

2020 ACR-RBMA Practice Leaders Forum



Using AI tools to Drive Quality and Ensure Patient Safety Across the Radiology Practice

Jason Wiesner, MD, MS, MBA, Diagnostic Imaging Service Line Medical Director, Sutter Health

@xrayviznMD

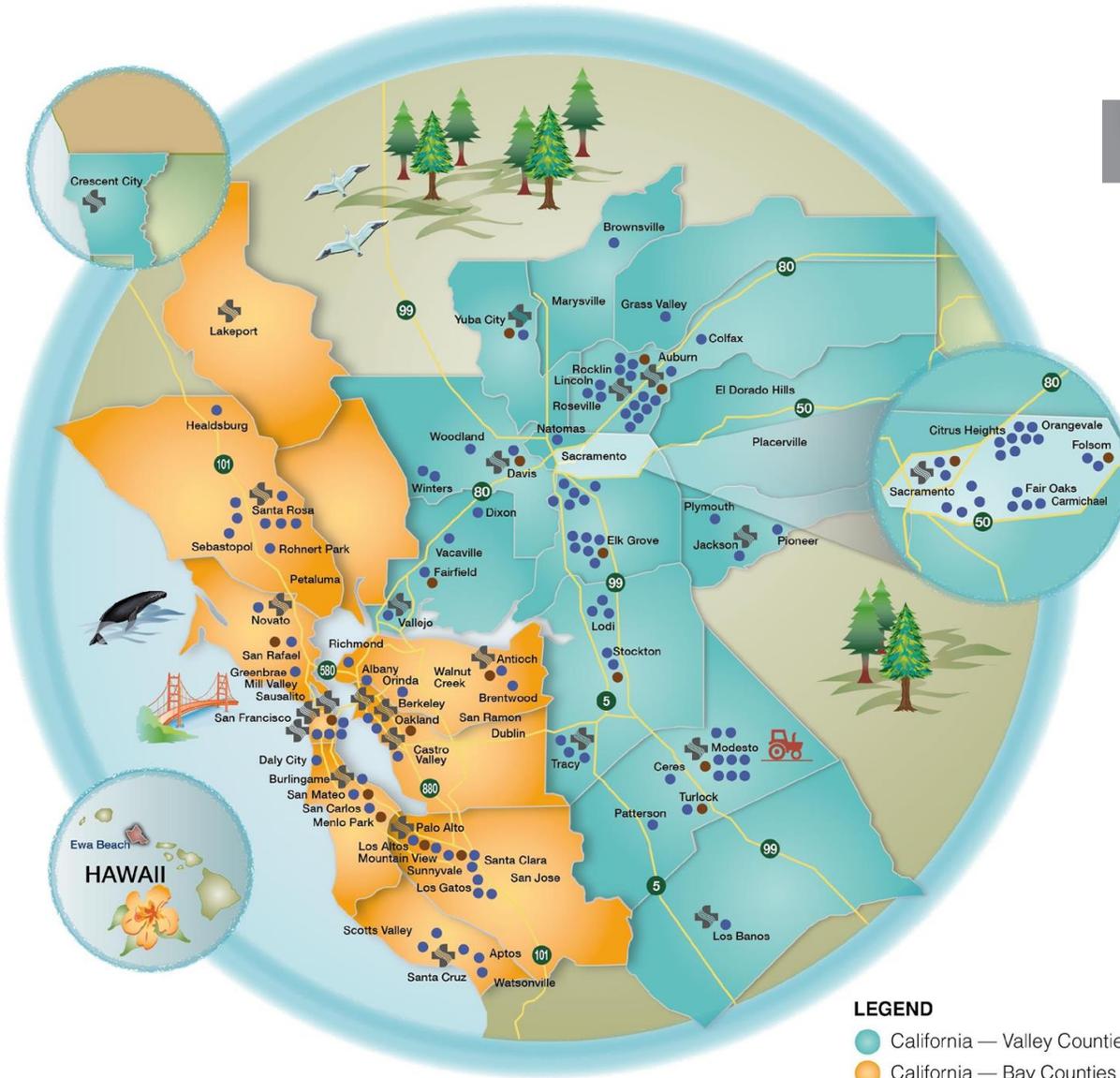
Disclosures

- Salary—Sutter Health system office, Sutter Health Information Services, Sutter Medical Foundation, Sutter Medical Group



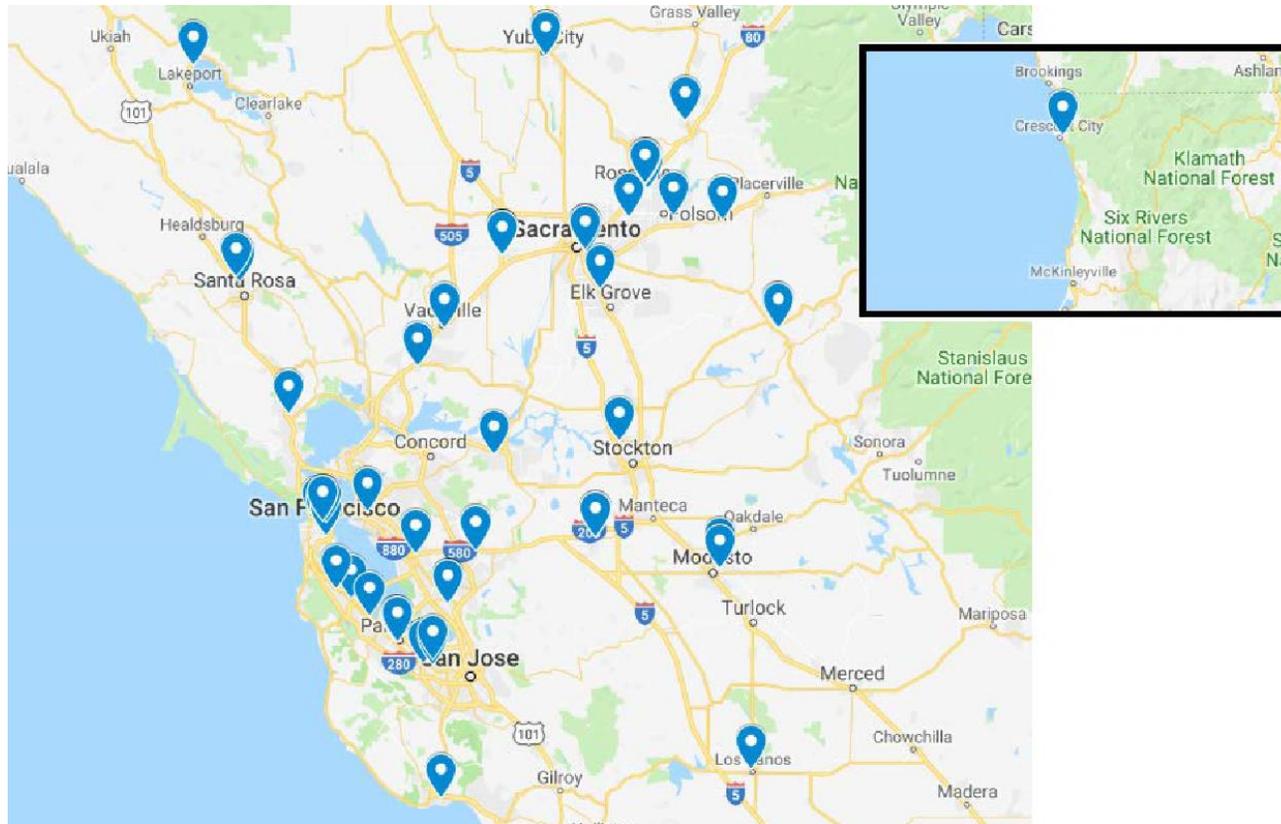
Sutter Health at a Glance

Physicians	18,000
Network and affiliate employees	53,000 +
Patients cared for each year	3.5 million



LEGEND

- California — Valley Counties
- California — Bay Counties
- Hospitals
- Physician Offices
- Urgent Care



Sutter Health Imaging Services
Annual Exams: Approx 3.2 million
Acute Facilities: 23
Outpatient Centers: 67
Radiology Groups: 11
Referring Physicians: 15-18K
Imaging Team Members: >2500

Total Diagnostic Imaging Assets Managed by Radiology \$438,930,000

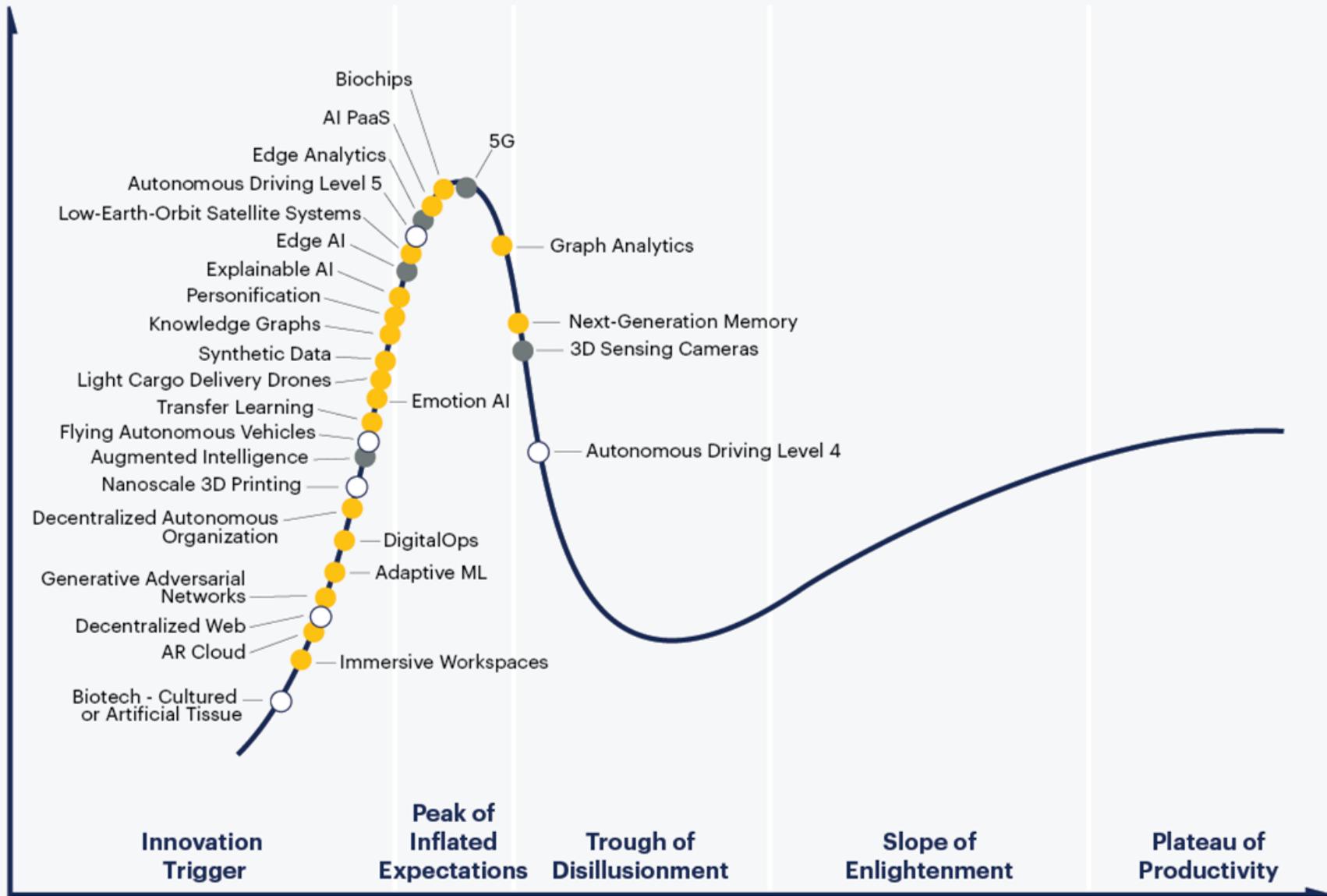
Learning Objectives

- Review the current literature relative to the clinical impact of sources of error in healthcare with a focus on Radiology
- Discuss why we are focusing on the wrong problems in AI and why I believe we should focus on “Augmented AI” in Radiology
- Share Sutter Health experience using AI tools in diagnostic imaging with a focus on Quality and Safety applications

