Diagnostic Excellence, AI, and Patient Safety

Town Hall Call

May 21, 2025

THE LEAPFROG GROUP

Webinar Reminders

Accessing the Audio

- If you are using computer audio, please select that option in the audio options pop up.
- If you are joining by phone, please dial in using the Toll Free 800 number provided. Then enter the Meeting ID when prompted, then your Participant ID.
  - The Meeting ID can be found in the confirmation email or in the Zoom meeting by clicking the audio button in the bottom left-hand corner.
  - The Participant ID can be found in the audio options in the bottom left-hand corner.
  - If you forgot to enter the Participant ID when dialing in, please dial # then your Participant ID again followed by #.

Use of the Zoom Chat Function

- The Town Hall Call includes a live Q&A during the presentation; therefore, we do not monitor the chat for questions. Please reserve the Zoom Chat Function for reporting technical issues only.

Accessing the Slides & Recording

- Following each session, a copy of the slides and recording will be posted and available for download on the Leapfrog website here: https://www.leapfroggroup.org/survey-materials/town-hall-calls

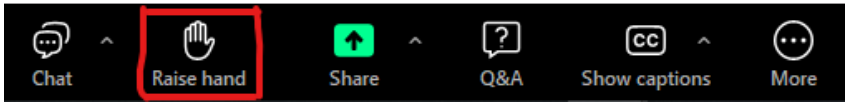
THELEAPFROGGROUP

Q & A

Participants will be able to ask questions throughout the presentation. Please select the Q&A icon at the bottom of your screen:

- Once the icon has been selected a Q&A box will appear for you to type your questions.
- All participants will be able to view the questions and answers during the duration of the webinar.
  - You will be receiving responses in real time from a member of our team.
  - We will include a transcript of the Q&A on the Leapfrog website here: https://www.leapfroggroup.org/survey-materials/town-hall-calls
  - Some questions may be answered live – please pay close attention.

Following the presentation, we will have a live Q&A session. Please use the Raise Hand icon at the bottom of your screen:



- Once the icon has been selected you will be placed in the queue. When it is your turn to ask your question, you will receive a prompt from the host asking you to unmute yourself.

THELEAPFROGGROUP

Introductions

Emily Patterson, PhD
Professor
Ohio State University
College of Medicine

Sangeeta Ahluwalia, PhD, MA, MPH
Senior Policy Researcher
RAND

Margie Shofer, BSN, MBA
Director, General Patient Safety Division, CQuIPS
AHRQ

Jean-Luc Tilly, MPA
Program Director
The Leapfrog Group

Our Agenda

- Implementing Diagnostic Excellence Across System (IDEAS)
  - Calibrate Dx,
  - Measure Dx
- AI, Human Factors, and Diagnostic Excellence
- Diagnostic Excellence in the 2025 Hospital Survey

Implementing Diagnostic Excellence Across Systems (IDEAS) Project

Project Overview

AHRQ Agency for Healthcare Research and Quality

PATIENT SAFETY

Diagnostic safety is critically important

- Diagnostic safety events contribute to preventable harm in 1 in 20 outpatient visits and up to 250,000 inpatient stays each year
- They are the most common root cause for medical liability claims
- Healthcare organizations and clinicians need actionable, practical, and flexible resources to support diagnostic safety improvement efforts.

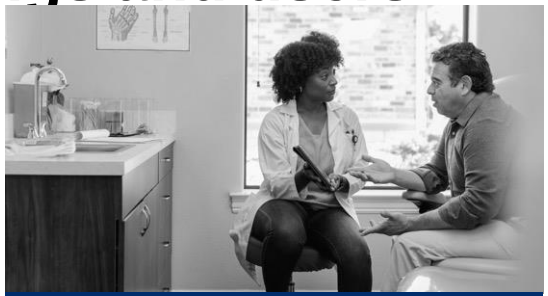
Singh et al. BMJ Qual Saf. Sep 2014;23(9):727-731.
Gunderson et al. BMJ Qual Saf. 2020 Dec;29(12):1008-1018.
Hanscom et al. Coverys.com. March 13, 2018.

AHRQ is committed to improving diagnostic safety

- The Agency for Healthcare Research and Quality (AHRQ) has developed diagnostic safety resources to support the delivery of high-quality healthcare
- Resources were created by Baylor College of Medicine and MedStar Health for AHRQ; tested with clinicians at multiple institutions
- RAND is leading the **IDEAS Project**, an AHRQ-funded national effort to recruit healthcare organizations to implement up to three diagnostic safety resources


SDI Productions/Getty (E+)

IDEAS supports diagnostic safety across settings and users



Toolkit for Patients To Improve Diagnostic Safety

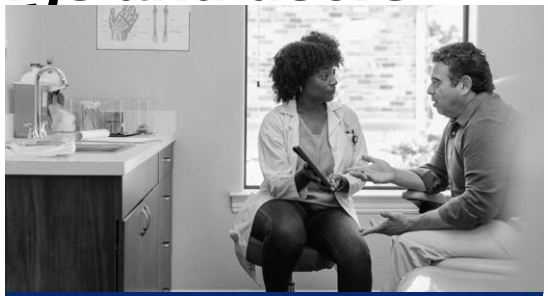
Measure Dx

Calibrate Dx

	Toolkit for Patients To Improve Diagnostic Safety	Measure Dx	Calibrate Dx
SETTING	Office-based setting	Any healthcare setting	Any healthcare setting
USERS	Clinicians and their patients	QI or patient safety teams	Individual licensed clinicians whose scope of practice includes diagnosis

NIKCOA/Adobe Stock; C Malambo/Adobe Stock; SDI Productions/Getty

IDEAS supports diagnostic safety across settings and users



Toolkit for Patients To Improve Diagnostic Safety

Measure Dx

Calibrate Dx

	Toolkit for Patients To Improve Diagnostic Safety	Measure Dx	Calibrate Dx
SETTING	Office-based setting	Any healthcare setting	Any healthcare setting
USERS	Clinicians and their patients	QI or patient safety teams	Individual licensed clinicians whose scope of practice includes diagnosis

C Malambo/Adobe Stock

Why use Measure Dx?

- Diagnostic errors are major contributors to patient harm, but they are difficult to identify and measure
- Measuring for improvement is key to a learning health system approach
- The measurement process itself can facilitate learning and improvement
- Several measurement strategies have already demonstrated “proof of concept” for use in health systems

