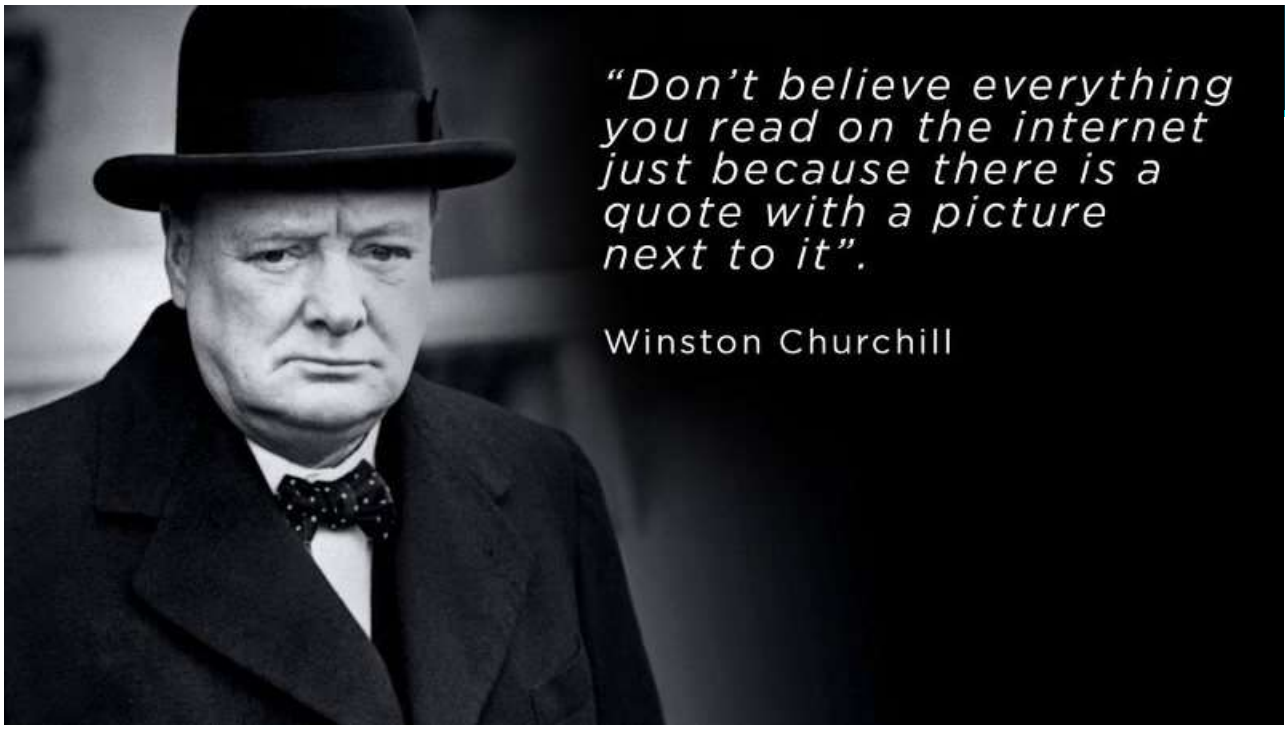
Risk Management Considerations in the Use of Diagnostic and Predictive AI Models

John Banja, PhD
Center for Ethics
Emory University
jbanja@emory.edu

When Models Outperform Human Clinicians

Outperform:

- Fewer false positives
- Fewer false negatives
- Fewer perceptual errors
- Fewer decision errors
- Faster/better disease detection
- Faster/better treatment
- Better outcomes
- Lower cost


“Don’t believe everything you read on the internet just because there is a quote with a picture next to it”.
Winston Churchill

Two Problematic Liability Scenarios in an Era of Advanced AI Models

- Scenario #1:

  The AI model makes a decision or recommendation with which the physician disagrees. However, AI has become the standard of care, so the physician complies with the model's recommendation. A poor outcome follows. The patient-plaintiff learns of the physician's unfortunate decision and decides to sue, arguing that the physician should have heeded his or her clinical experience and not listened to the advice of some computer software.

- Scenario #2:

  The AI model makes a decision or recommendation with which the physician disagrees. Although AI has become the standard of care, the physician decides to follow his or her clinical experience and intuitions, not the model's. A poor outcome follows. The patient-plaintiff learns of the physician's decision and decides to sue, arguing that the AI obviously knew something that the physician didn't and that, as the standard of care, the physician should have followed the AI's recommendation.