Artificial Intelligence—hype versus hope

- AI is already all around us, whether we are conscious of it or not.
 - Google Translate
 - Gmail Autocomplete
 - Voice assistants (Siri, Alexa)
 - Vehicle Navigation and driver assist
 - NLP use in Coding and billing
 - NLP use in radiology report analysis, data mining, and quality review

Expectations



Plateau will be reached:

○ less than 2 years

● 2 to 5 years

● 5 to 10 years

○ more than 10 years

● obsolete before plateau

As of August 2019

Artificial Intelligence—hype versus hope

TECH

Here's why one tech investor thinks some doctors will be 'obsolete' in five years

PUBLISHED FRI, APR 7 2017•2:28 PM EDT | UPDATED FRI, APR 7 2017•5:15 PM EDT



Christina Farr
@CHRISSEYFARR

SHARE

Artificial Intelligence for Medical Imaging Market to Top \$2B

  [AI to help the sleep-deprived catch more Zzs](#)

People who struggle to get a good night's sleep and seek medical help for the problem are producing mega data on things like eye movement, breathing, brain activity and restless legs. Which is to say sleep medicine is as ripe as any field in healthcare for help from AI.

Featured Articles

  [NLP identifies cancer patients suffering in silence from social isolation](#)

A natural language processing algorithm has achieved 90% precision in automatically spotting signs of social isolation in cancer patients by "reading" clinical notes in a hospital's electronic health record.

 [AI company Health at Scale raises \\$16M](#)

A San Jose-based company focused on machine learning for care optimization, Health[at]Scale, has raised \$16 million in a Series A funding round, with Optum as its sole investor.

 [Mental-health professionals urged to step up human oversight of 'robot therapists'](#)

Academic and popular writings on the use of "embodied" AI in mental healthcare are piling up fast. But where's the guidance for psychiatrists, psychotherapists and clinical social workers looking to use robots, avatars and chatbots with real patients?

  [AI could soon diagnose disease without office visit](#)

Instead of Googling symptoms to when feeling an ailment and landing on an incorrect diagnosis, AI could soon provide accurate diagnosis without needing to go to a doctor's office.

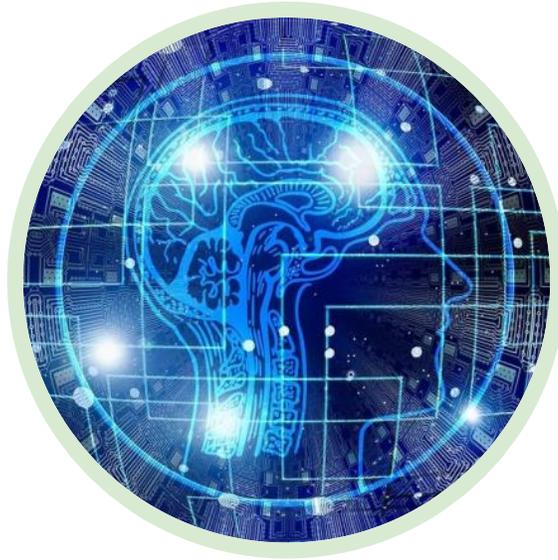
Artificial Intelligence—hype versus hope



Assisted Intelligence

Improves the value of existing activities, such as face recognition at the airport, so that agents who perform immigrations checks can proceed more quickly.

Ex: CAD



Augmented Intelligence

Brings new capabilities to human activity, such as a chess program that suggests possible moves to human players, who then play better than an un-aided human could

Ex: Identifying at-risk populations through EHR data analysis



Autonomous Intelligence

Makes decisions without human input or oversight.

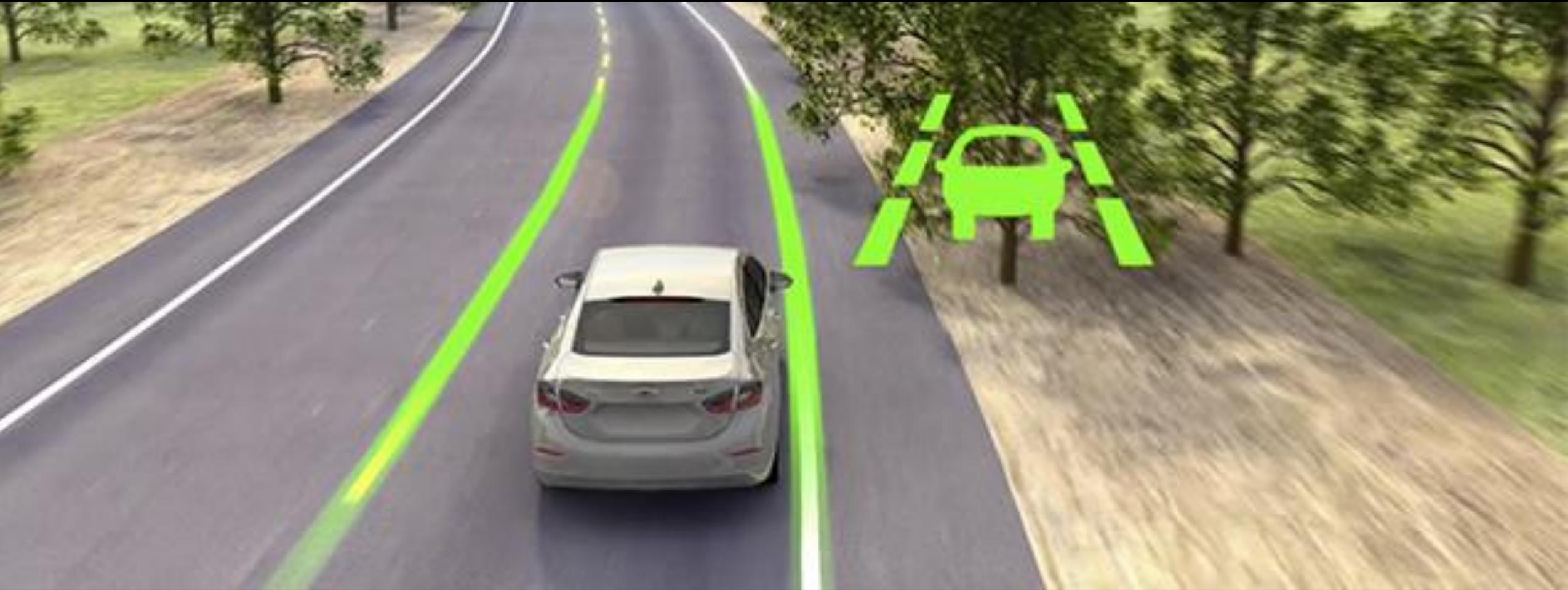
Ex: Full autonomous vehicles.
Fully autonomous trading.

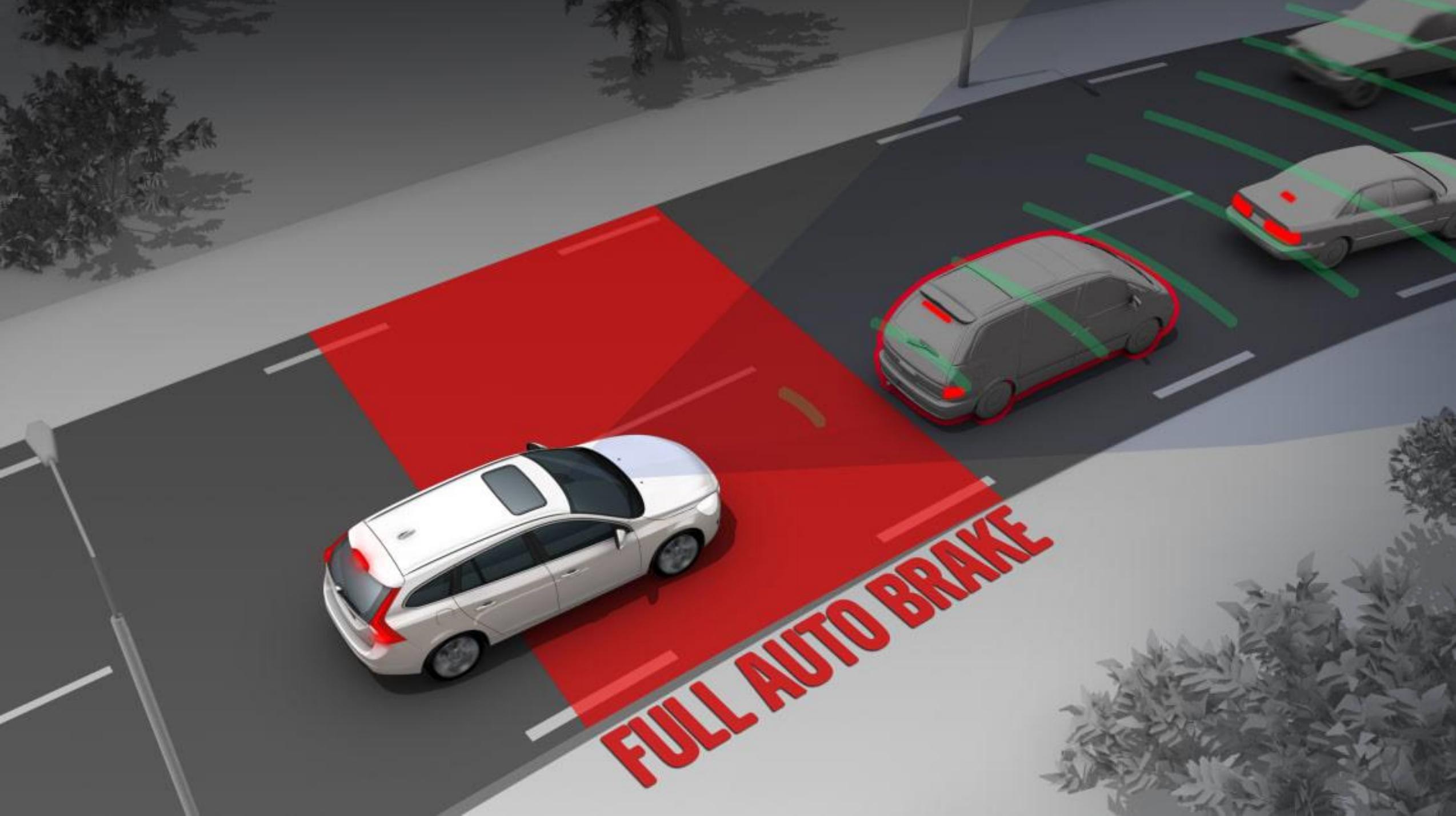
Ask the Audience:
Who here owns a car with...