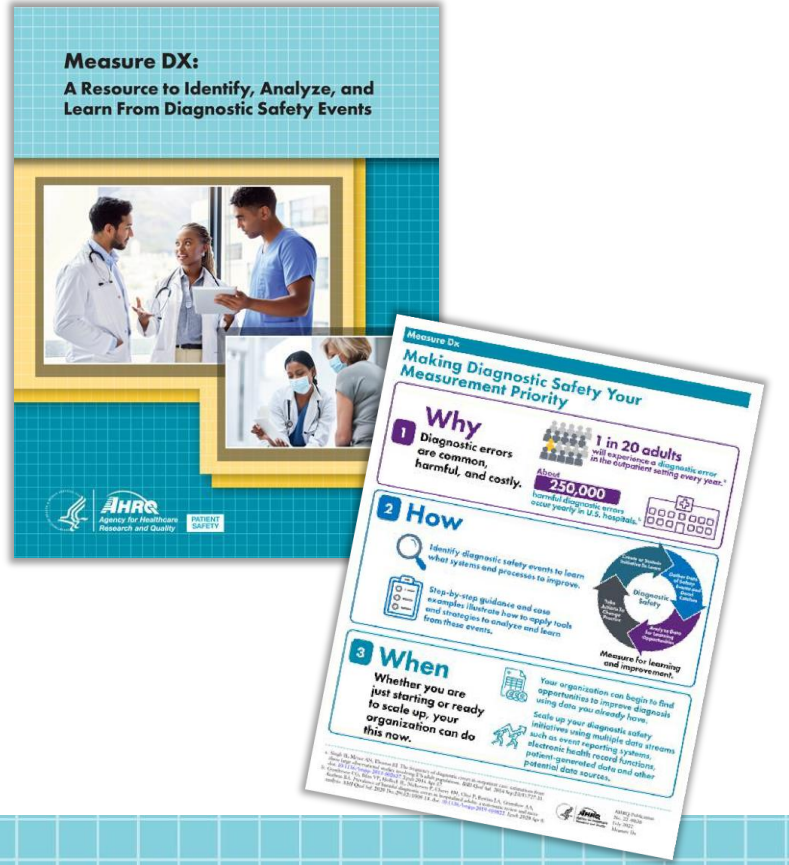
Measure Dx was developed by

How does Measure Dx work?

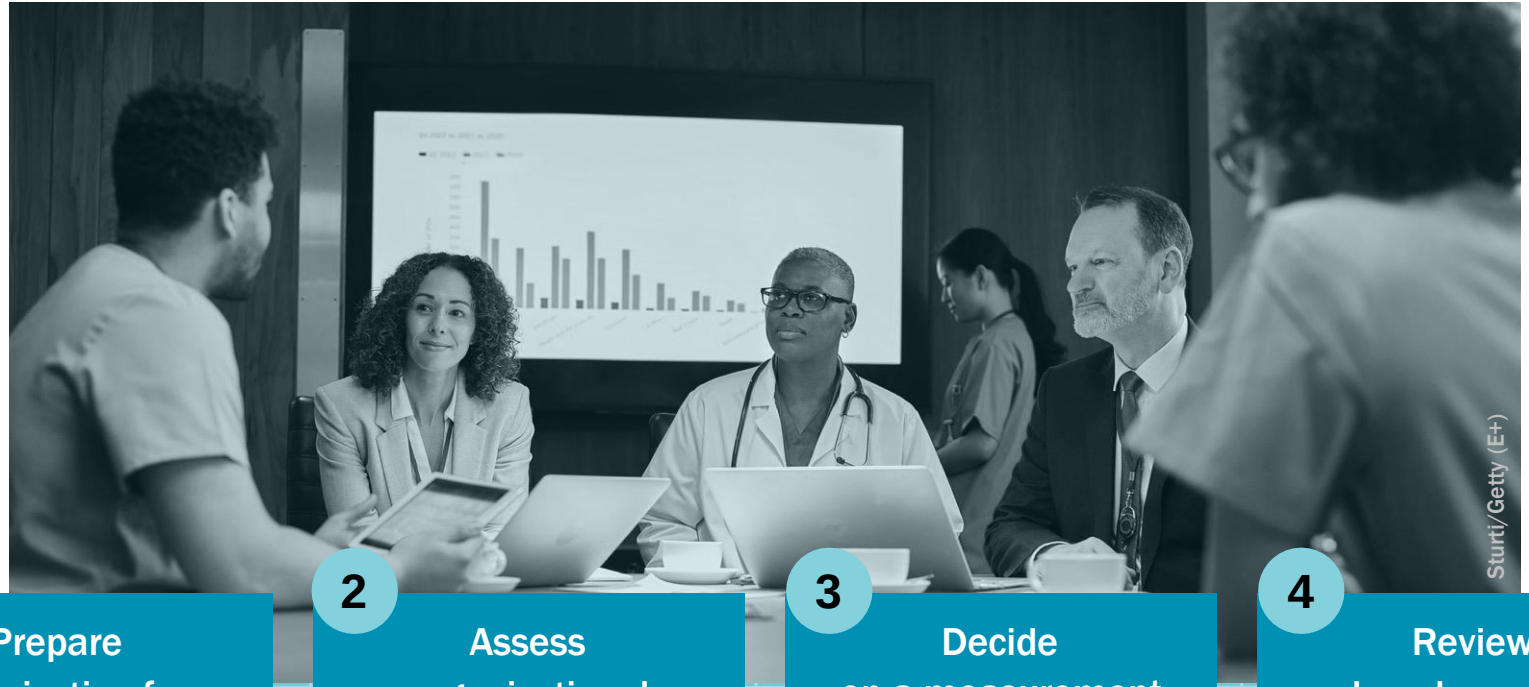
- Helps quality improvement (QI) teams identify and use specific strategies to detect, analyze, and learn from diagnostic safety events

VALUE ADDED

Use existing data to reduce diagnostic harms at an organizational level



Measure Dx guides QI teams through four steps



1. **Prepare** organization for diagnostic safety measurement
2. **Assess** organizational readiness (i.e., self-assessment)
3. **Decide** on a measurement strategy
4. **Review** and analyze cases of interest

Who can test

- Any healthcare setting or site, including those new to learning about diagnostic safety
- Sites that have or that can create a QI team with access to patient-level data
- Entire hospitals or specific units (e.g., emergency departments, intensive care units) or ambulatory clinics

What is needed to implement Measure Dx in this project?

- Commit to this 2-year effort
- Demonstrate executive support
- Identify a site champion to help coordinate project activities and communicate with the project team
- Identify or assemble a QI team that includes a clinical lead and a quality manager
- Engage in project and data collection activities

It was very easy for me to follow. I like the fact that I could share a document with the team, which had everything laid out **very easy to understand**.

—Measure Dx pilot user

The project became an opportunity, it really put some structure around what we were going to do. It was a vehicle to **get safety leadership in the hospital to come on board**.

—Measure Dx pilot user

One of the things that's very clear in the tool is you don't have to solve all these problems. Start small, pick an area you want to focus on, and start with the resources that you have. **Being able to fit the resource to people's local environment is huge.** Measure Dx does a good job of helping with that.

—Measure Dx pilot user

Why use Calibrate Dx?

