

How to evaluate AI in the context of Patient Safety

Ruth Bergman, PhD



GE HealthCare

First do no harm

Patient safety

11 September 2023

[World Health Organization](#)

Key facts

- Around 1 in every 10 patients is harmed in health care and more than 3 million deaths occur annually due to unsafe care. In low-to-middle income countries, as many as 4 in 100 people die from unsafe care (1).
- Above 50% of harm (1 in every 20 patients) is preventable; half of this harm is attributed to medications (2,3).
- Some estimates suggest that as many as 4 in 10 patients are harmed in primary and ambulatory settings, while up to 80% (23.6–85%) of this harm can be avoided (4).
- **Common adverse events that may result in avoidable patient harm are medication errors, unsafe surgical procedures, health care-associated infections, diagnostic errors, patient falls, pressure ulcers, patient misidentification, unsafe blood transfusion and venous thromboembolism.**
- Patient harm potentially reduces global economic growth by 0.7% a year. On a global scale, the indirect cost of harm amounts to trillions of US dollars each year (1).
- Investment in reducing patient harm can lead to significant financial savings, and more importantly better patient outcomes (5). An example of a good return on investment is patient engagement, which, if done well, can reduce the burden of harm by up to 15% (4).

To vaccinate or not to vaccinate?

1997 Recommendation:
Do not vaccinate for chicken pox



1998 Recommendation:
Vaccinate



Confused mother asks – what happened?
Dr: We have seen increased incidence of death due to infected pox

The data:
Chickenpox used to be very common in the United States. In the early 1990s, more than 4 million people got chickenpox, 10,500 to 13,000 were hospitalized, and 100 to 150 died each year. Chickenpox vaccine became available in the United States in 1995. During the first 25 years of the U.S. chickenpox vaccination program, the vaccine has prevented an estimated 91 million cases, 238,000 hospitalizations, and 2,000 deaths.

CDC recommends two doses of chickenpox vaccine for children, adolescents, and adults who have never had chickenpox and were never vaccinated. Children are routinely recommended to receive the first dose at 12 through 15 months of age and the second dose at 4 through 6 years of age.

<https://www.cdc.gov/vaccines/vpd/varicella/public/index.html>

Risk of AI



Assessing risk of AI

Business Risk

Violation of user privacy

Financial loss

Loss of data

Loss of reputation

Risk to Patient Safety

Violation of patient privacy

Confuse or misguide clinical decision making

Incorrect or missing information

Data for the wrong patient

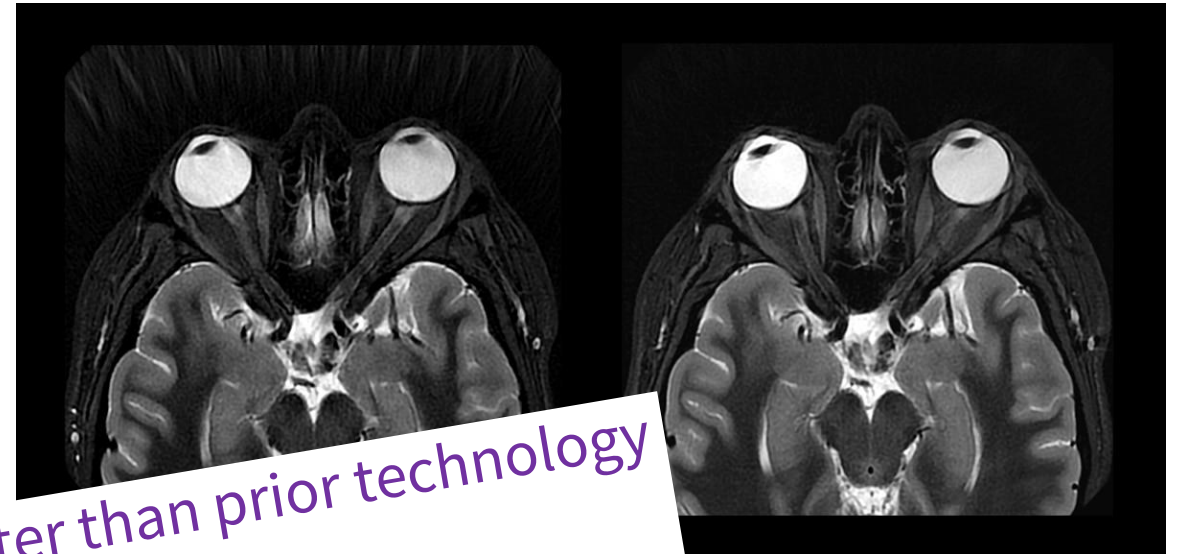
Does it add or reduce risk
to patient safety?