Blind Spot Detection



Lane Assistance





FULL AUTO BRAKE

Fully Autonomous Driving?





Radiology applications of AI in Quality and Safety

Automation is hard!

So how do we assist and augment our radiologists?

Making Radiology “Better” means either...



Making Radiology “Better” means either...

Increasing productivity

Considerations

- Practice patterns vary between sites and radiologists
- Top 100 most common diagnoses cover only 50% of average radiologist workflow
- Interfaces are proprietary, hard to plug and play solutions.

Increasing quality

Considerations

- Relatively uncommon (3-5% of all scans)
- Time spend on higher quality means less time spent on productivity.
- Sensitive topic to address
- Quality is not always financially justifiable

Sutter Safe Care



Making Radiology “Better” means either...

Increasing productivity

Considerations

- Practice patterns vary between sites and radiologists
- Long tail - 100 most common diagnoses cover only 50% of average radiologist workflow
- Workflow interfaces are proprietary, hard to plug and play solutions.

Increasing quality

Considerations

- Relatively uncommon (3-5% of all scans)
- Time spend on higher quality means less time spent on productivity.
- Sensitive topic to address
- Quality is not always financially justifiable

Diagnostic Errors are a System-level Problem



Healthcare professionals along the patient journey



What are we doing today for Quality Improvement?

Index RADPEER Input Form Review Password Help Logout

RADPEER Input Form

> [Please review the 2 case(s) returned by Administrator]

Your ID:002 Group ID:999

Reviewed Physician ID:

Site:

Modality:

Score:	Clinical Significance
1 Concur with interpretation	<input checked="" type="radio"/>
2 Discrepancy in Interpretation/ not ordinarily expected to be made (understandable miss)	<input type="radio"/> 2a. Unlikely to be significant <input type="radio"/> 2b. Likely to be significant
3 Discrepancy in Interpretation/ should be made most of time	<input type="radio"/> 3a. Unlikely to be significant <input type="radio"/> 3b. Likely to be significant
4 Discrepancy in Interpretation/ should be made almost every time - misinterpretation of findings	<input type="radio"/> 4a. Unlikely to be significant <input type="radio"/> 4b. Likely to be significant

*:Score 2b, 3 or 4 should be reviewed by QA Chair/Committee
for internal peer review before submission to ACR.

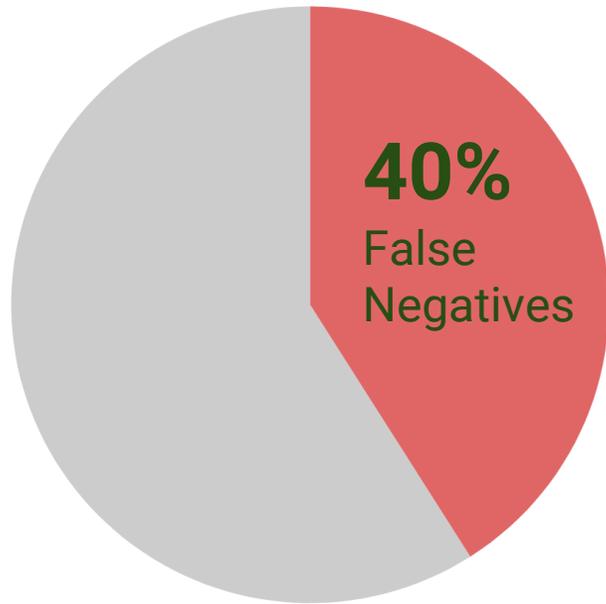
[For Score 2b, 3 or 4]
Has the score been reviewed by Chair/Committee? Yes No N/A
Required explanatory comments:(if Score 2b/3/4 and has not been reviewed. Max. 450 characters)

Radiology applications of AI in Quality and Safety

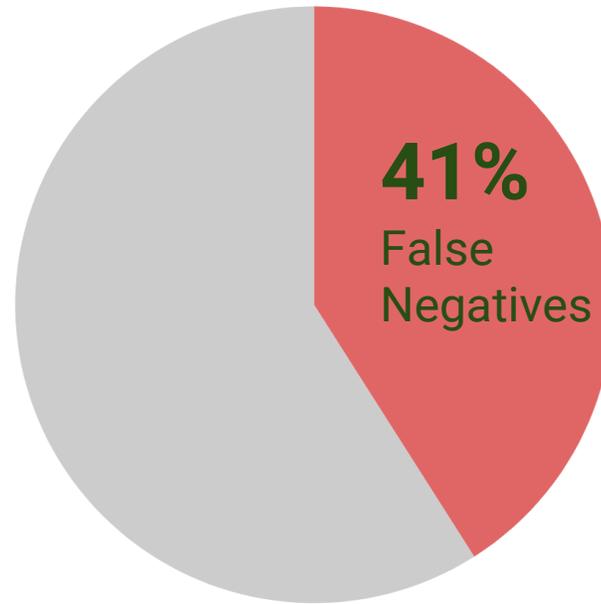
- **Peer Review:** Partial (1-3%), unfocused, and time delayed.
- **Peer Learning:** Opportunistic and focused. Attempts to improve quality.
- **Consequently, medical error is universally **under-reported** and **under-resourced****