- Delayed, wrong, or missed diagnoses are common challenges for patients, families, and clinicians, yet clinicians rarely receive feedback on their diagnostic decision making
- Lack of feedback leads to a “calibration gap,”[1] or a difference between one’s perceived and actual diagnostic performance.
- Clinicians who are aligned on both their diagnostic accuracy and their confidence in diagnosing patients will achieve diagnostic excellence

[1]Omron R, et al. *AEM Educ Train* 2018;2:339-342

Why use Calibrate Dx? (Continued)

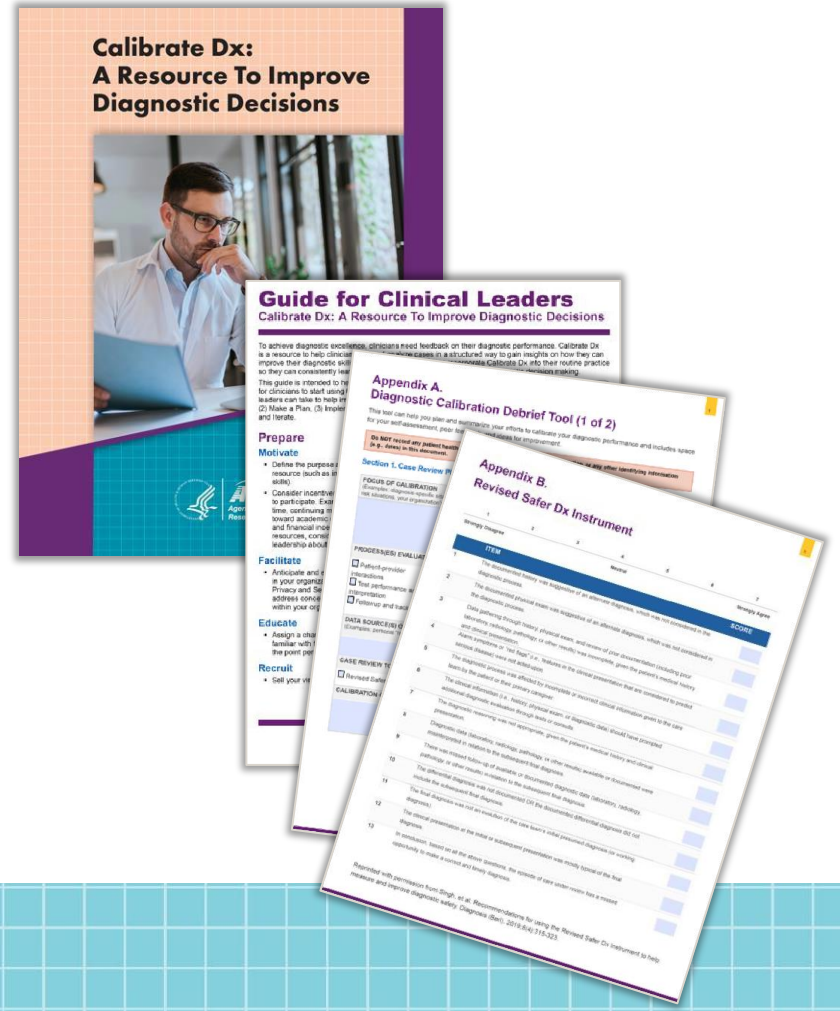
- Calibrate Dx was developed at Baylor College of Medicine in partnership with MedStar and tested in collaboration with clinicians at multiple institutions
- It uses active reflection and peer feedback to improve diagnostic calibration and confidence

How does Calibrate Dx work?

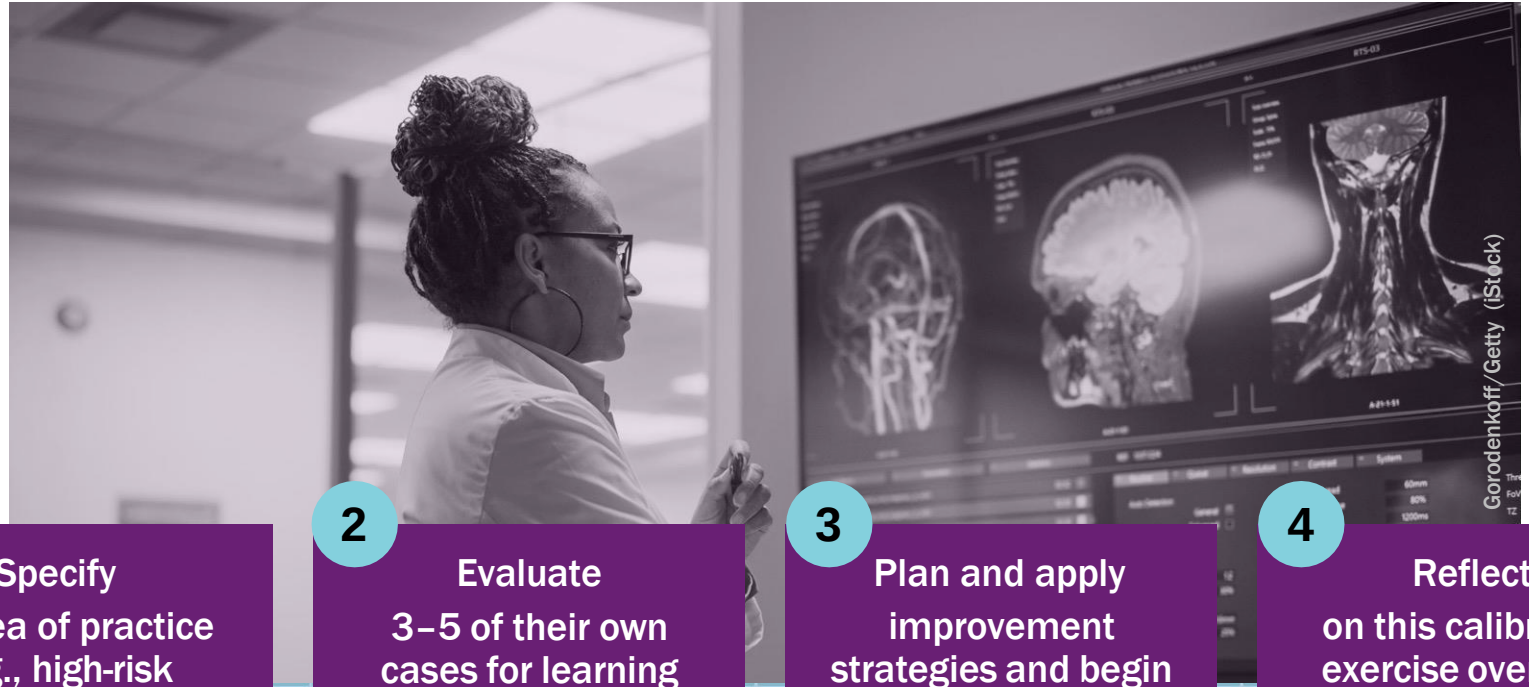
- Self-evaluation tool that guides clinicians in reviewing cases, reflecting on their diagnostic performance, and debriefing with trusted colleagues

VALUE ADDED

Empower clinicians to learn from their own clinical practice through structured exercises and tools



Calibrate Dx guides clinicians through 4 steps



1

Specify

an area of practice (e.g., high-risk situations, commonly missed or delayed diagnoses)