Do AI errors always increase risk to patient safety?



GE HealthCare

GE Healthcare's AIR™ Recon DL* technology



Conventional (1:59 min.) vs. AIR™ Recon DL (3:33 min)

Conventional (1:59 min.) vs. AIR™ Recon (3:33 min)

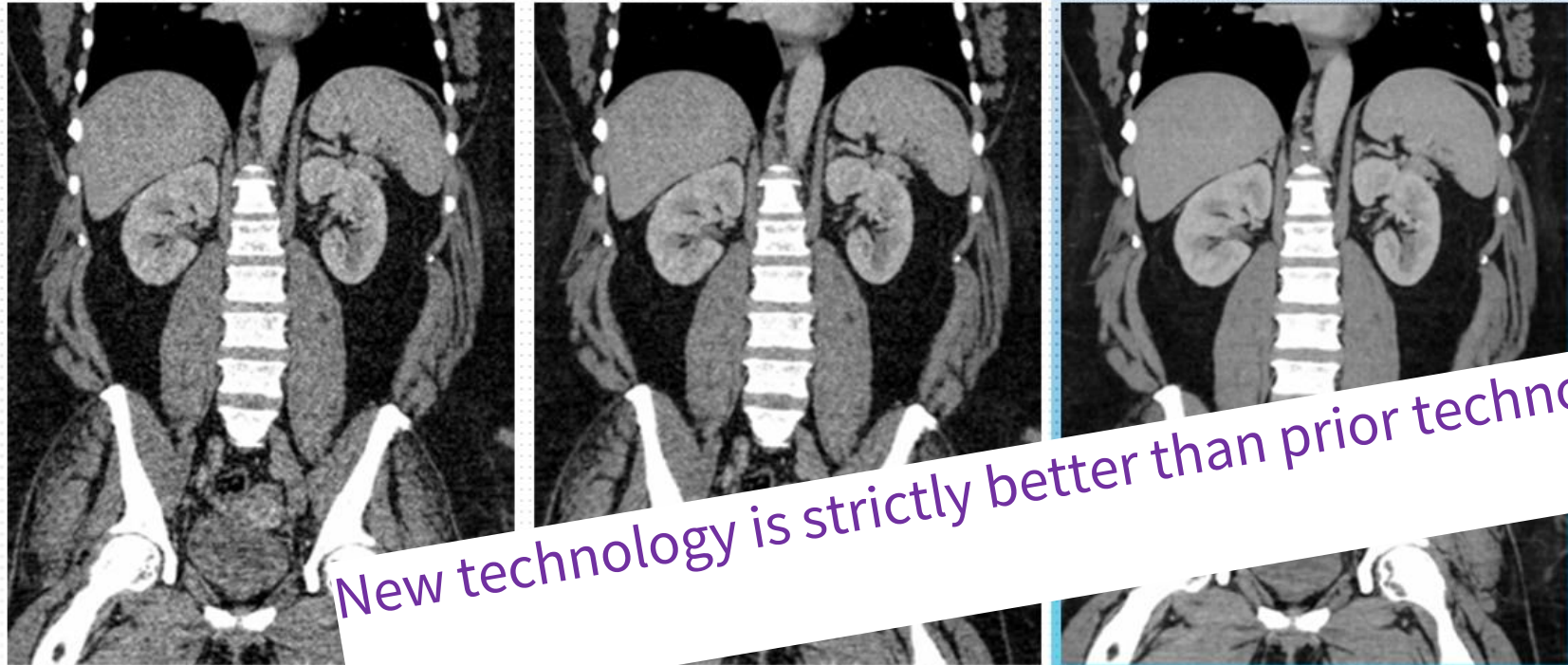
New technology is strictly better than prior technology

- Exceptional image quality with shorter scan times
- Improve SNR and image sharpness
- Prevent artifacts, see true-to-MR detail

"...We can drastically increase the contrast-to-noise ratio, which really improves lesion conspicuity. It allows us to push protocols or applications to levels that would otherwise be incompatible with conventional reconstruction and works in any anatomy."