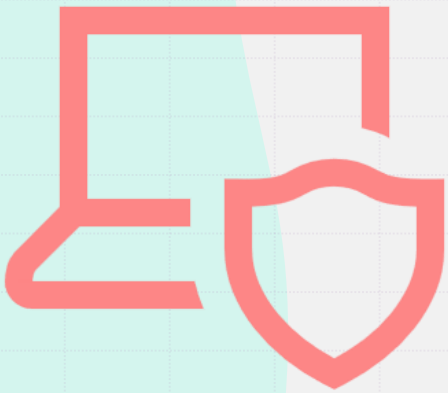
This is not fanciful: "The Appeal to Authority Phenomenon"

- **"Incorrect AI causes radiologists to make incorrect follow-up decisions when they were correct without AI. This effect is mitigated when radiologists believe AI will be deleted from the patient's file or a box is provided around the region of interest…(but) the more accurate an AI system is perceived, the more a radiologist will be influenced by incorrect feedback from that system…Radiologists were less likely to disagree with AI (even when AI is incorrect) if there is a record of that disagreement occurring." Bernstein, et al, "Can incorrect artificial intelligence (AI) results impact radiologists…" European Radiology, June 2, 2023; https://doi.org/10.1007/s00330-023-09747-1.**

But look at this one!

- At some point in the near future, AI has become the standard of care. Radiologist Smith has performed an abdominal study as requested. At Smith's clinic, an AI model does the first read and in this case, it declares a finding of pancreatic cancer and places a box around the affected area. Smith inspects the images very carefully and finds absolutely nothing irregular or concerning. He asks a colleague to look at the images as well. Smith's colleague agrees there is no sign of cancer. Smith decides the model has simply made an error and reports it to the AI/IT team. However, he does not share this information with the patient. Six months later, the patient returns complaining of abdominal pain. Another study is performed which clearly shows pancreatic cancer. Dr. Smith is dumbfounded and perseverates on whether he has done anything wrong.

Work with your professional organizations and liability carrier on these problems



Patient Data Privacy and Hackers

Scenario: Senator X is a prominent national political figure who is receiving mental health services at one of your facilities. Her medical records are hacked, and her patient data are stolen. Subsequent (federal) investigation discovers that the security measures to protect (all) patient data were inadequate. Adding to the problem is that Senator X's political rivals acquired her mental health data and have disclosed it to media and the public. Senator X files a lawsuit against the facility, which now also faces fines for HIPAA violations.

Reliability and Consistency of AI Outputs: Navigating the “Standard of Care”

- Scenario: A patient is transferred from Hospital A to Hospital B. The referring hospital had sophisticated scanning devices and diagnostic AI models. Hospital B does not have such technology and treated the patient with what was available. The patient’s condition rapidly deteriorated and he died. When clinicians at the referring hospital learned of this and Hospital B’s (out of date, in their opinion) technology, they recommended the family sue Hospital B for providing inferior, substandard care.

Informed Consent for Data Use

Scenario: Mrs. X had gynecologic surgery at a university research hospital and signed an informed consent form that stated: "I allow the hospital to use my medical data for research purposes." Some months later, a newspaper featured the hospital's latest research on birth control interventions. Mrs. X suspected her data were used and became very upset because any form of birth control violated her religious beliefs. When she contacted the hospital, she was told about her voluntary consent to research, to which she replied, "But I didn't know you were going to use my data for THAT."

Explainability

- Scenario: Mr. Smith's physician tells him that a new AI system the hospital purchased has recommended a novel chemotherapy for his cancer. The physician remarked that he had not ever seen this type of chemo recommended before. Mr. Smith asks, "Why is the AI recommending this?" Dr. Jones says, "Well, I guess the AI has digested millions of patients' records and thinks this is the best course of treatment for you." Smith responds, "Yeah, but why?" Dr. Jones says, "Well, to be honest, we really don't know. The AI model is a "black box," and the AI developer won't allow us to open up the model and analyze how it works. It's intellectual property and protected from external investigation." Smith responds, "You're recommending something for my cancer, but you don't know why? Dr. Jones, this stuff better work."

What happens when you don't know how the model reached its decision?

