A flagged value typically appears in red with an up or down arrow to indicate high or low. This is a common EHR/LIS functionality that is intended to make abnormal results more prominent and ensure they are noticed by clinicians and laboratorians scanning the electronic test results for an abnormal result.

The flagging range is usually defined to match the reference range, but this is not guaranteed and is not always the case. In some cases, such as with diurnal reference ranges, the reference range and the flagging range are not the same. In this adverse event, a random cortisol test in the test menu had two distinct (a.m./p.m.) reference ranges. However, no flagging range was programmed into the EHR or LIS because diurnal flagging ranges are not supported. The EHR and LIS were unable to be programmed to flag for two different flagging ranges for a single test. Therefore, no flagging or red text would appear even when those test results were abnormal. In

other words, an undefined flagging range will cause the test result to always be presented the same way as a normal test result.

Table 3 summarizes the differences between the reference range and the flagging range.

Table 3: Reference Range Versus Flagging range

		Reference Range	
		Defined	Undefined
Flagging Range	Defined	<p>About 85% of the reviewed PST tests fall into this category. Any values in the flagging range will be flagged red.</p> <p>The flagging range thresholds usually will match the reference range thresholds, but not always. Following the adverse event, cortisol tests were moved to this category. At the time of this report, the flagging range does not match the reference range for p.m. cortisol tests. Instead, the flagging range matches the a.m. reference range for cortisol tests.</p>	<p>Values outside the flagging range will be flagged, even though there is no standardized reference range. This category is possible but typically not intentional and likely indicates an error in either the reference range or the flagging range. It is unknown whether these cases exist at the medical center as there is no process to systematically catch all of these cases across the entire test menu.</p>
	Undefined	<p>Any values outside the reference range will not be flagged. Instead, the value is displayed in the same way as a typical result that is verified to be normal and within the reference range (black) even if it is outside the reference range. About 7% of the reviewed PST tests fall into this category. During the adverse event, cortisol tests were in this category.</p>	<p>Values must be interpreted by the clinician as there is no reference range and no flagging range to flag abnormal values. Instead, the value will appear in the same way as a typical result that is verified to be normal and within the reference range (black). About 7% of the reviewed PST tests fall into this category.</p>

After the adverse event, a cortisol flagging range was created to be consistent with the cortisol a.m. reference range, which is 5-25 mcg/dL. This flagging range ensures that extremely low levels, like the one in this adverse event, will always be flagged. However, if the blood sample is taken during a p.m. visit, then the flagging range will not match the reference range. Exceptions like this

make false positive and false negative flags possible. For example, a p.m. test value can be outside the reference range but not outside the defined flagging range and therefore, not flagged.

After the adverse event, the decision to change all cortisol tests to adopt an a.m. flagging range was discussed with clinicians, including the possibility of false positives and false negatives. The decision was approved, as the perceived benefits were deemed to outweigh the risks of false positive and false negative flags.

Another challenge in defining flagging ranges arises when tests are used to monitor how drugs are affecting the patient, such as in anticoagulation therapy. The flagging range may actually be abnormal for healthy individuals, but it reflects the expected range for the amount of drug the patient is taking. A flagged result could indicate that the patient is receiving too little or too much medication. These ranges are called “therapeutic ranges”, and they do not fit the definition of a reference range for normal healthy individuals.

In some instances, the most appropriate range may depend on the patient’s disease and the goals of the physician. Most information systems do not currently have the ability to discern the appropriate range based on disease or other contextual factors. The end goal of the ordering clinician is also not always clear or documented, making it difficult to determine what flagging range is most appropriate. In these cases, it is common practice to leave the flagging range undefined and use text notes to help guide the clinician in interpreting the test result. However, it is not always clear to clinicians when the flagging range is undefined or when hidden text notes exist.

Critical Value Range

The critical value range defines what test results are severe enough to require an immediate phone call to the provider. Many tests do not have critical values defined because there are risks in requiring too many unnecessary phone calls. Numerous and frequent healthcare provider interruptions can increase human error, reduce working memory, and create alarm overload/alarm fatigue.

This facility uses Lundberg’s definition of a critical value as representing “a pathophysiologic state at such variance with normal as to be life-threatening unless something is done promptly and for which some corrective action can be taken”. All critical values are approved by the facility’s medical board. Lab values that are in the critical range must be “reported directly” to the ordering clinician within 60 minutes of the results being in the LIS. If the ordering clinician is not available, then another appropriate clinician must be notified. Notification and acknowledgment can be verbal or digital.