2

Evaluate

3–5 of their own cases for learning opportunities

3

Plan and apply

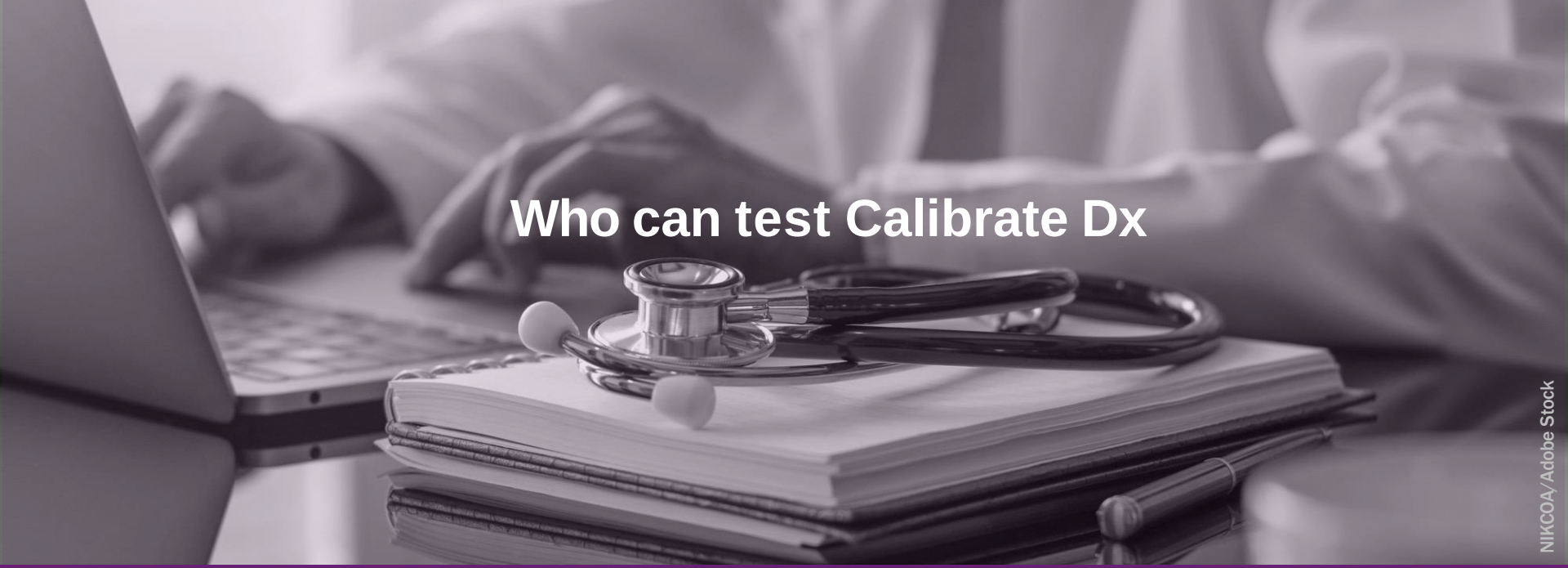
improvement strategies and begin to take actions

4

Reflect

on this calibration exercise over time

Who can test Calibrate Dx



NIKCOA/Adobe Stock

- Sites with at least 4–6 clinicians willing to commit to testing Calibrate Dx
- Small and large group practices, departments within clinics or hospitals, and physician or other clinician training programs

What is needed to implement Calibrate Dx in this project?

- Commit to this 2-year effort
- Demonstrate executive support
- Identify 4-6 clinician participants; engage in project and data collection activities
- Identify a site champion to help coordinate project activities and communicate with the project team
- Ensure clinician access to EHR data for case reviews
- Identify site-specific contacts who can help troubleshoot privacy, security, or confidentiality barriers

So, I think it's helped me reason more and explain more of my reasoning in my documentation. I also think it's **helped me naturally think about other factors at play in my own decision**

—Calibrate Dx pilot user

This is awesome. Like **I'm absolutely going to keep using this**.

—Calibrate Dx pilot user

Hearing the **reinforcement from my colleague and recognition** of, "this sounds like it was a really challenging case, but you did the best that you could, maybe here's one or two things that I would've done differently that you could apply in the future." I think it was good camaraderie building and also reinforcement that…we did the best we could.

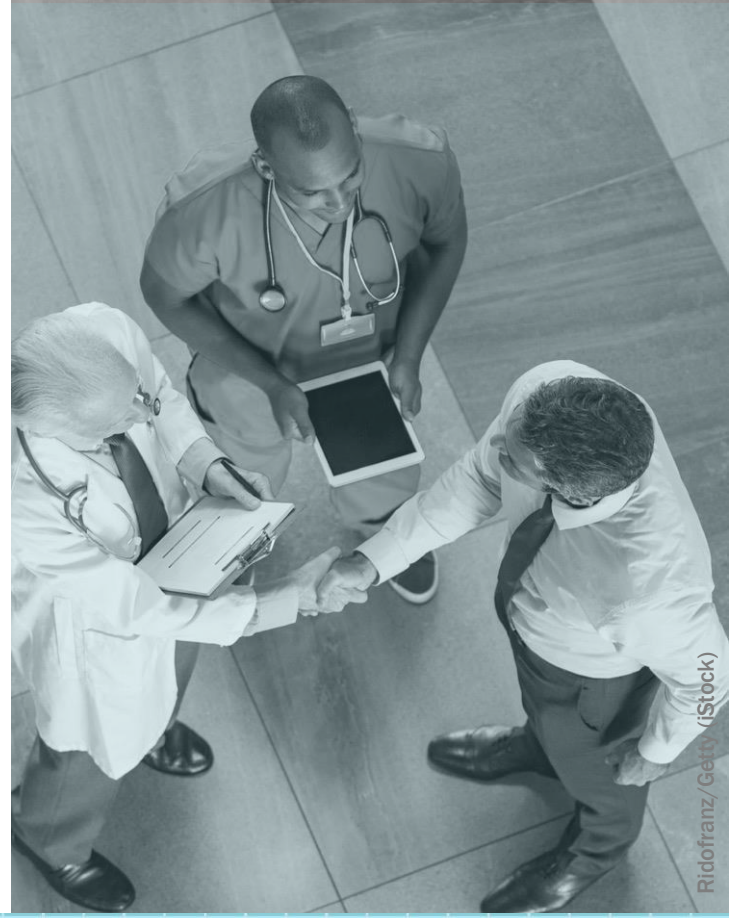
—Calibrate Dx pilot user

Why participate in IDEAS?

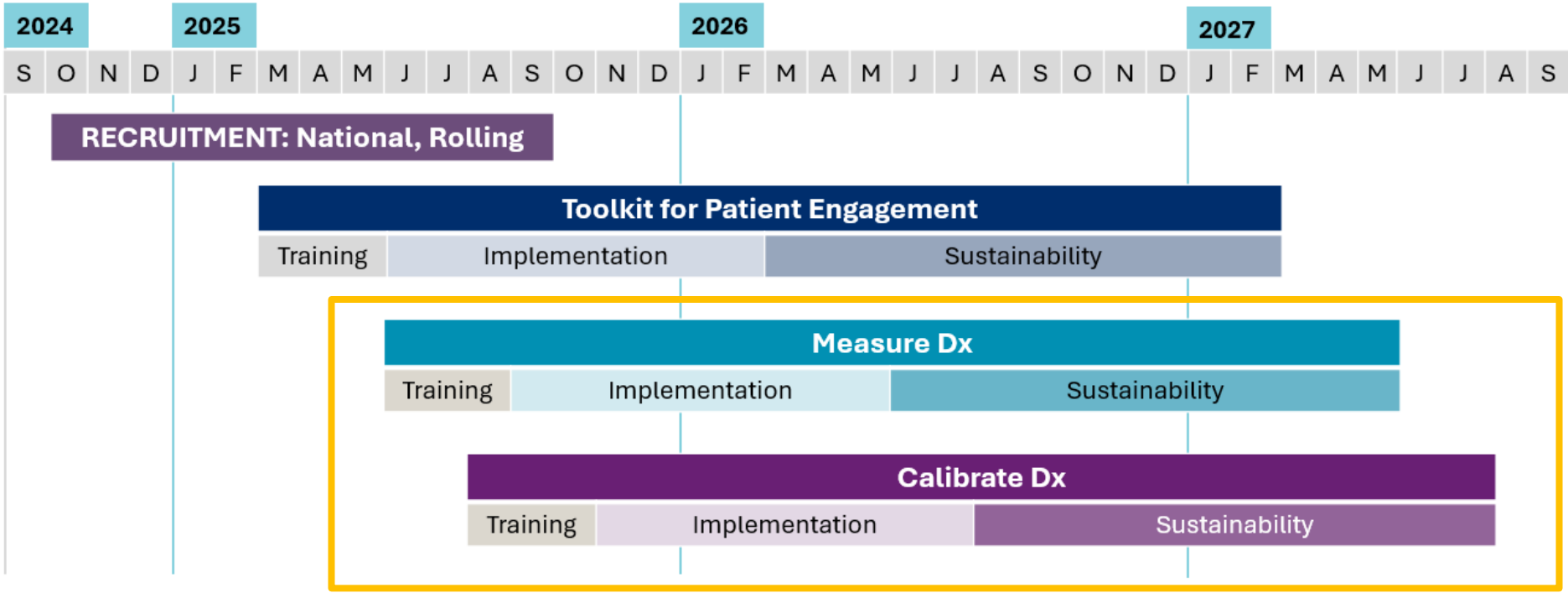
- Implement innovative methods for advancing diagnostic safety
- Improve care, minimize risk of harm
- Receive tailored support and training on the AHRQ resources
- Connect with the project's Learning Collaboratives
- Reduce opportunities for legal claims and lawsuits
- Obtain free continuing education (CE) credits
- Engage in Maintenance of Certification (MOC) improvement activities
- Receive a modest stipend