- AI model predicts patient's race
- Was able to predict regardless of anatomical region
- Prediction not due to:
  - Surrogate markers, e.g., BMI, breast density, etc.
  - Corrupted, cropped, degraded images

Judy Gichoya, et al. AI recognition of patient race in medical imaging: a modelling study. The Lancet. 4, June 2022:e406-e414.

Bias and Discrimination

- Scenario: In order to conserve revenues, Leviathan Hospital is using an AI model to predict patient no-shows and then double book the slot. If the model predicts wrongly, however, patient treatment times will be decreased, waits and delays will be longer, and non-physicians might have to see the patient rather than a physician. Some hospital staff believe that double booking unfairly discriminates against the less well-off who often lack transportation or are too sick to make it to their appointments. Hospital administrators defend double booking and say, “But these patients are the reason we have to double book. If there are burdens connected with double booking, then this population of patients should bear those burdens rather than more reliable patients who do show up.”

Who owns the data and how will it be used?

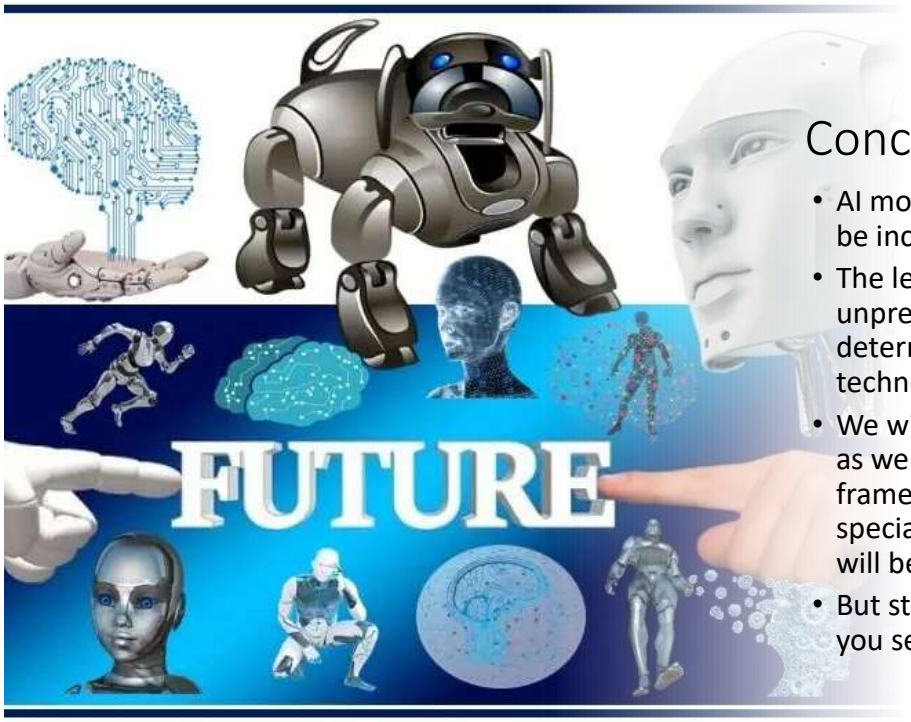
Scenario: Patients with a very rare form of cancer have learned that data from their genetic tests are providing new and valuable insights into the general treatment of cancer. These patients have received no remuneration for their medical data, but they are beginning to feel that since their genetic data are so valuable, they should be paid. A few of them schedule an appointment with a prestigious law firm to analyze the matter further.

Automation

1. Scenario: As AI models become increasingly used in health care, they will assume many functions that humans previously performed. Models will likely be used to read all kinds of radiographs, perform microsurgeries, and even do counseling and psychotherapy. Some veteran health professionals are extremely concerned about how the “human touch” will disappear from health care. Should they be and should risk managers be concerned?

Hospital Liability

- Scenario: AI models depend on the accuracy and representativeness of the training data that are used to create them. That data is itself generated from extant technology, e.g., new cameras, new EMR technologies, new genetic tests, that is constantly being developed. This means that AI models will have to be continuously upgraded because the older training data and programs that depended on them have become outdated. In turn, that means that the hospital that is lax in upgrading its AI technology is assuming a considerable malpractice risk if it continues to use outmoded models that do not meet the standard of care.



Conclusion

- AI models will only improve and be increasingly used
- The legal system is presently unprepared to make liability determinations for advanced AI technology
- We will invent the liability rules as we do for all legal and moral frameworks; that means that special interest group lobbying will be intense
- But still: Don't believe everything you see or hear.


The End