The critical values policy states that “result notification will not be made if critical results have been previously notified twice consecutively within the past 24 hours”.

Cortisol was reviewed by the facility and found not to meet the definition of a critical value. Critical values for cortisol have not been defined before or after the adverse event.

Regulatory Requirements

The Clinical Laboratory Improvement Amendments (CLIA) [22] were established by Congress in 1988 to ensure the accuracy, reliability, and timeliness of laboratory testing across the United States. These regulations apply to all clinical laboratory testing performed on humans, with the goal of improving the quality of diagnostic testing and, consequently, patient care. CLIA sets forth standards for laboratory operations, including required data that must be included in a test report, processes to periodically conduct reviews to identify gaps, and requirements to address deficiencies. Laboratories must obtain CLIA certification to legally perform diagnostic tests, and they are regularly inspected to ensure compliance with these standards.

However, as technology has advanced, particularly with the widespread adoption of Electronic Health Records (EHRs), there have been concerns that CLIA regulations, originally designed for paper-based systems, may not adequately address the complexities of modern digital health environments. These gaps in regulation contributed to some of the systemic factors identified during the investigation, particularly in relation to the display and interpretation of laboratory results.

Excerpts from CLIA that are most relevant to the adverse event are shown in Table 4.

Table 4: Relevant excerpts from CLIA [22]

<p>§ 493.1241 Standard: Test request.</p> <p>[...]</p> <p>(c) The laboratory must ensure the test requisition solicits the following information:</p> <p>[...]</p> <p>(8) Any additional information relevant and necessary for a specific test to ensure accurate and timely testing and reporting of results, including interpretation, if applicable.</p>
<p>§ 493.1291 Standard: Test report.</p> <p>[...]</p> <p>(c) The test report must indicate the following:</p> <p>[...]</p> <p>(6) The test result and, if applicable, the units of measurement or interpretation, or both.</p> <p>[...]</p> <p>(d) Pertinent “reference intervals” or “normal” values, as determined by the laboratory performing the tests, must be available to the authorized person who ordered the tests and, if applicable, the individual responsible for using the test results.</p>

**§ 493.1407 Standard; Laboratory director responsibilities, and
§ 493.1445 Standard; Laboratory director responsibilities**

The laboratory director is responsible for the overall operation and administration of the laboratory, including the employment of personnel who are competent to perform test procedures, and record and report test results promptly, accurately, and proficiently, and for assuring compliance with the applicable regulations.

[...]

(e) The laboratory director must—

(4) Ensure that the laboratory is enrolled in an HHS-approved proficiency testing program for the testing performed and that—

[...]

(iii) All proficiency testing reports received are reviewed by the appropriate staff to evaluate the laboratory's performance and to identify any problems that require corrective action; and

(iv) An approved corrective action plan is followed when any proficiency testing results are found to be unacceptable or unsatisfactory;

[...]

(8) Ensure that reports of test results include pertinent information required for interpretation;

**§ 493.1419 Standard; Clinical consultant responsibilities, and
§ 493.1457 Standard; Clinical consultant responsibilities**

The clinical consultant provides consultation regarding the appropriateness of the testing ordered and interpretation of test results. The clinical consultant must—

[...]

(c) Ensure that reports of test results include pertinent information required for specific patient interpretation; and

(d) Ensure that consultation is available and communicated to the laboratory's clients on matters related to the quality of the test results reported and their interpretation concerning specific patient conditions.

§ 493.1239 Standard: General laboratory systems quality assessment

(a) The laboratory must establish and follow written policies and procedures for an ongoing mechanism to monitor, assess, and, when indicated, correct problems identified in the general laboratory systems requirements specified at §§ 493.1231 through 493.1236.