<https://www.gehealthcare.com/article/intelligently-efficient-with-deep-learning-reconstruction-in-mri>

TrueFidelity CT Images



New technology is strictly better than prior technology

- Low-dose
- Improved image quality
- Reduced noise
- Fewer artifacts

Filtered Back
Projection
1972-2008



Iterative
Reconstruction
2008-2018

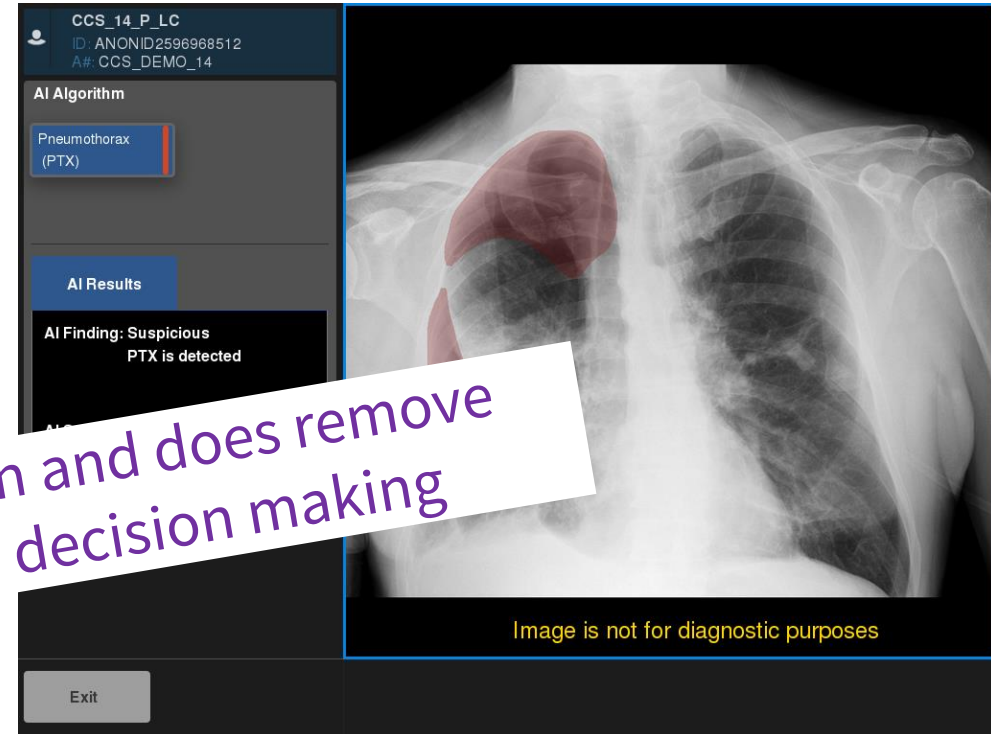
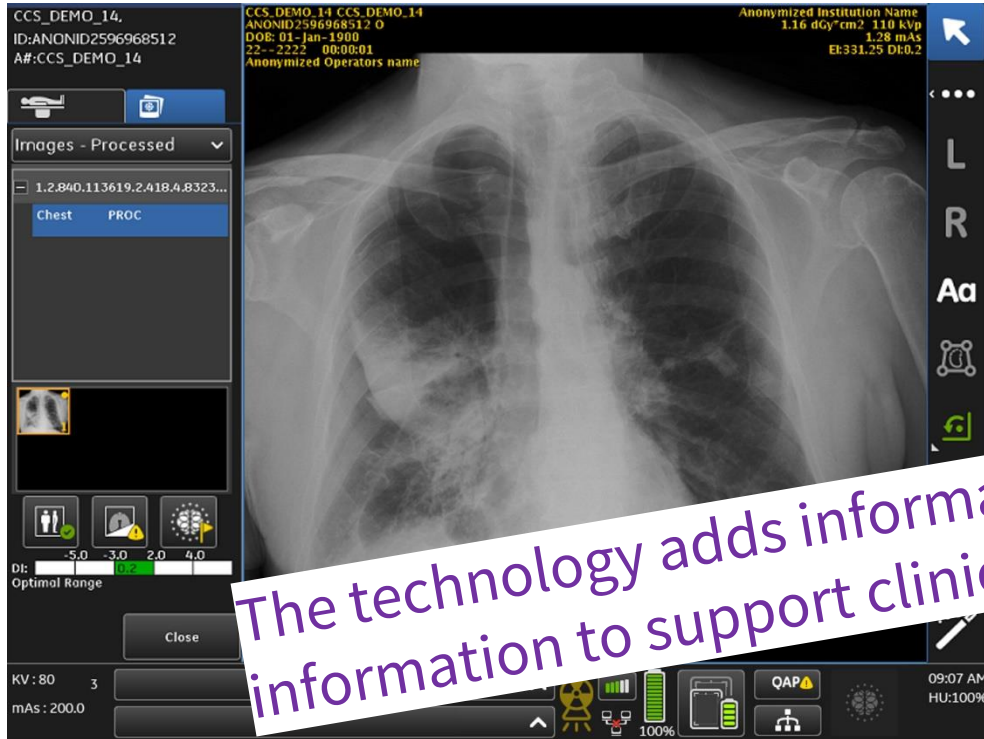