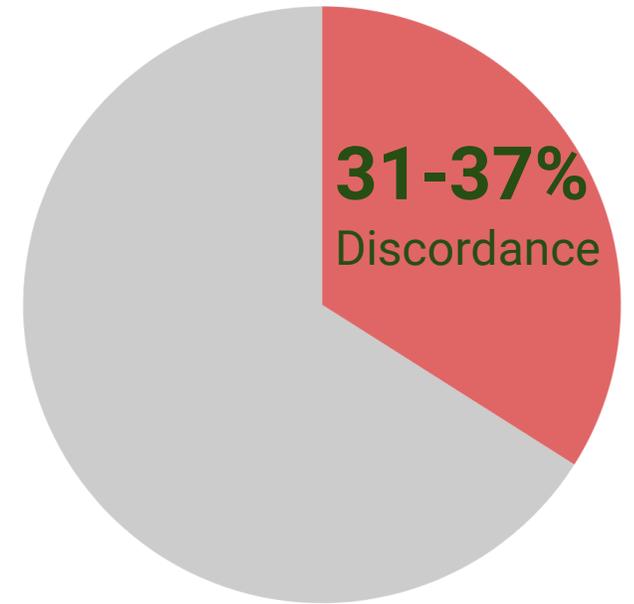
The Literature on Diagnostic Errors



CXR



**Mammograph
y**

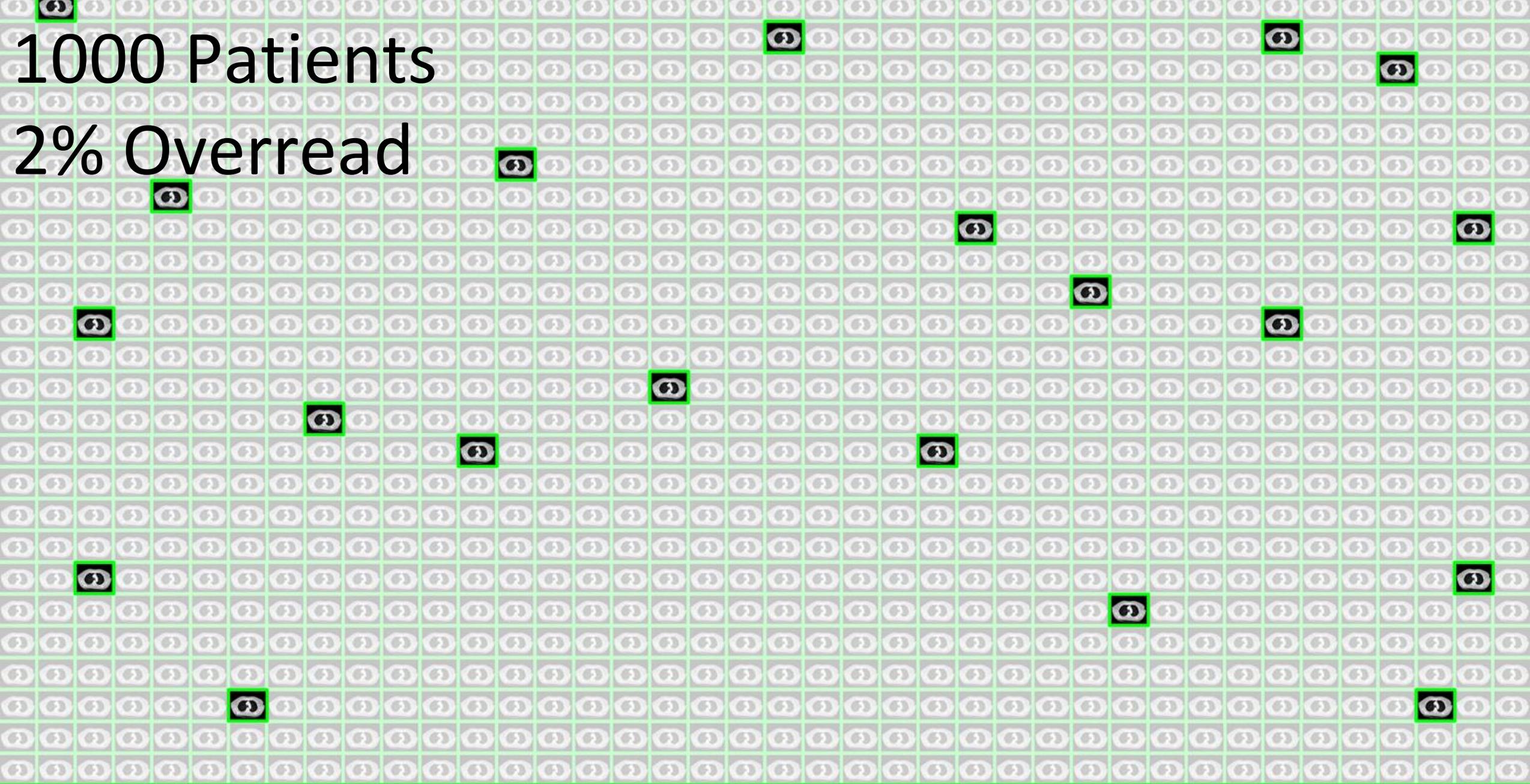


CT

1000 Patients

1000 Patients

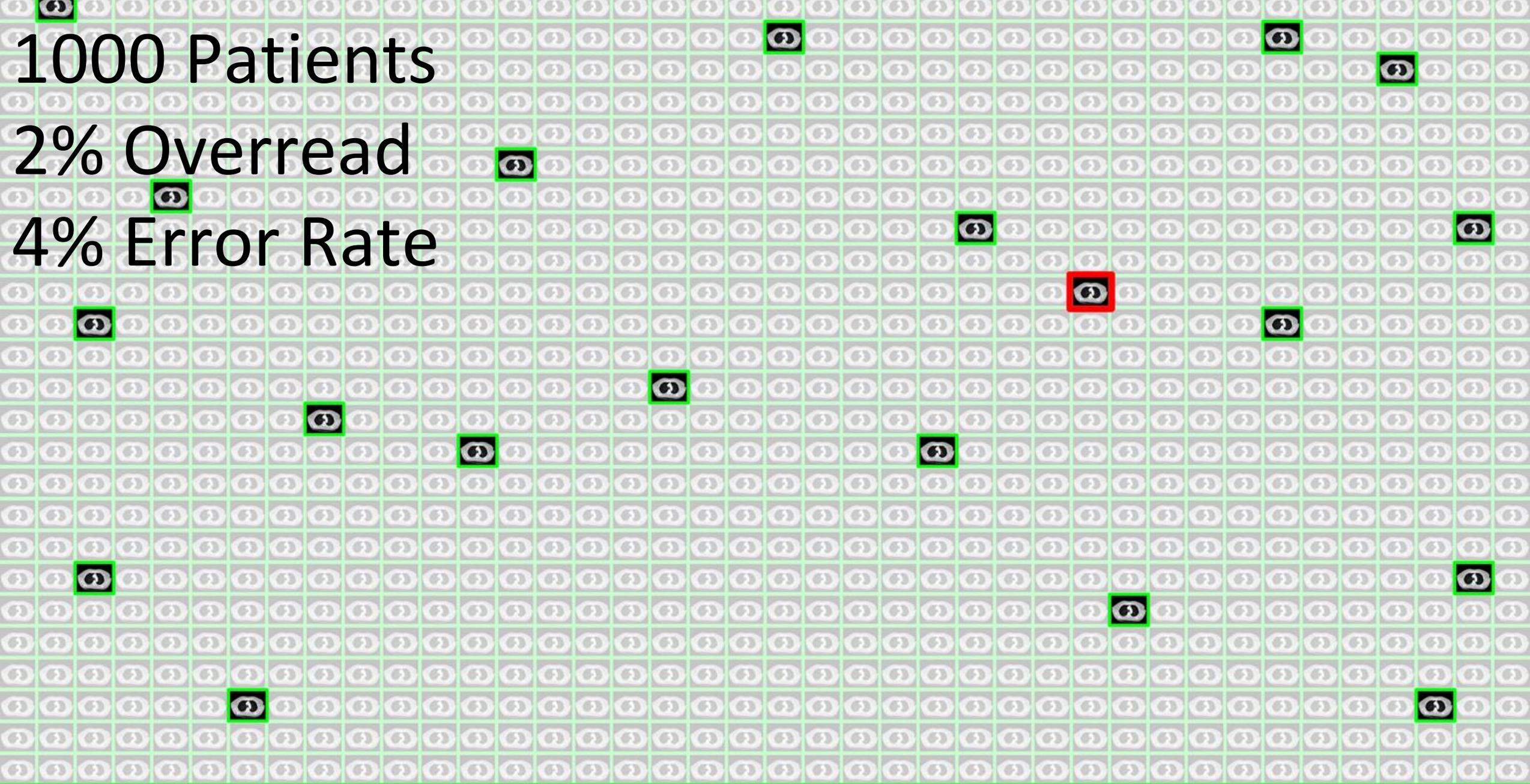
2% Overread

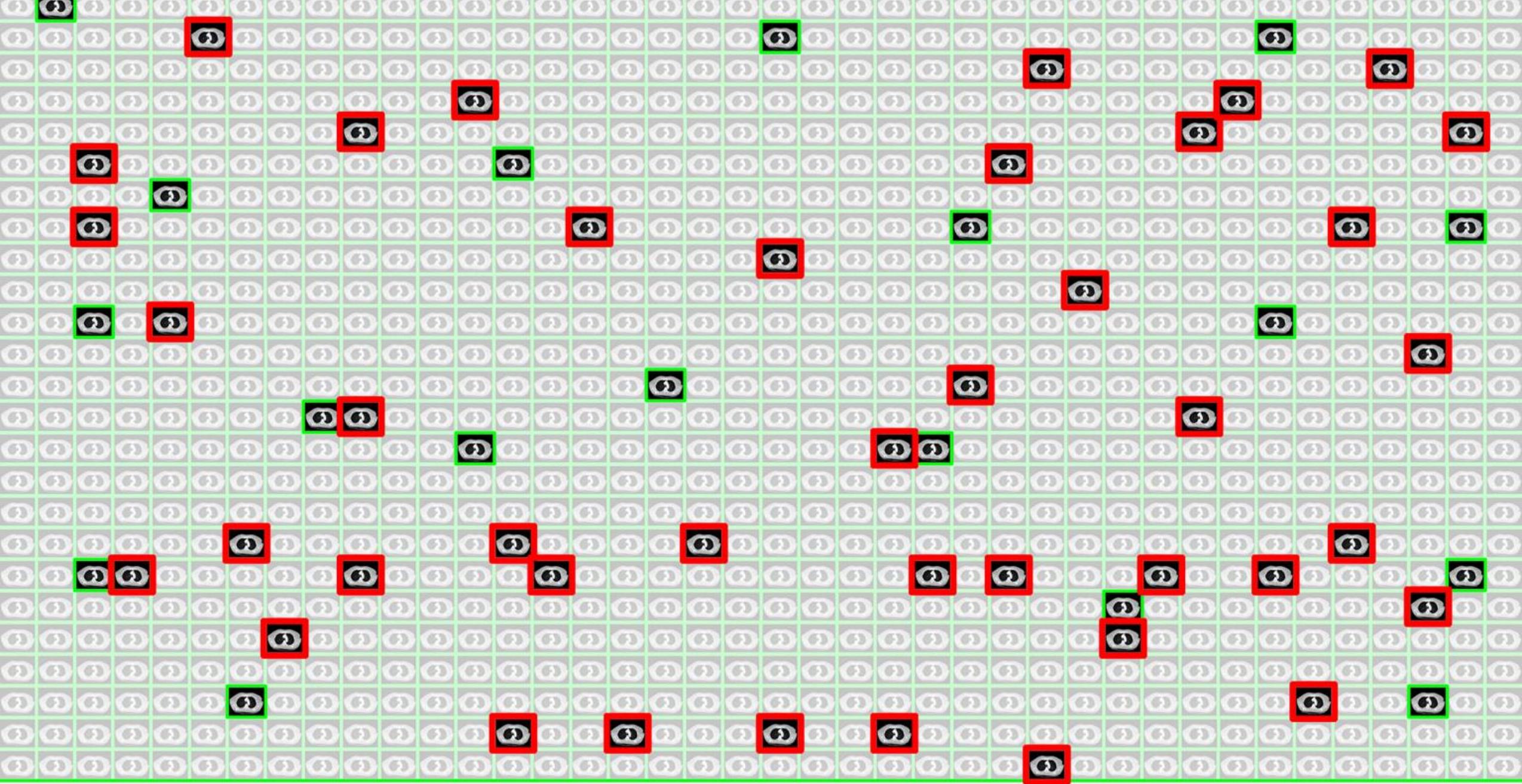


1000 Patients

2% Overread

4% Error Rate





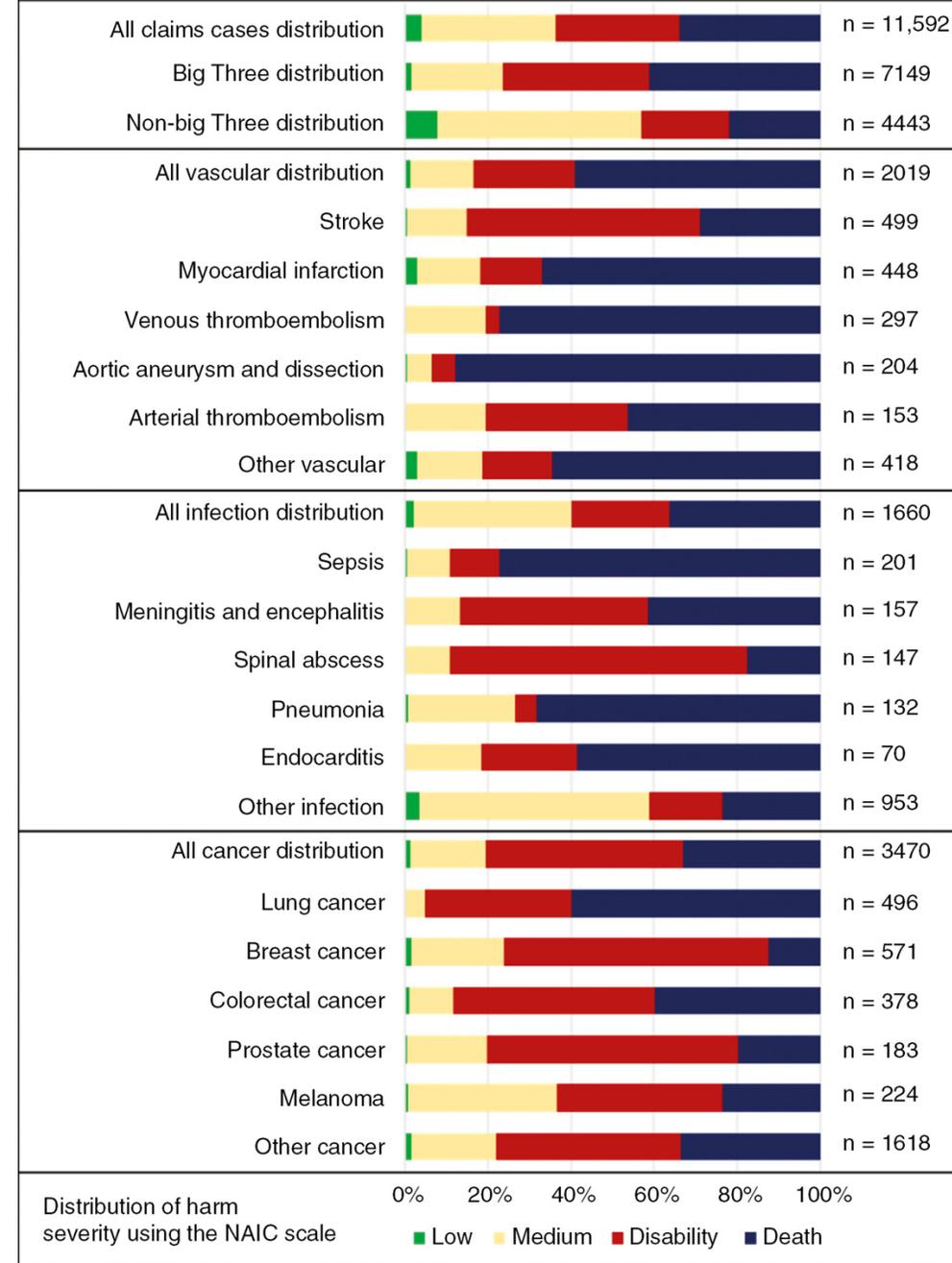


DE GRUYTER

Diagnosis 2019; 6(3): 227–240

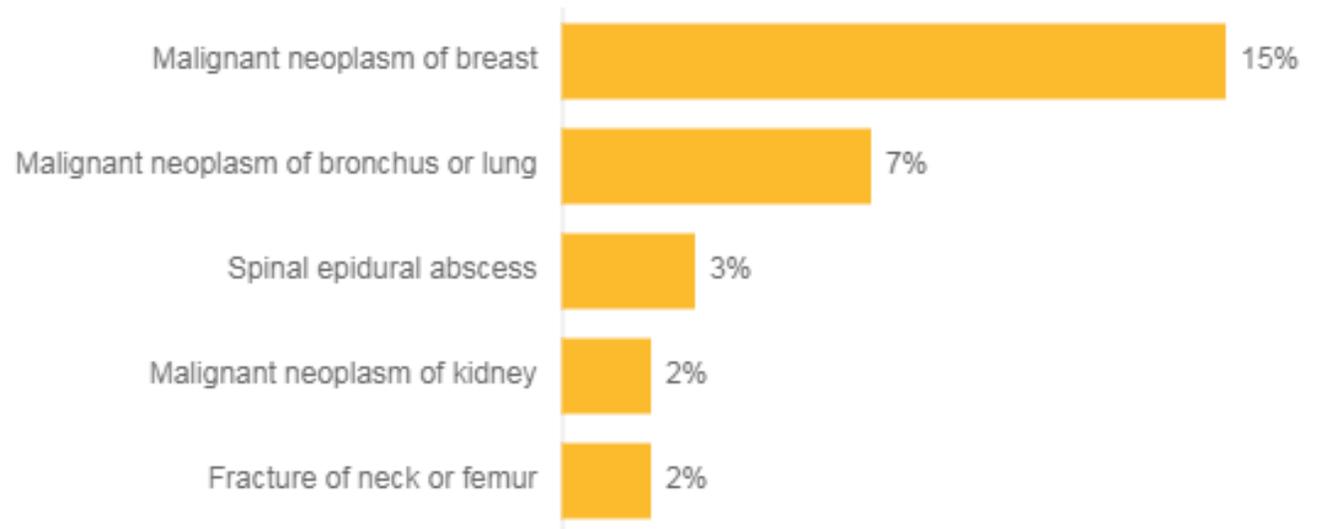
David E. Newman-Toker*, Adam C. Schaffer, C. Winnie Yu-Moe, Najlla Nassery, Ali S. Saber Tehrani, Gwendolyn D. Clemens, Zheyu Wang, Yuxin Zhu, Mehdi Fanai and Dana Siegal*

Serious misdiagnosis-related harms in malpractice claims: The “Big Three” – vascular events, infections, and cancers



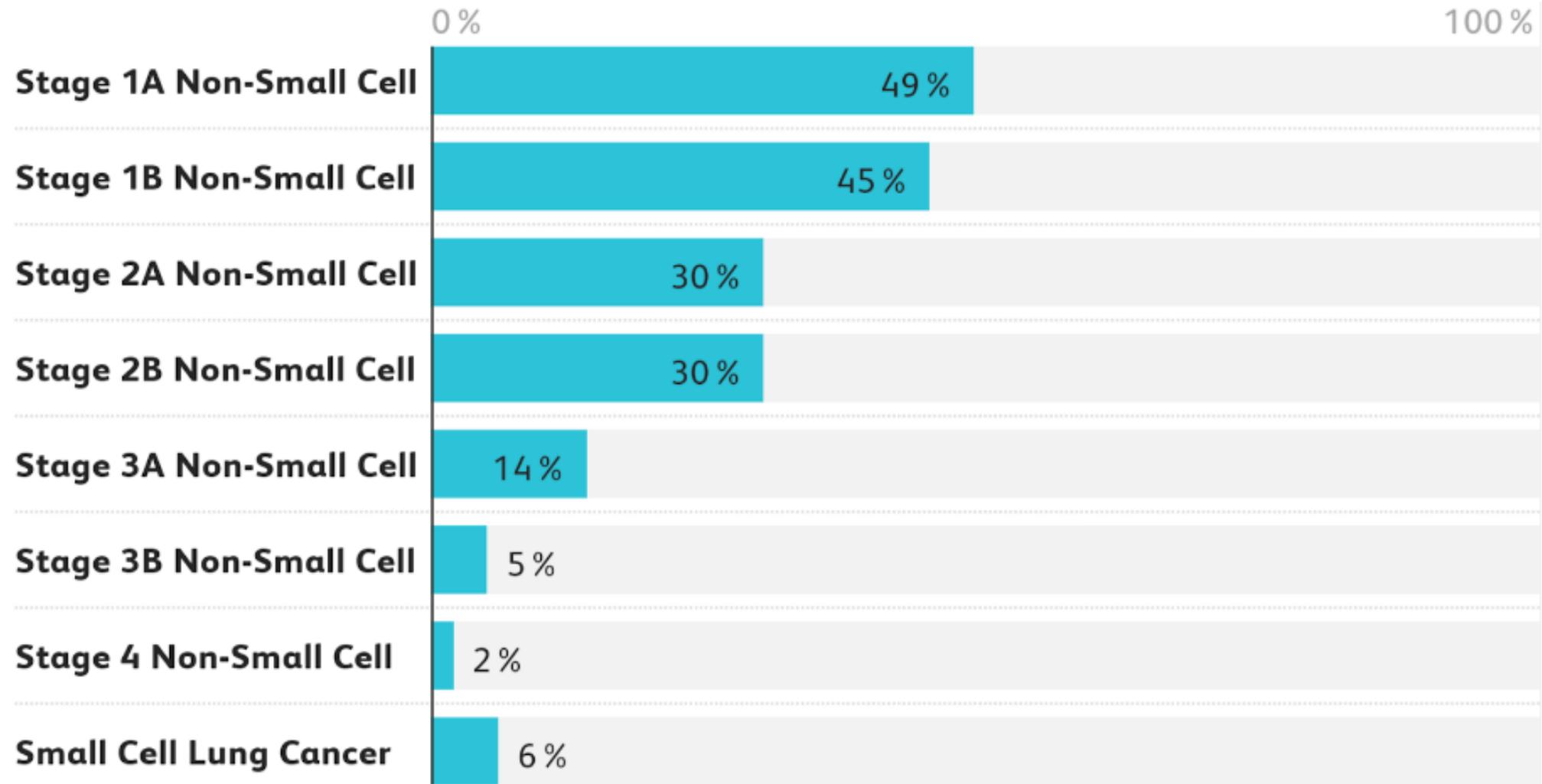


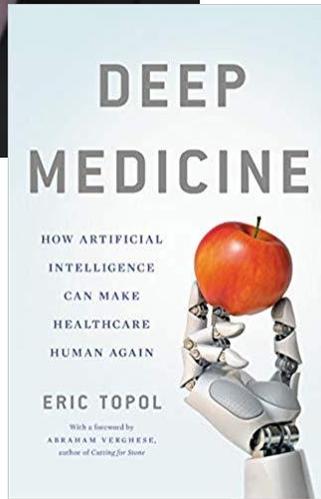
The most common diagnoses associated with diagnosis-related claims (failure, delay, or wrong) were:



<https://www.thedoctors.com/articles/diagnostic-and-interventional-radiology-closed-claims-study/>

Clear clinical impact





“Limits on analysis of data by humans alone have clearly been exceeded, necessitating an increased reliance on machines”.

Topol, Nature 2019.

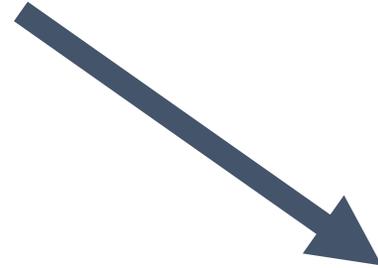
Algorithms are plentiful

Home > Articles > Lung Cancer Among the Most Likely Conditions to Harm Patients who are...

Articles **Must Read Articles**

Lung Cancer Among the Most Likely Conditions to Harm Patients who are Misdiagnosed

By Peter Hofland, Ph.D - July 12, 2019



RSNA 2019 AI Vendors



Guiding Principles