What support will IDEAS provide?

- Flexible training for all sites including resource-specific materials and tip sheets
- CE credits and ABMS-approved MOC credits
- Hands-on assistance including dedicated email for questions and “office hours” with content experts
- Learning Collaboratives with other participants
- Feedback on the impact of the intervention (available to participants at end of study)



Implement multiple resources with staggered starts



Want to learn more?

VISIT: https://www.ahrq.gov/ideas-project

EMAIL: Courtney Gidengil, MD, MPH

Sangeeta Ahluwalia, PhD

IDEASproject@rand.org

AHRQ Agency for Healthcare Research and Quality

PATIENT SAFETY

Toolkit for Engaging Patients To Improve Diagnostic Safety

- Combines two easy-to-launch approaches to improve patient-clinician communication

VALUE ADDED

Engage patients directly, increase patient satisfaction, and enhance clinician reflection to reduce diagnostic harms in office-based practices

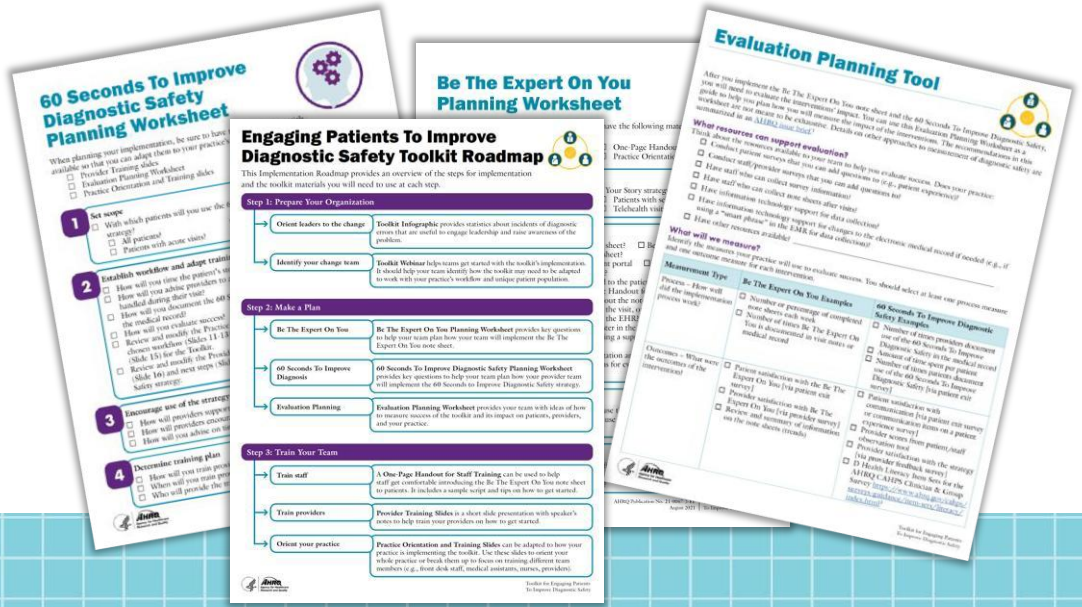
Toolkit has two components

1 Be The Expert On You

Patient-facing component

2 60 Seconds To Improve Diagnostic Safety

Clinician-facing component



Who can test the Toolkit

- Any office-based practice with at least two to three clinicians
- Primary care and specialty practices of any type

SDI Productions/Getty

UNLOCKING THE FUTURE OF HEALTHCARE: AI AND HUMAN FACTORS CONFERENCE

April 8-11, 2025 | Hershey PA
https://research.med.psu.edu/ai-human-factors-conference/

Emily S. Patterson

May 21, 2025

The Ohio State University

ACKNOWLEDGMENTS

This work is supported by a conference grant (Bruno PI) from the Agency for Healthcare Research and Quality. The contents, findings, and conclusions do not necessarily represent the views of AHRQ. Additional sponsorship was provided by Penn State Health, The Society for Medical Image Perception, Medical Image Perception Society, RadAI, AIDoc, The Leapfrog Group, Siemens Healthineers.

THE OHIO STATE UNIVERSITY

THE CHALLENGE WITH AI

When implementing AI-based technology, it could:

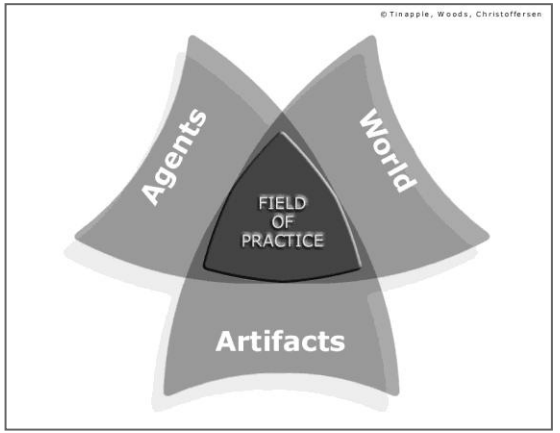
- Fail to be **adopted**
- Fail to be **used as intended** (e.g., workarounds)
- Be trusted when it is wrong (**brittle**)
- Negatively affect **workload** and **workflow**
- **Interfere** with existing sources of system resilience

KNOWLEDGE GAPS WITH AI

1) Develop, validate, implement, maintain embedded AI
2) Redesign health systems to safely deploy
3) Cooperative human-autonomy teaming (HAT)
4) Achieve learning (e.g., foundation models)
5) Balance innovation, standardization, regulations
6) Clinical, cultural, computational considerations
7) Guide a patient's trajectory through the system