Deep Learning
Image Reconstruction
2018-Future



TRUEFIDELITY IMAGE

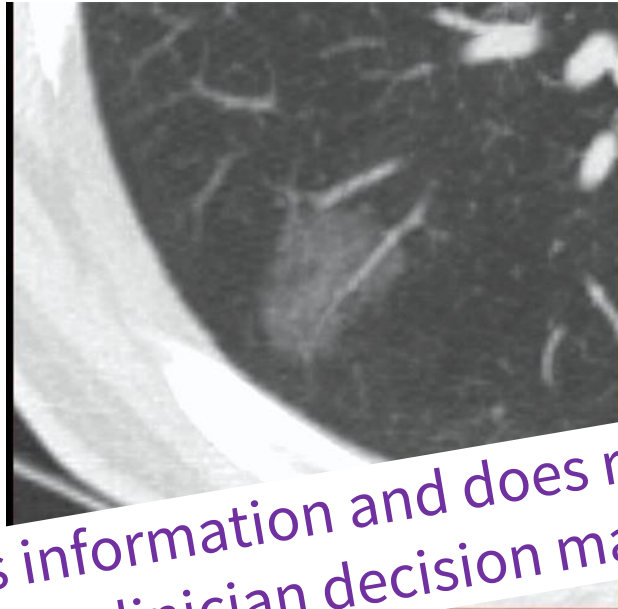
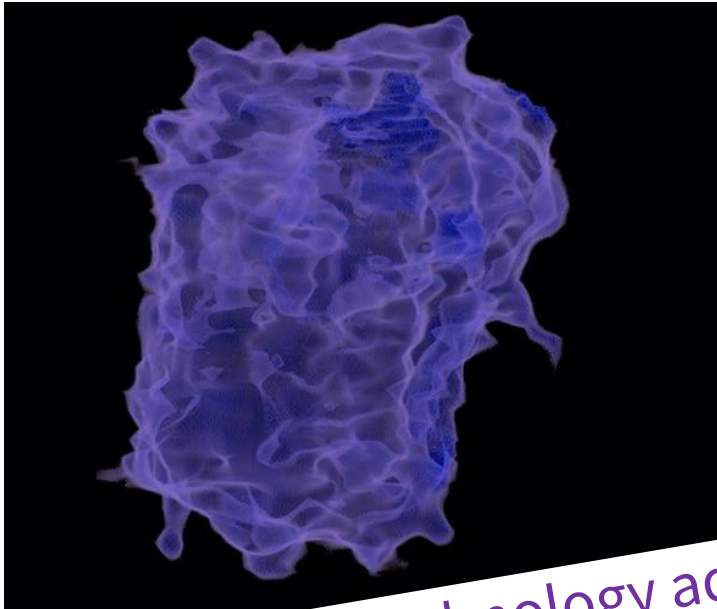
Triaging and prioritizing urgent cases in X-Ray



- Detects nearly all large pneumothoraces (96% sensitivity).
- Identifies 3 out of 4 small pneumothoraces (75% sensitivity).
- Limits false alerts (94% specificity).
- An Area Under Curve (AUC) of 0.96.