- **No Workflow Impact** – The technology is not intrusive, allowing physicians to practice without interruption and without changing the way they work. Physicians make the decisions and maintain control of the diagnostic process.
- **On-Premises Deployment** – HIPAA compliant, on premises. *PHI never leaves the four walls of the institution and is completely secure.*
- **Vendor-Neutral** – There are 700+ AI companies, 200 founded in the past year, with the market leader changing every quarter for every diagnostic use case.

Guiding Principles

- **Error Reduction = Better Patient Care** – the best predictor of outcomes for critical diagnosis like cancer is early detection.
- **Improve Patient Follow-up** – patients are prompted to return for valuable follow-up care and state within the health system.
- **Risk Mitigation** – the platform helps avoid malpractice suits that lead to financial loss, reputational harm, and physician burnout.

Sutter's Solution

- ✓ **Set** success metrics (KPI's) in collaboration with clinical, administrative, risk/compliance, and operational leaders.
- ✓ **Identify** a well-backed startup that offered a fully vendor-managed on-premises algorithm deployment platform that met our IT requirements.
- ✓ **Implement** machine vision and natural language algorithms in a post-interpretation workflow as a “second read.”
- ✓ **Convene** a group of senior subject matter expert radiologists to be deputized by our quality committee as reviewers.
- ✓ **Develop** a closed loop workflow that ensured rapid resolution of identified findings.
- ✓ **Tracking and Reporting** on the outcomes of the findings generated.