HUMAN FACTORS

Applies theoretical frameworks from behavioral and social sciences to enable **experts** in specialized roles to **effectively and easily use** complex systems with embedded **automation**



THE OHIO STATE UNIVERSITY

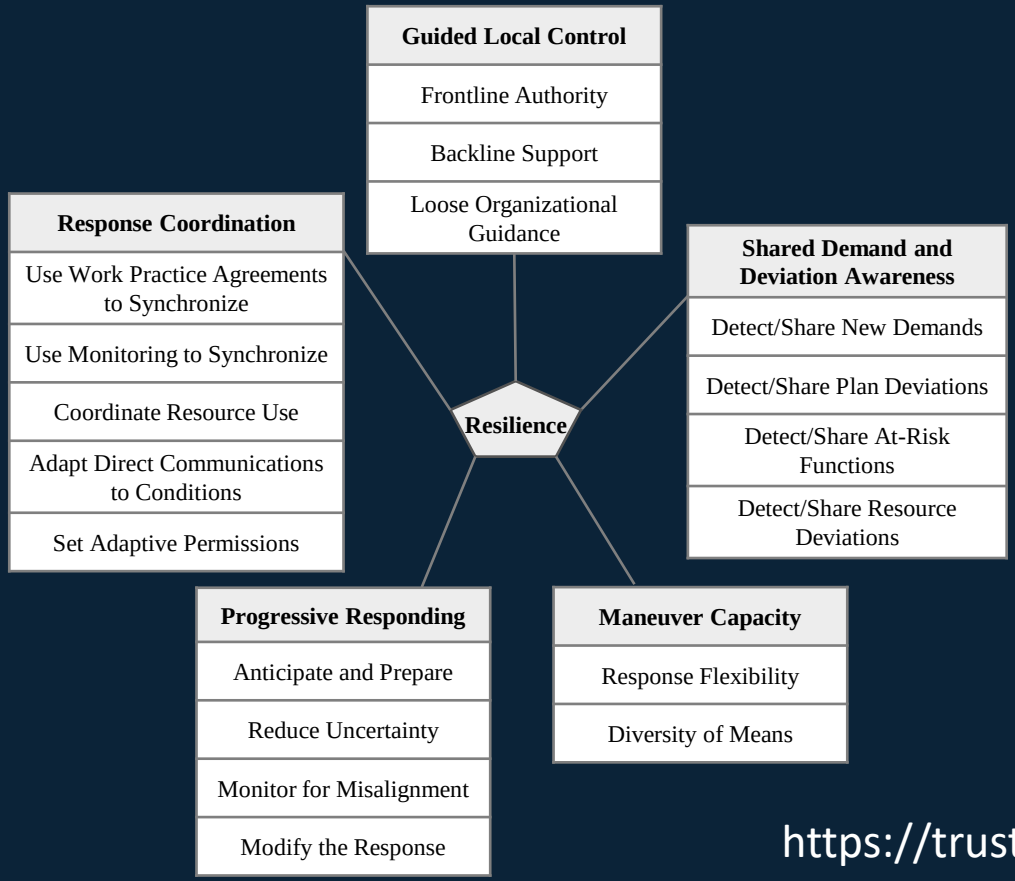
WORK SYSTEM RESILIENCE

The intrinsic ability of a work system to **adaptively adjust its functioning** prior to, during, or following events, changes, disturbances, and opportunities, and thereby **sustain operations** required to achieve the system's goals and mission

Adapted from Nemeth, Wears, Woods, Hollnagel, & Cook, 2008

MITRE's TRUSTS Framework 1.0

Tiers 1 and 2: Factors and Subfactors



https://trusts.mep.mitre.org

MITRE

EXAMPLE: REQUIREMENT FOR SHARED D&D AWARENESS (SDDA)

Agents are informed (and inform others) about new and developing demands

ED Status Board had information used to recognize a patient overflow hidden behind the primary display

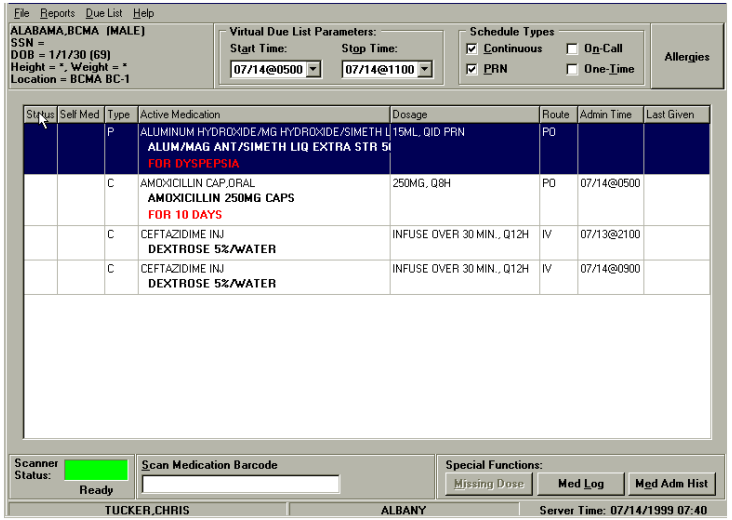
Perry & Wears 2012

EXAMPLE: REQUIREMENT FOR RESPONSE COORDINATION (RC)

Anticipation and coordination are enabled by agents that know about other units and how to coordinate by observing or “listening in” to their activities

Bar Code Medication Administration V1.0 displayed administration times based on scanning bar code for nurses, but physicians did not have access

Patterson, Cook & Render 2002



THE OHIO STATE UNIVERSITY

7 OPPORTUNITIES FOR AI NOW FROM AN HF PERSPECTIVE

1. Fix EHR configuration errors
2. AI critiques diagnoses of human experts
3. AI critiques plans of human experts
4. Reduce alarm overload
5. Improve care transitions (second sign-out, consults)
6. Align AI with context-based goal prioritization
7. Text conversion (role, health literacy, language)

EHR CONFIGURATION ERRORS (1)

- Adult dose ranges in pediatric hospital
- ‘Broken data pipe’ for sepsis alert
- Attending physician set to medical student
- Physician mapped to wrong hospital
- Patient scheduled for the wrong type of provider
- Alerts, lab data sent to wrong provider
- User assigned to incorrect roles
- Incorrect permissions for site-specific roles

IT CRITIQUE OF DIAGNOSIS (2)

Error Type	Group	Case 1	Case 2	Case 3	Case 4
1. Rule out hypothesis incorrectly	Control	40.6%	37.5%	40.6%	40.6%
	Treatment	28.1%	34.4%	18.8%	28.1%
2. Failure to rule out when appropriate	Control	37.5%	15.6%	34.4%	34.4%
	Treatment	34.4%	21.9%	9.4%	25.0%
3. Failure to collect converging evidence	Control	37.5%	34.4%	31.3%	34.4%
	Treatment	3.1%	9.4%	9.4%	12.5%
4. Data implausible given answer	Control	15.6%	21.9%	3.1%	21.9%
	Treatment	0.0%	15.6%	6.3%	15.6%
5. Answer implausible given prior probabilities	Control	21.9%	15.6%	12.5%	40.6%
	Treatment	3.1%	25.0%	9.4%	12.5%