- Sends a secondary capture DICOM image to PACS and presents the AI results to the radiologist
- Image flags help enable worklist prioritization and have the potential to expedite review of critical findings

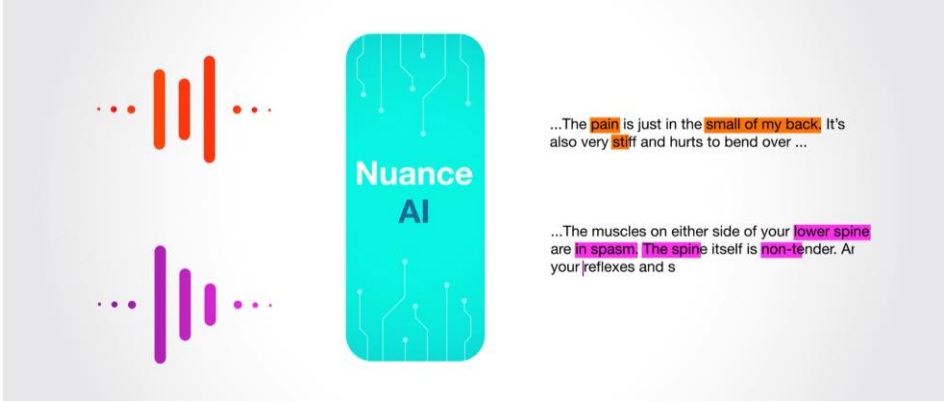
Lung VCAR



The technology adds information and does remove information to support clinician decision making

- Synchronized CT , $CBCT$ and segmentation analysis.
- Automatic nodule visualization.
- Automatic nodule analysis (volume, doubling time, % growth).

Dragon Ambient eXperience by Nuance

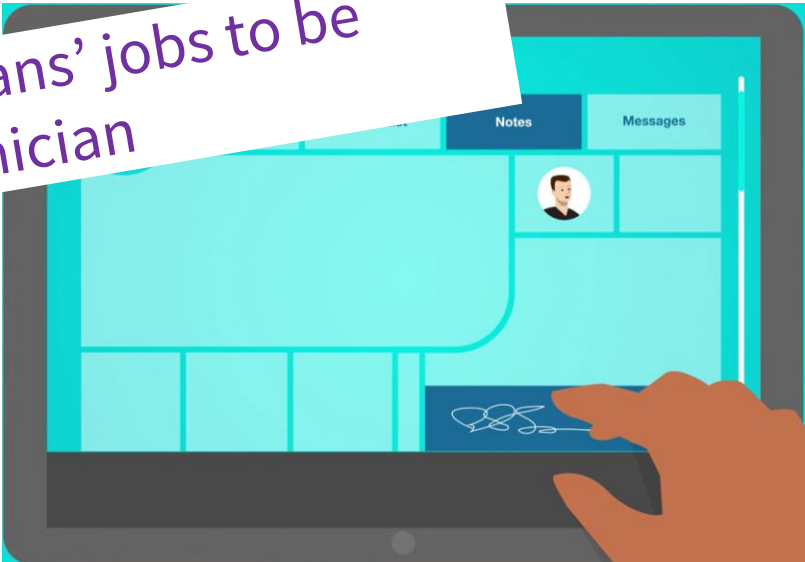


The physician- patient conversation is recorded and securely streamed to the Nuance AI platform, where the speech is diarized, automatically translated into a clinical note, and entered directly into the EHR as structured data. The draft note goes through a quality review process to ensure accuracy before it's delivered to the physician through the EHR for final review and signature.

<https://videos.nuance.com/healthcare/watch/VUbqh7UJZnhfw94KBfunBc?>



The technology assists with clinicians' jobs to be done and does not replace the clinician



Do AI errors always increase risk to patient safety?