(b) The general laboratory systems quality assessment must include a review of the effectiveness of corrective actions taken to resolve problems, revision of policies and procedures necessary to prevent recurrence of problems, and

discussion of general laboratory systems quality assessment reviews with appropriate staff

Experts who were consulted during this investigation interpreted the CLIA language in different ways. For example, § 493.1291 requires that the “test report” must include an “interpretation”, but only “if applicable”. “Test report” is not defined, but it is widely understood by laboratory professionals to refer to the full set of data that is entered into the LIS. Clinical experts and some regulatory experts understood “test report” to mean the information that is actually presented to clinicians at the final destination. “Interpretation” is not defined, but most experts indicated that for quantitative tests, a normal/abnormal flag such as “H” or “L” would typically meet the requirement. No guidance is provided to determine when such an interpretation is applicable and must be included. Nearly all of the clinicians consulted during this investigation incorrectly believed that the cortisol test results would include the flags used to support clinical interpretation. Meanwhile, laboratory experts believed that a text note with sentences describing the typical cortisol reference range was sufficient to provide the test result “interpretation”, even though such notes can be hidden on the screens used by clinicians. Some CLIA experts said they would have cited the medical center for the missing cortisol flags while other experts said they would not.

More important than the ambiguity in the CLIA requirements are the gaps where no requirement exists. For example, there is no specific or general requirement that would compel a facility to ensure that results that are capable of flagging (e.g., with a positive/negative interpretation) are clearly distinguished from those that are not capable of flagging.

One of the most significant gaps in CLIA regulations is the lack of coverage for the end-to-end transmission of laboratory data to clinicians. CLIA requirements pertain to the "test report", which may not be utilized by clinicians in their standard workflow unless a specific concern arises. Instead, clinicians frequently rely on EHR screens that summarize laboratory results through dashboards, flowsheets, tables, or other formats. Currently, no regulatory requirements exist within CLIA or elsewhere to ensure that relevant laboratory data is presented to clinicians or that the tools they use accurately display the relevant laboratory information.

A few states, like New York, define laboratory requirements [23] that are recognized by federal authorities as equivalent or superior to CLIA. However, these state-level requirements would not address the identified gaps relevant to this adverse event.

Table 5: Relevant Excerpts from New York clinical laboratory requirements [23]

Standard	Guidance
<p>Test Request Standard of Practice 4 (TR S4): Urgent Test Request</p> <p>The laboratory must have standard operating procedures and/or policies for the receipt, labeling, processing, and reporting of specimens that are marked as urgent or STAT.</p>	<p>[none]</p>

<p>The procedure must include instructions for reporting critical and alert values.</p>	
<p>Result Review Standard of Practice 1 (RR S1): Result Review Criteria</p> <p>The laboratory must have standard operating procedures for the review of test results for accuracy and reliability.</p> <p>[...]</p> <p>Review of all test results must verify that:</p> <ul style="list-style-type: none"> a) [...] f) patient test results that are consistent with relevant patient information such as age, gender, diagnosis, and relationship are identified; g) reference ranges are appropriate; h) reporting interpretations are appropriate for the test results; and i) abnormal results are flagged, and alert or panic values are communicated according to the laboratory’s established standard operating procedures, protocols or policies 	<p>[none]</p>
<p>Reporting Standard of Practice 2 (REP S2): Test Report Content</p> <p>[...]</p> <p>Test results, whether transmitted electronically or by hard copy, must include all required report information, including:</p> <ul style="list-style-type: none"> [...] c) the date, and hour if required, when the specimen was collected; [...] f) test results, and if applicable, units of measure, reference ranges, or a similar method for identifying abnormal values; [...] 	<p>[none]</p>
<p>Investigation and Corrective Action Standard of Practice 3 (ICA S3): Actionable Events</p>	<p>[none]</p>

<p>The laboratory must define a nonconformity to include any aspect of the test process that does not follow the laboratory’s established standard operating procedure and/or policies, requirements of the Quality Management System (QMS) or client specifications including:</p> <p>[...]</p> <p>f) reference ranges for a test procedure are inappropriate for the laboratory’s test population.</p>	
--	--

Standards and Guidance

Standards and guidance documents exist, although they are not mandated, and they may or may not be used by medical centers. One example is SAFER [24] by the Office of the National Coordinator for Health Information Technology. Table 6 shows excerpts from SAFER that are related to flagging and abnormal values.