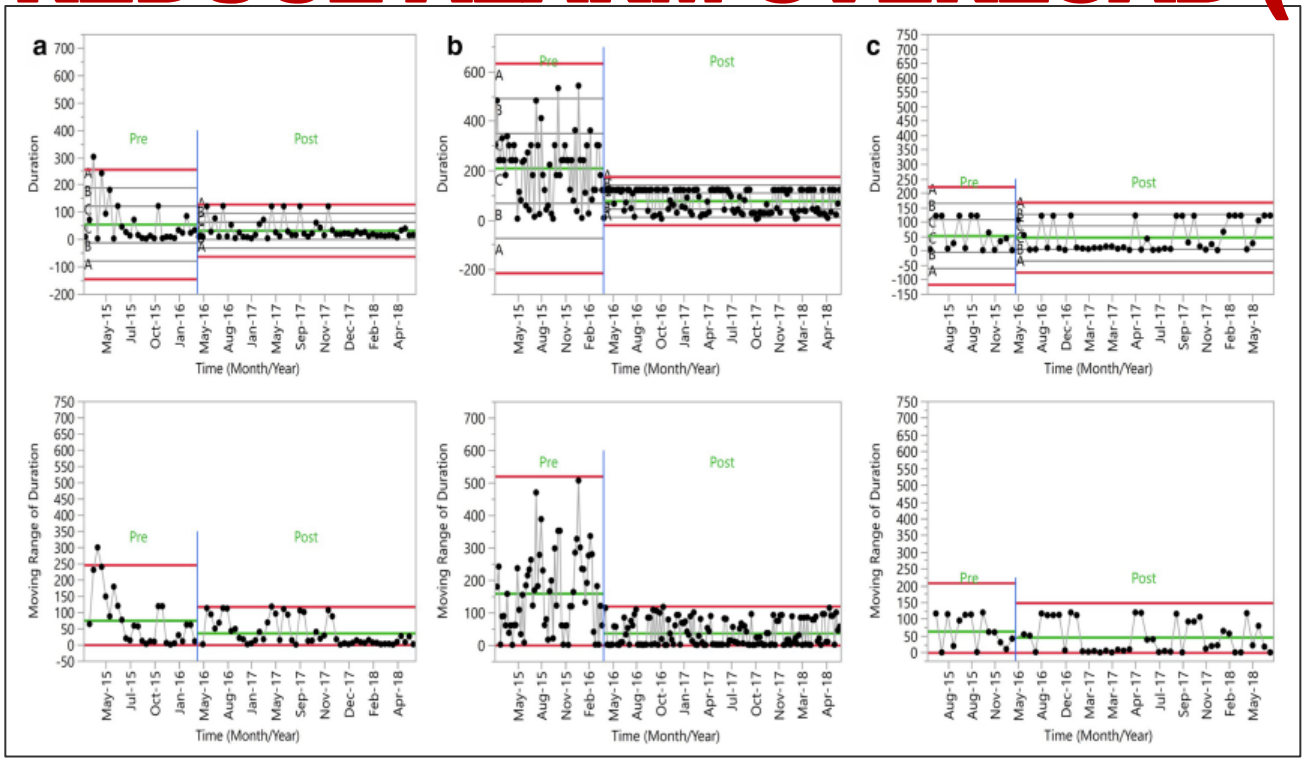
https://doi.org/10.1518/001872099779577363

IT CRITIQUE OF FLIGHT PLAN (3)

- Suggest ‘good’ route (waypoints), altitude, air speed
- Suggest alternative routes, altitudes, air speeds
- Alert when new data are available
- Identify significant problems with a proposed flight plan
  - Goes through forecast weather pattern
  - Fuel consumption
  - Suboptimal destination parameters

https://doi.org/10.1177/001872089403600106

REDUCE ALARM OVERLOAD (4)



THE OHIO STATE UNIVERSITY

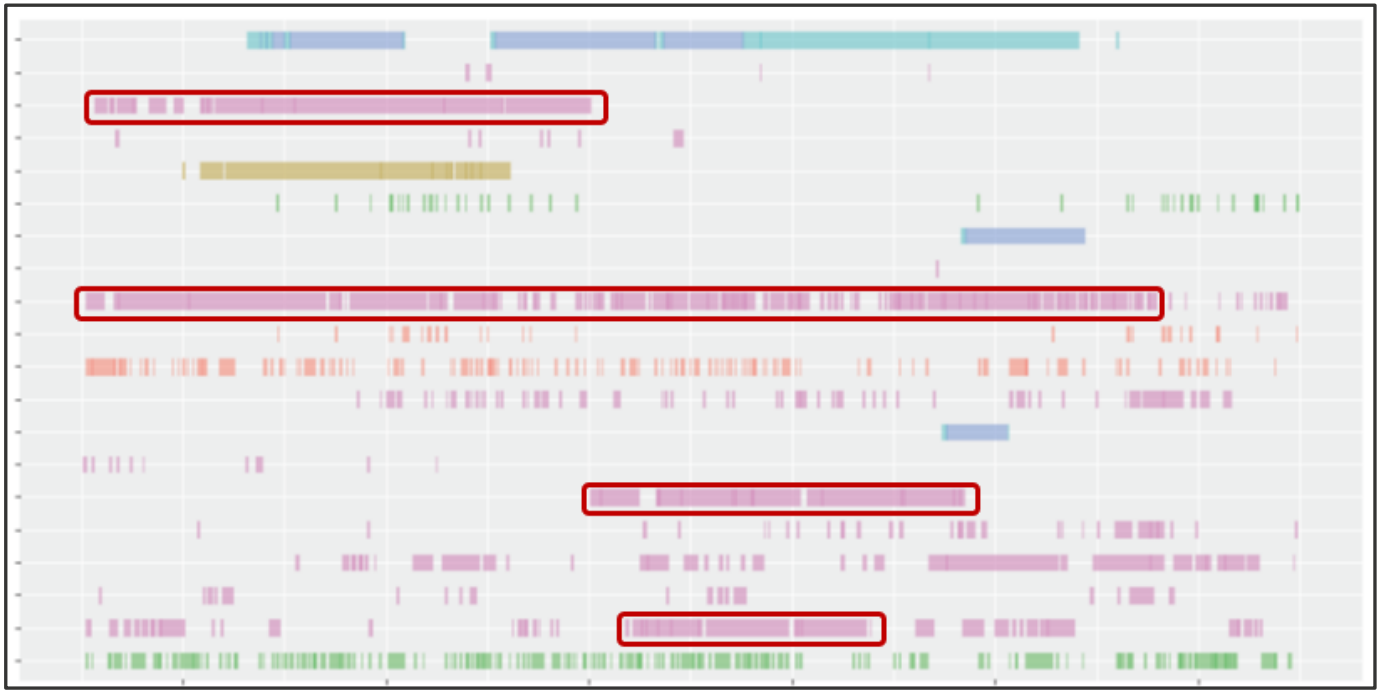
Funded by AHRQ, ONR

15-59% faster to respond
20-38% reduction in alarm burden

DETECT NUISANCE ALARMS (4)

SPO_2 low 37-69% alarms on the unit

NON-ACTIONABLE ALARMS (4)



Most low heart rate alarms have no new information

DETECT ARTIFACTS (4)

Leads fail, ‘no telemetry’ alarms are artifacts during baths and ‘jostles’

IMPROVE CARE TRANSITIONS (5)

- Transcribe first sign-out and make summary available during second sign-out
  - [This patient] has been here 16 hours. He is admitted for an upper GI bleed. I don't know anything else.
- ChatGPT prompts for ED consult request
  - Reason for the consult
  - Pertinent history and physical findings
  - Relevant labs/imaging already done
  - Urgency of the request
  - Clear contact information or who to follow up with

ALIGN AI WITH CONTEXT-BASED GOAL PRIORITIZATION (6)

Safety, Production, Quality of Work Life

Security, Cost/Revenue, Quality of Care, Workflow Efficiency

Predictability-Flexibility-Adherence to Best Practice-Compliance

Documentation, Equitable Access

TEXT CONVERSION (7)

- Prompt: Please translate a radiology report into plain language that is easy to understand.
  - PLEURA: No pleural thickening or effusion. No pneumothorax. HEART: Heart size normal. No pericardial effusion. CORONARY ARTERY CALCIFICATION: None. MEDIASTINUM/HILUM/AXILLA: No adenopathy
- Response:
  - The pleura, heart and blood vessels are normal, and there is no sign of cancer in the lymph nodes.