Not necessarily, if we can validate that

- New technology is strictly better than prior technology
- The technology adds information and does not remove information to support clinician decision making
- The technology assists with clinicians' jobs to be done and does not replace the clinician
- The technology does not substitute existing clinical workflow

Demonstrate you have reduced the risk as far as possible

Role of the regulator



GE HealthCare

Sanofi Pasteur 450/477 Fluzone® Quadrivalent

Efficacy

Table 8: Estimated Efficacy of Fluzone (Trivalent Influenza Vaccine) Against Culture-Confirmed Influenza in Children Aged 6 through 24 Months during the 1999-2000 and 2000-2001 Influenza Seasons – Intent-to-Treat Analysis Set^a

Year	Fluzone ^b				Placebo ^c				Fluzone vs. Placebo	
	n ^d	N ^e	Rate (n/N) ^f	(95% CI)						
Year 1 ^b (1999-2000)	15	273	5.5	(3.1; 8.9)						
Year 2 ^l (2000-2001)	9	252	3.6	(1.6; 6.7)	4	123	3.3	(0.9; 8.1)	1.10 (0.34; 3.50)	-10 (-250; 66)

Risk - adverse reactions

Table 2: Study 1^a: Percentage of Solicited Inj... ion-site and Systemic Adverse Reactions Within 7 Days After Vaccin... Months Through 35 Months of Age (Safety Analysis Set)^{a,b}

	Any (%)	Grade 2 ^h (%)	Grade 3 ⁱ (%)	TIV-1 ^{d,e} (B Victoria) (N ^g =310)			TIV-2 ^{d,f} (B Yamagata) (N ^g =308)		
				Any (%)	Grade 2 ^h (%)	Grade 3 ⁱ (%)	Any (%)	Grade 2 ^h (%)	Grade 3 ⁱ (%)
Injection-site adverse reactions									
Pain^j	57.0	10.2	1.0	52.3	11.5	0.8	50.3	5.4	2.7
Tenderness^k	54.1	11.3	1.9	48.4	8.2	1.9	49.7	10.3	0.0
Erythema	37.3	1.5	0.2	32.9	1.0	0.0	33.3	1.0	0.0
Swelling	21.6	0.8	0.2	19.7	1.0	0.0	17.3	0.0	0.0
Systemic									

Determine if clinical value outweighs patient risk

<https://www.fda.gov/media/119856/download>

Artificial Intelligence and Machine Learning in Software as a Medical Device

Change management

In the framework described in the discussion paper, the FDA envisions a “**predetermined change control plan**” in premarket submissions. This plan would include the types of anticipated modifications and the associated methodology to implement those changes in a manner that manages risk.

Real-world performance monitoring

In this potential approach, the FDA would expect a commitment from manufacturers on **transparency and real-world performance monitoring for artificial intelligence and machine learning-based software as a medical device** as well as periodic updates to the FDA whenever changes were implemented as part of the approved pre-specifications and the algorithm change protocol.

Assess your process and validation

<https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device>

Business risk – AI & Data



GE HealthCare

Federal Trade Commission Actions



TECH

Amazon to pay over \$30 million in FTC settlements over Ring, Alexa privacy violations

PUBLISHED WED, MAY 31 2023•2:15 PM EDT | UPDATED WED, MAY 31 2023•6:07 PM EDT

- Don't use behind-the-scenes tracking technologies that contradict your privacy promises or otherwise harm consumers.
- Insist your technology people and compliance staff communicate about your company's data intake and privacy policies
- "HIPAA Compliant," "HIPAA Secure," and similar claims may deceive consumers.