Table 6: Excerpts from the SAFER guide

<p>Recommended Practice 1.3 Functionality for ordering and reporting results is tested pre- and post-go-live</p> <p>Examples of Potentially Useful Practices/Scenarios</p> <ul style="list-style-type: none"> - Efforts are made to proactively identify failure points related to EHR-enabled test results delivery
<p>Recommended Practice 2.7 Written policies specify unambiguous responsibility for test result follow-up with a shared understanding of that responsibility among all involved in providing follow-up care.</p> <p>Examples of Potentially Useful Practices/Scenarios</p> <ul style="list-style-type: none"> - Ordering clinicians in any setting assume responsibility for follow-up care, unless that responsibility is unambiguously transferred to another clinician who accepts responsibility.
<p>Recommended Practice 2.8 Workflows that are particularly vulnerable to mishandling of test results, especially critical ones, are identified, and back-up procedures ensure test results are received by someone responsible for the affected patient’s care.</p> <p>Examples of Potentially Useful Practices/Scenarios</p> <ul style="list-style-type: none"> - Situations that are vulnerable to test results follow-up failures are identified. These include handoffs between clinicians (e.g., between residents, part-time physicians, ER physicians, and hospitalists), and care transitions between clinical settings (e.g., between different units of a hospital; between the hospital and home or a post-acute

<p>facility). In these situations, processes should be in place to ensure that test results are communicated to a clinician responsible for follow-up care.</p>
<p>Recommended Practice 2.9 Results outside normal reference ranges, or otherwise determined to be abnormal, are flagged (i.e., presented in a visually distinct way).</p> <p>Examples of Potentially Useful Practices/Scenarios</p> <ul style="list-style-type: none"> - Abnormal results are flagged (e.g., bolded font, asterisk beside values, use of “H” or “L”, different colors) or marked for better visualization in the EHR. - Color is not used as the only visual indicator of clinical significance.
<p>Recommended Practice 2.10 Display of results (e.g., numeric, text, graphical, image) should be easily accessible, clearly visible, not easily overlooked, and understandable.</p> <p>Examples of Potentially Useful Practices/Scenarios</p> <ul style="list-style-type: none"> - Displays of test results undergo usability testing for the intended clinical users. - Result details are reported on one screen, eliminating the need for horizontal scrolling. For example, providers should not have to use additional scrolling (e.g., on the “next page”) to access critical information.
<p>Recommended Practice 3.1 As part of quality assurance activities, organizations monitor selected practices related to test result reporting and follow-up. Monitored practices include clinician use of the EHR for the test results review and clinician follow-up on abnormal test results.</p> <p>Examples of Potentially Useful Practices/Scenarios</p> <ul style="list-style-type: none"> - Clinicians document communication of test results to patients in the EHR - Organizational quality assurance activities select and measure test results-related benchmarks for ongoing monitoring, starting in areas of identified concern and high risk. For example, an organization could develop a measurement system for test results reporting using measures along the following lines: <ul style="list-style-type: none"> o Test results with the lowest follow-up rate are investigated to understand the root causes of the problem.
<p>Recommended Practice 3.2 As part of quality assurance, the organization monitors and addresses test results sent to the wrong clinician or never transmitted to any clinician (e.g., due to an interface problem or patient/provider misidentification).</p> <p>Examples of Potentially Useful Practices/Scenarios</p> <ul style="list-style-type: none"> - Error Logs are used to detect results, such as those that were never delivered, results without any ordering provider, or results with unidentifiable providers.

Unfortunately, guidance documents like SAFER have limited impact. Many medical centers face ongoing financial pressures, causing leadership to prioritize required regulations over optional guidance. As a result, leadership may choose not to invest time, money, and resources in implementing recommendations that are not mandated. The prevailing assumption is often that if a laboratory regulation does not require a particular action, then it is not essential for safe operation.

Most guidance documents tend to be built around general checklists that are broadly applicable. However, they do not account for the unique characteristics of each medical center. Documents like SAFER often lack methods or processes to help facilities systematically identify weaknesses that fall outside the general checklist.

While there is guidance related to flagging and other factors, the available guidance does not appear to address the exact EHR, clinical, and laboratory factors that led to the adverse event described in this report. For example, existing guidance ensures that flagged values are visually distinguished from non-flagged values. However, that wasn't a problem in the adverse event. There is no guidance to recommend that normal values (unflagged because they do not exceed the flagging range) should be presented differently from unflaggable values (unflagged due to a lack of a defined flagging range).

There seems to be no existing guidance to conduct a human factors analysis or even a basic hazard analysis for clinical use of EHRs and laboratory data, as required in most other safety-related industries. Such an analysis would have anticipated many of the technical weaknesses that contributed to this event.