AI Workflow:



Report dictated



Scans reviewed by AI: Use machine learning and computer vision analysis to detect lesions



Natural Language Processing to analyze reports for findings and match to image

SME review of any studies flagged with a discrepancy

Addendum and follow-up care

Finding Info

Accession Number: RD9127471118

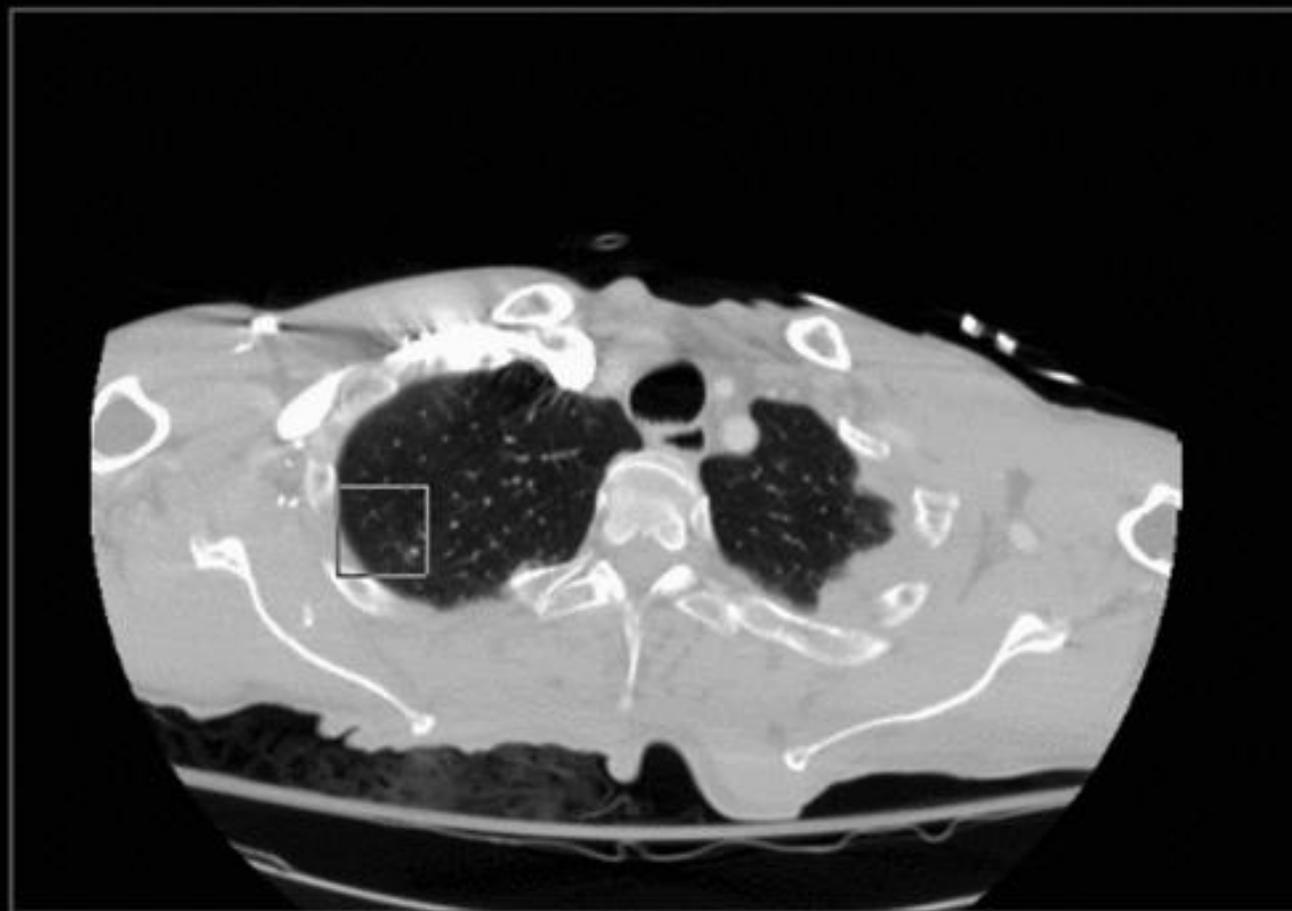
Study description: CT Abdomen

Study date: 2018-09-26

Series Number: 847

Finding:

- Diameter: 7.7594337 mm



Axial Slice

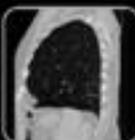
Coronal Slice

Sagittal Slice

Axial Animation

Coronal Animation

Sagittal Animation



Report

A 3.6 cm ill-defined early enhancing lesion in S6 of liver (Ser/Img: 3/25), with nodular enhancement in arterial phase and centripetal enhancement in portal venous phase. Hemangioma was suggested. **Some poor enhanced areas in the lesion, maybe due to partial thrombosis.** A 7 mm early enhancing lesion in S7 of liver (Ser/Img: 3/16), with persistent enhancement in portal venous phase. Hemangioma was suggested. A tiny ill-defined early enhancing lesion in S4 of liver (Ser/Img: 3/11), with isodensity in porto venous phase, favor AP-shunting. **Fatty infiltration of liver.** A stone in left kidney.

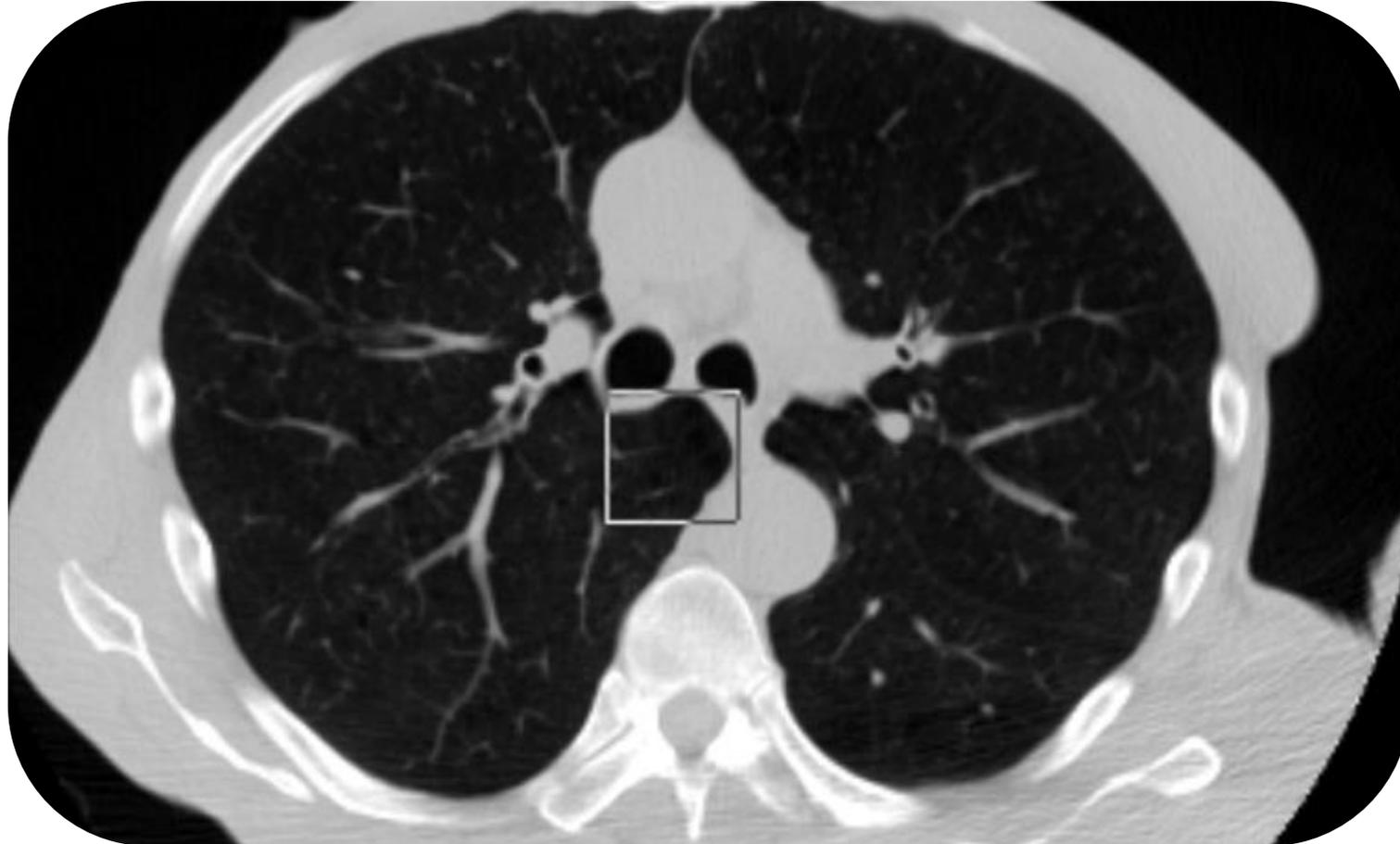
No mass or nodular lesion in the visible lung. No abnormal finding in the spleen, pancreas, gallbladder and bilateral adrenal glands. No definite evidence of enlarged lymph nodes was noted at paraaortic retroperitoneum. Fatty infiltration of liver. A stone in left kidney.

This image viewer is not approved for diagnostic use. Image flags are limited by Federal law to investigational use only. Ferrum algorithms are for investigational use only.

©2019 Ferrum Health Inc.