Potential of Generative AI for Healthcare

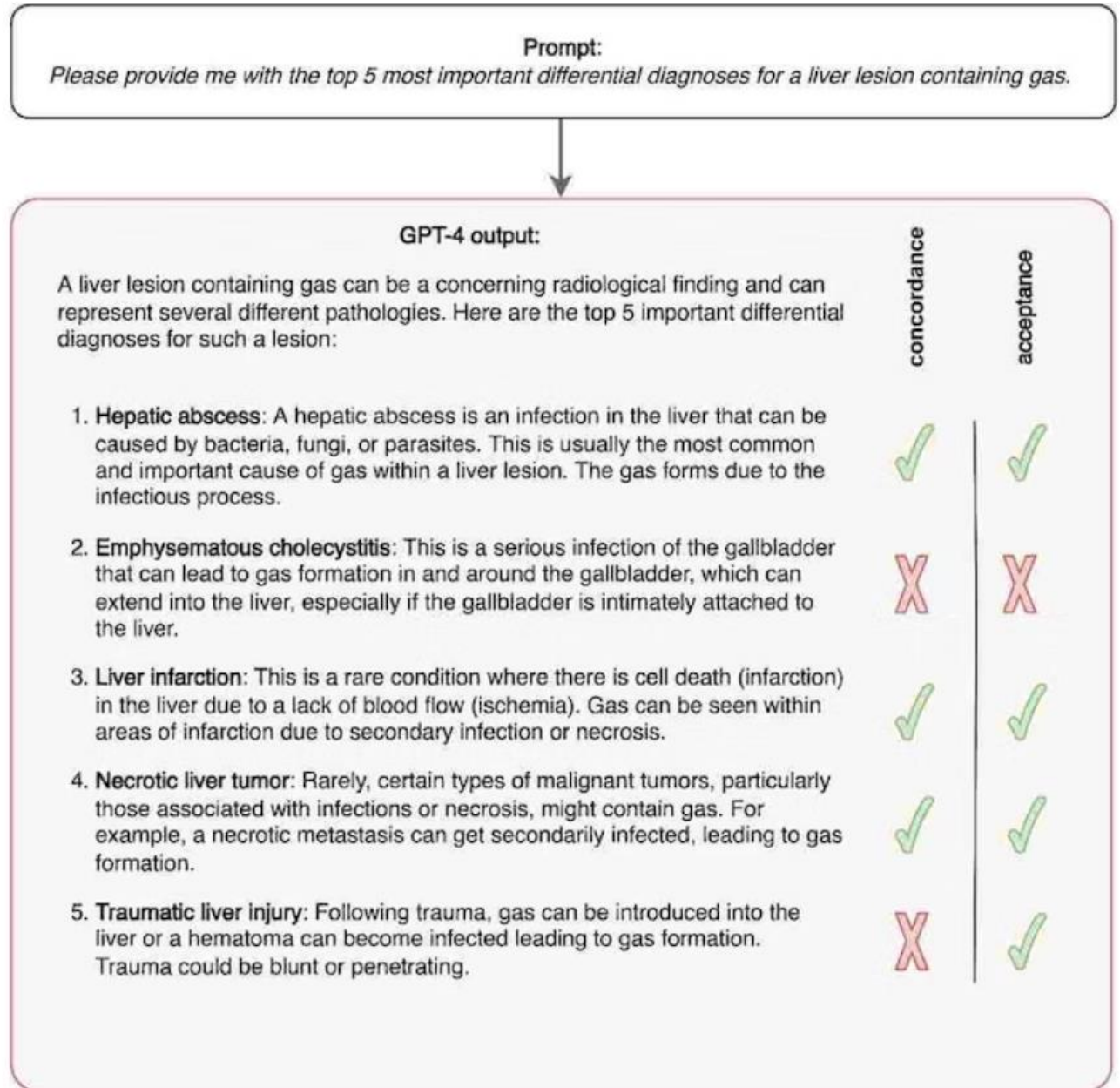
IMAGING INFORMATICS | ARTIFICIAL INTELLIGENCE

Can ChatGPT support diagnostic imaging decisions?

Will Morton
Jul 6, 2023

A proof-of-concept study suggests that ChatGPT-4 can generate relevant differential diagnoses for specific imaging patterns, according to a group in Germany.

Example of a prompt and subsequent outputs of GPT-4 that attained lower concordance (60% [3/5]) and acceptance (80% [4/5]). Image courtesy of Radiology through CC BY 4.0.



Summary

We must learn how to use the immense opportunity for better healthcare that Foundational Models enable

Does it add or reduce risk to patient safety?

Assess the risk added by using this algorithm or system

Reduce risk **as far as possible**

Turn to regulators to weigh in on your risk assessment



GE HealthCare