Lyu, Q., Tan, J., Zapadka, M. E., Ponnatapura, J., Niu, C., Myers, K. J., ... & Whitlow, C. T. (2023). Translating radiology reports into plain language using ChatGPT and GPT-4 with prompt learning: results, limitations, and potential. Visual Computing for Industry, Biomedicine, and Art, 6(1), 9.

TO HAVE RESILIENT AI, WE CAN…

- Combine structure and adaptation
- Proactively defend against ‘brittleness’
- Facilitate sources of resilience:
  - Detect and share demands and deviations
  - Progressively probe and respond
  - Local control supported by guidance
  - Synchronize activities and coordinate resources
  - Protect margin of maneuver

THANK YOU!

patterson.150@osu.edu

Leapfrog's National Initiative for Hospitals

Recognizing Excellence in Diagnosis
A Program of The Leapfrog Group

2022: Published Recommended Practices Report describing 29 options for hospitals looking to reduce diagnostic errors

2023: Measured implementation progress in pilot survey of 95 hospitals across the country

April 2024: Introducing new process and measures for inclusion in the 2024 Leapfrog Hospital Survey

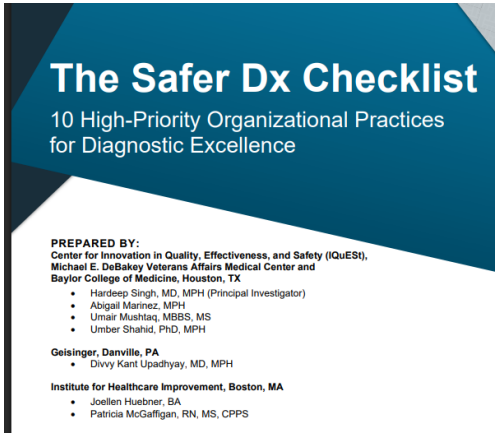
July 2024: Update to the Recommended Practices Report

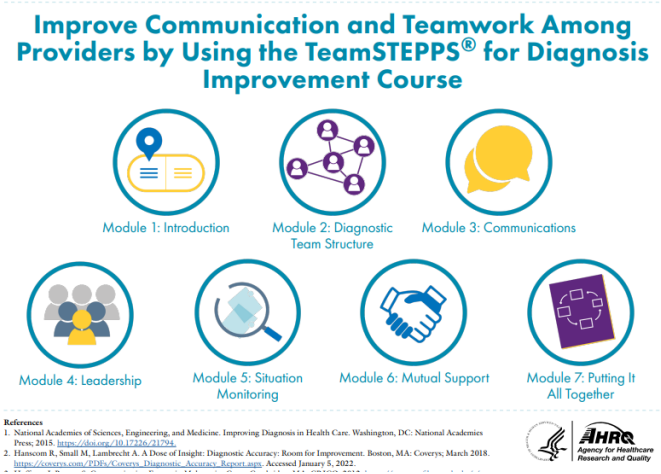
2025
Revised
Diagnostic Excellence in 2025 Hospital Survey

THELEAPFROGGROUP

Updates to Diagnostic Excellence in the 2025 Hospital Survey

- Section remains **optional** and results will not be scored or publicly reported
- Specific Tools:

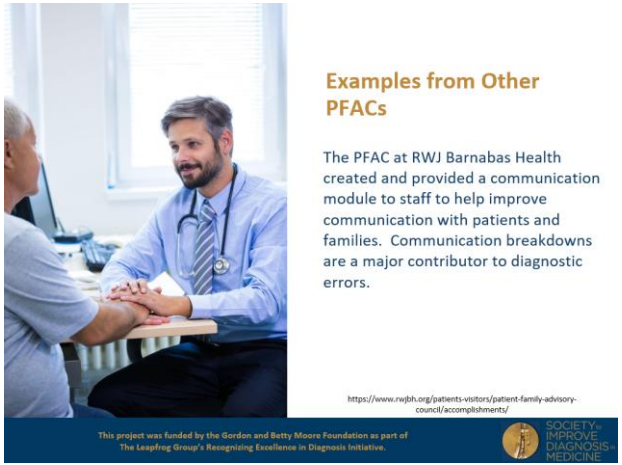




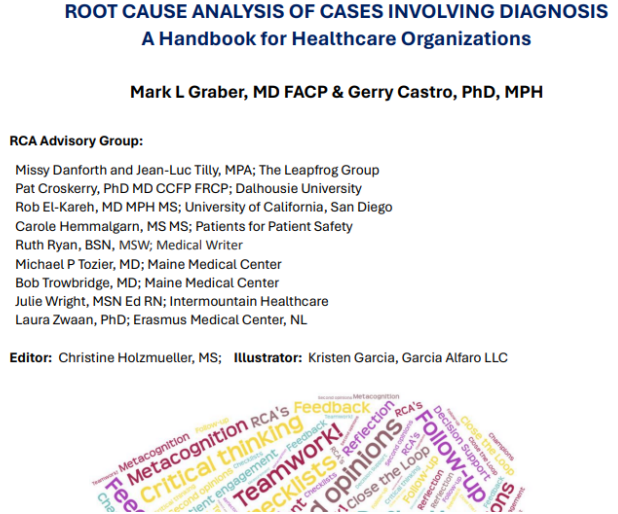
Updates to Diagnostic Excellence in the 2025 Hospital Survey

- Focusing on diagnosis:
  - Question #1: commitment to diagnostic excellence must specifically mention diagnosis or diagnostic errors, and must be by current CEO or CMO
  - Question #4: PFAC engagement must be specific to diagnosis at the hospital itself
  - Multidisciplinary teams focused on diagnosis should be:
    - At the hospital campus level, not system level
    - Focused on diagnosis, not the established patient safety committee

New Resources for Hospitals



<u>Patient and Family Advisory Council (PFAC) Toolkit</u>
To help hospital PFACs learn about diagnosis and diagnostic quality and explore ways to reduce diagnostic error in their institutions



<u>Root Cause Analysis of Cases Involving Diagnosis</u>
A handbook for hospitals conducting RCAs of cases involving diagnosis.

THELEAPFROGGROUP

Thank you for joining us today.

Questions? Contact the Help Desk at https://leapfroghelpdesk.Zendesk.com

THELEAPFROGGROUP