Cases

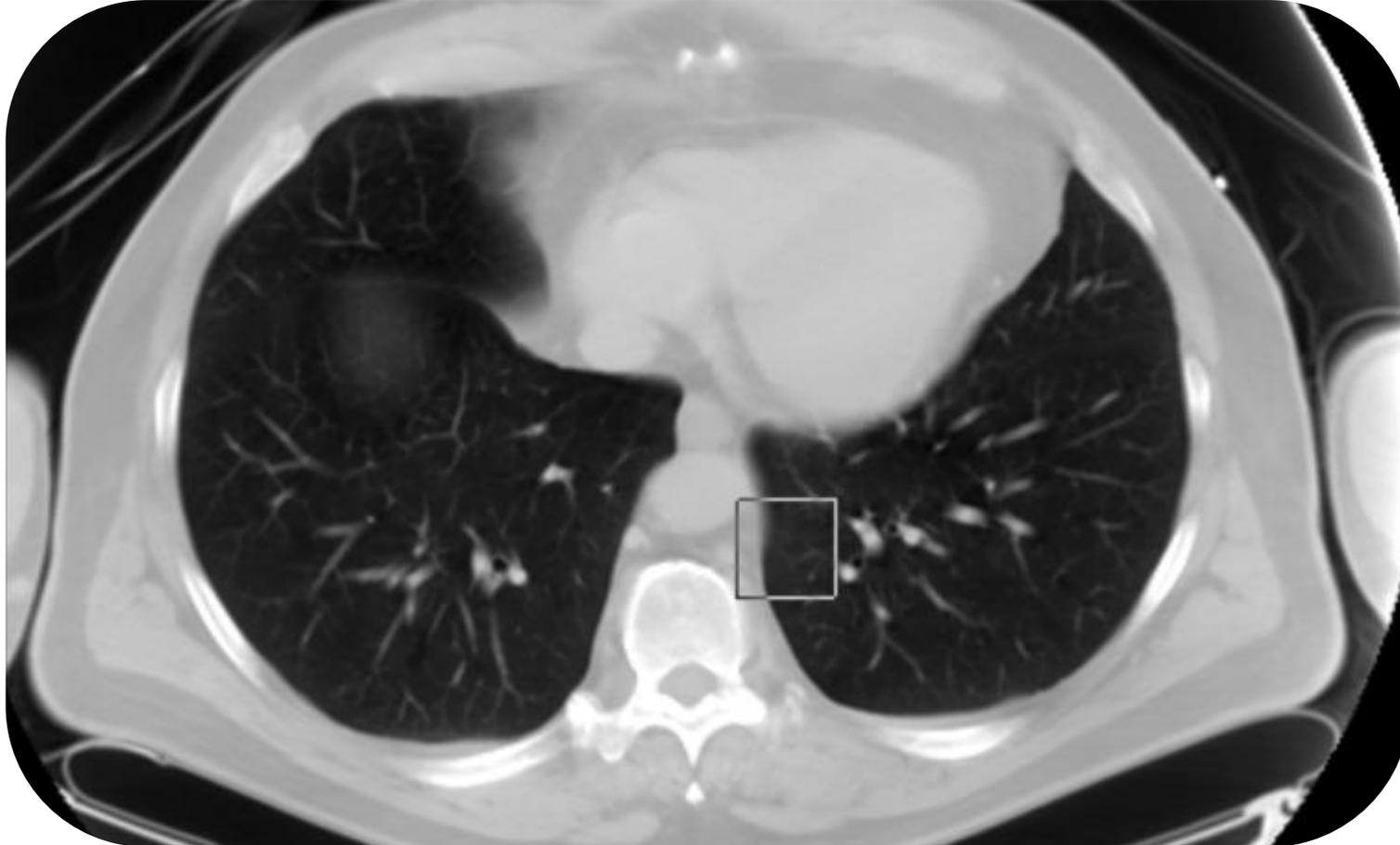
Sutter experience



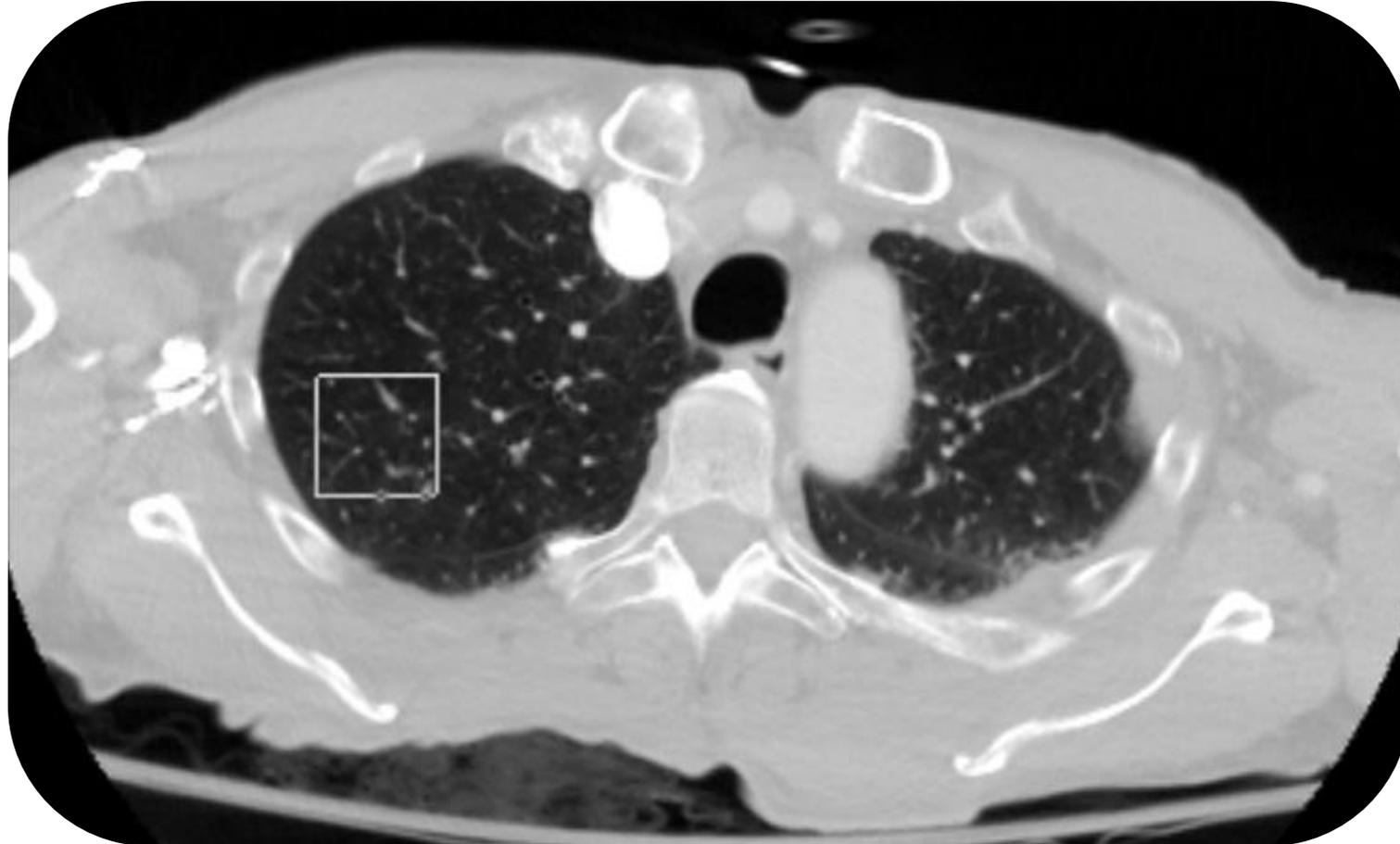
Sutter experience



Sutter experience



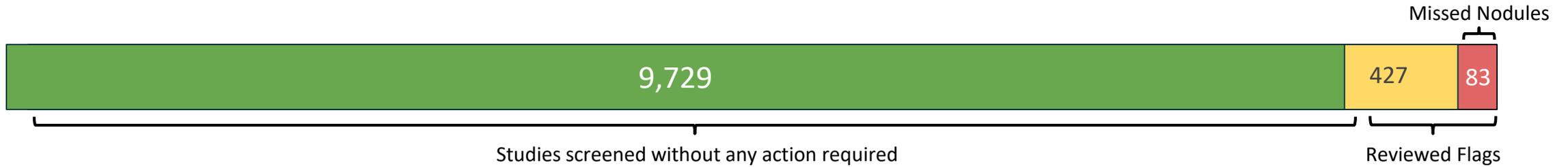
Sutter experience



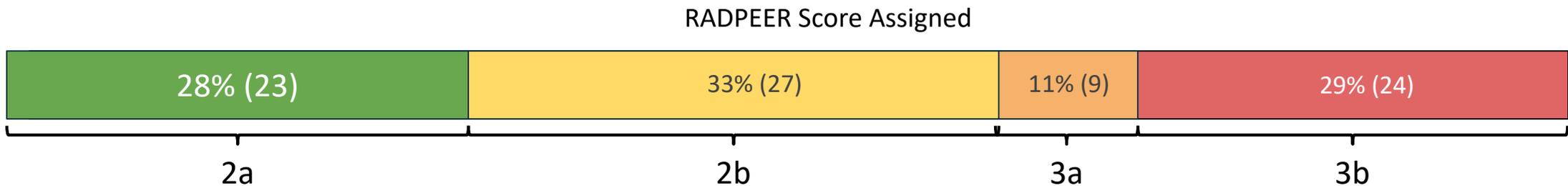
Sutter experience - 90 Days

Data from 8/11 to 11/20

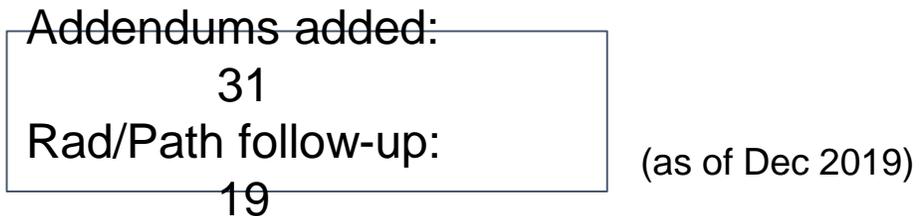
All Studies:

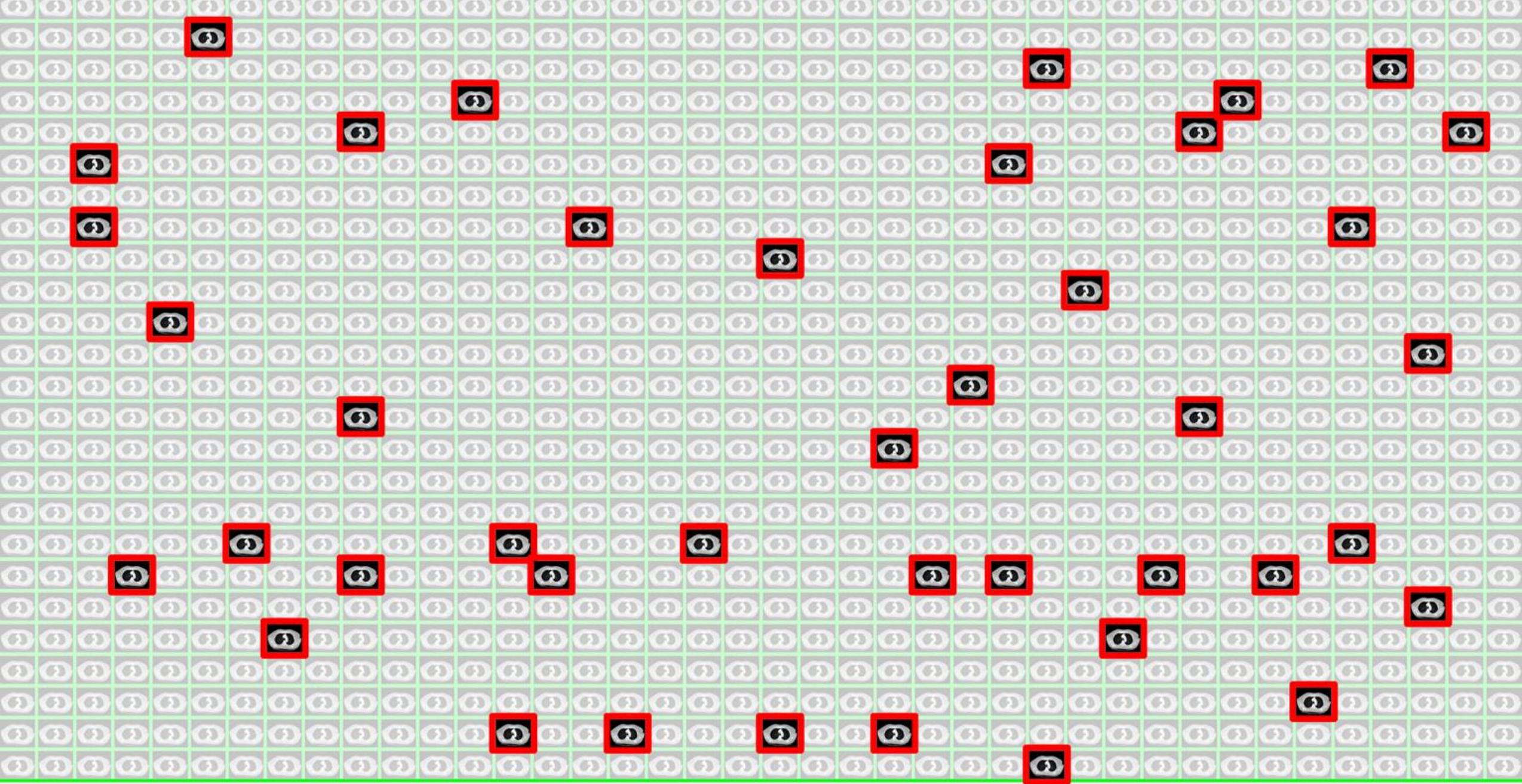


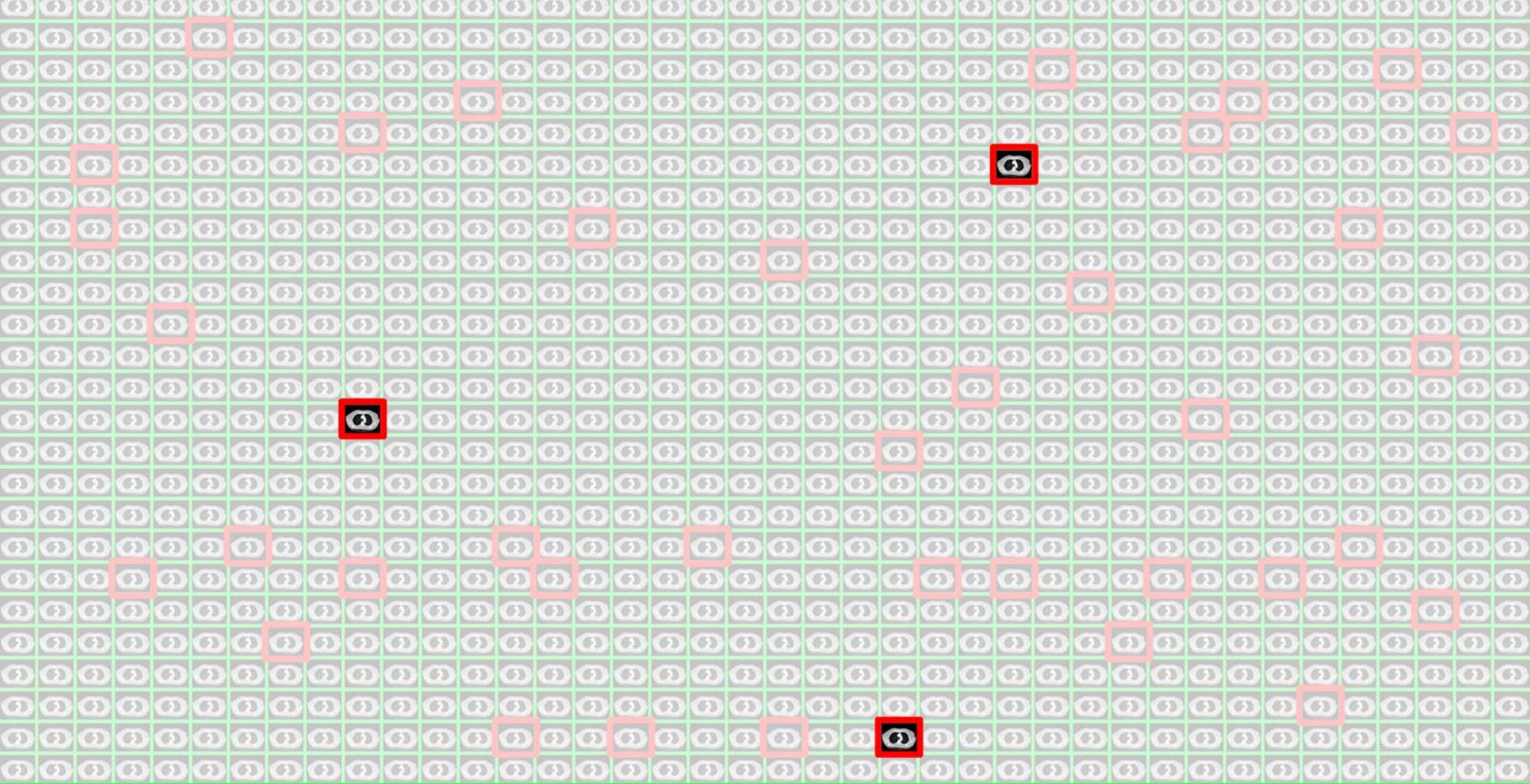
Radiologist confirmed missed nodules:



Changes in downstream care:









Proud to have the most comprehensive cancer diagnosis quality program in Sacramento.



Sutter experience

Summary:

- We conclude that the use of AI tools in peer review is of great benefit to peer review efficiency as well as to the detection of clinically significant missed findings in a large patient population.
- As AI tools proliferate and the coverage of pathologies and modalities increases, the impact of AI aided peer review and quality assurance will result in a dramatic reduction in medical error in radiology.

Use of Post-Interpretation AI Tools to Catch Potential Errors Before Clinical Impact

Other Key Learnings:

Use of Post-Interpretation AI Tools to Catch Potential Errors Before Clinical Impact

Key Learnings:

- Data
- *Partnership* with vendors
 - Open platform
 - Processing power
 - Talent
- Leadership willingness and “buy in”

Use of Post-Interpretation AI Tools to Catch Potential Errors Before Clinical Impact

Key Learnings:

- Data:
 - Sutter EHR Imaging Archive:
 - How many Images are in the VNA today?
 - **>3.5 Billion Images**
 - How many patients exist in the VNA as of 12/2019?
 - **>13 million unique patients**
 - How much storage has been utilized?
 - **3.0 PB**
 - Data security!
 - On-premises storage, processing



Use of Post-Interpretation AI Tools to Catch Potential Errors Before Clinical Impact

Key Learnings:

- *Partnership with Vendor*

Use of Post-Interpretation AI Tools to Catch Potential Errors Before Clinical Impact

Key Learnings:

- *SH Leadership Buy-in*

“Leading Change: Why Transformation Efforts Fail”,
John P. Kotter HBR Jan. 2007



Radiology applications of AI in Quality and Safety—Sutter Health experience

Key learnings:

- Explainability: the level to which the AI's internal mechanics are explainable in human terms
 - Critical to adoption and success in healthcare
 - The human element: “black box AI” is not sufficient in healthcare
- 2019 Article in the Journal of Consumer Research:
 - consumers are reluctant to utilize healthcare provided by AI
 - consumers exhibit lower reservation prices for healthcare provided by AI
 - Derive negative utility if a provider is automated rather than human

Radiology applications of AI in Quality and Safety—Sutter Health experience

Key learnings 2:

- User acceptance
 - Disillusionment: *we still miss nodules*
- What to do with the False Positives?
 - 2 hours/week of non-interpretive time!
 - Will improve over time

Radiology applications of AI in Quality and Safety—Sutter Health experience

What about liability concerns for Radiologists that use AI?

- No case law on liability involving medical AI
- “because current law shields physicians from liability as long as they follow the standard of care, the “safest” way to use medical AI from a liability perspective is as a confirmatory tool to support existing decision-making processes, rather than as a source of ways to improve care.”

Price WN, Gerke S, Cohen IG. Potential Liability for Physicians Using Artificial Intelligence. *JAMA*. 2019;322(18):1765–1766. doi:10.1001/jama.2019.15064



Quality will be how we move our specialty forward

- **Error Reduction = Better Patient Care** – the best predictor of outcomes for critical diagnosis like cancer is early detection.
- **Improved Patient Follow-up/Retention** – patients are prompted to return for valuable follow-up care and stay within the health system.
- **Risk Mitigation** – “Safety net” AI tool helps avoid malpractice suits that lead to financial loss, reputational harm, and physician burnout.
- **Brand and Market Share** – negative publicity is driven by missed diagnoses and malpractice. We can differentiate on quality.
- **Transition to Value** – improved quality ultimately positions us better for payer negotiations and government incentives under increasingly common risk-based contracting.

Competing on cost/productivity vs quality will be a race to the bottom



© marketoonist.com

“It is important to take on the challenge of identifying success measures for AI systems by their impact on people’s lives”

-Barbara Grosz, Higgins Professor of Natural Sciences at Harvard University

References

- Panetta K. 5 Trends appear on the Gartner Hype Cycle for Emerging Technologies”. Retrieved from:
<https://www.gartner.com/smarterwithgartner/5-trends-appear-on-the-gartner-hype-cycle-for-emerging-technologies-2019/>
- Rao, A.S., & Verweij, G. (2017). Sizing the prize. What’s the real value of AI for your business and how can you capitalize?” Retrieved from:
<https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html>
- Bloomberg, 2018. “Think you know how disruptive artificial intelligence is? Think again.” Retrieved from:
<https://www.forbes.com/sites/jasonbloomberg/2018/07/07/think-you-know-how-disruptive-artificial-intelligence-is-think-again/#23ea48d13c90>

References

- The Doctors Company: Diagnostic and Interventional Radiology Closed Claims Study. Retrieved from: <https://www.thedoctors.com/articles/diagnostic-and-interventional-radiology-closed-claims-study/>
- [The frequency of diagnostic errors in outpatient care: estimations from three large observational studies involving US adult populations.](#) [H Singh](#), [AND Meyer](#), [EJ Thomas](#) - BMJ Qual Saf, 2014 - qualitysafety.bmj.com
- Brady A, Laoide RÓ, McCarthy P, McDermott R. Discrepancy and error in radiology: concepts, causes and consequences. *Ulster Med J.* 2012;81(1):3–9.\
- [Cognitive and System Factors Contributing to Diagnostic Errors in Radiology](#)
[Cindy S. Lee](#), [Paul G. Nagy](#), [Sallie J. Weaver](#), and [David E. Newman-Toker](#) *American Journal of Roentgenology* 2013 201:3, 611-617
- 10. [AJR Am J Roentgenol.](#) 2017 Apr;208(4):739-749. doi: 10.2214/AJR.16.16963. Epub 2016 Dec 27. **Interpretive Error in Radiology.**
- 11. [Systemic Error in Radiology.](#) [Stephen Waite](#), [Jinel Moore Scott](#), [Alan Legasto](#), [Srinivas Kolla](#), [Brian Gale](#), and [Elizabeth A. Krupinski](#) *American Journal of Roentgenology* 2017 209:3, 629-639