DIGITAL TRANSFORMATION IN MEDICINE:
HOW INNOVATIONS AND ARTIFICIAL INTELLIGENCE ARE SHAPING THE FUTURE OF HEALTHCARE

Beyond the Heatmap:
Explainable AI in Clinical Decisions

Balancing Diagnostic Accuracy, Transparency, and Ethical Responsibility

Presenter : Mohammed Aatir Shaikh

MD Program

East European University

Epistemic Crisis in Medicine AI

71%

Trust Deficit

Intensive care professionals express doubt regarding reliable AI use in ICU decision-making.

The Black Box Problem

Deep learning models provide high predictive accuracy but operate with opaque, non-linear reasoning that clinicians cannot decode or verify.

Decision Paralysis

Clinicians face immense risk integrating unexplainable algorithmic outputs into high-stakes environments, hampering timely and confident interventions.

False Hope of Post-Hoc Tools

- **Current Standard** Post-hoc tools like Saliency Maps, LIME, and SHAP attempt to explain a decision after the fact.

- **The Critique** Broad explanations are often highly unreliable or superficial for individual, patient-level clinical decisions.

- **The Danger** The "illusion of understanding" can actually hamper a clinician's ability to detect sizeable and critical model errors.
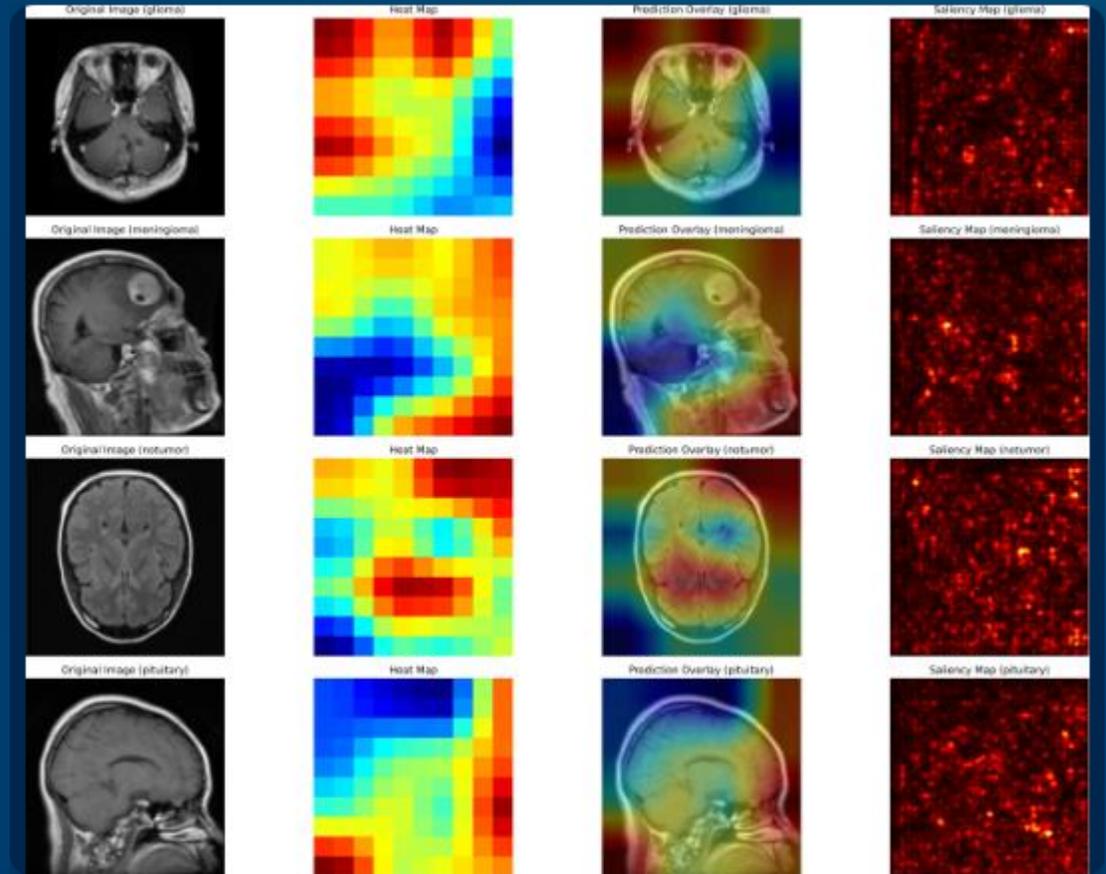
Inherent Interpretability

The Paradigm Shift
Moving away from post-hoc visualizers to transparent "glass-box" architectures that explicitly define their reasoning.

Concept Bottleneck Models
Advanced CBMs predict intermediate, human-interpretable concepts before generating the final diagnostic output, keeping the physician securely in the loop.

Oncology Case Study
The 2025 ConSurv breast cancer model predicts risk based on transparent biological pathways (e.g., hypoxia, angiogenesis) without sacrificing baseline predictive accuracy.
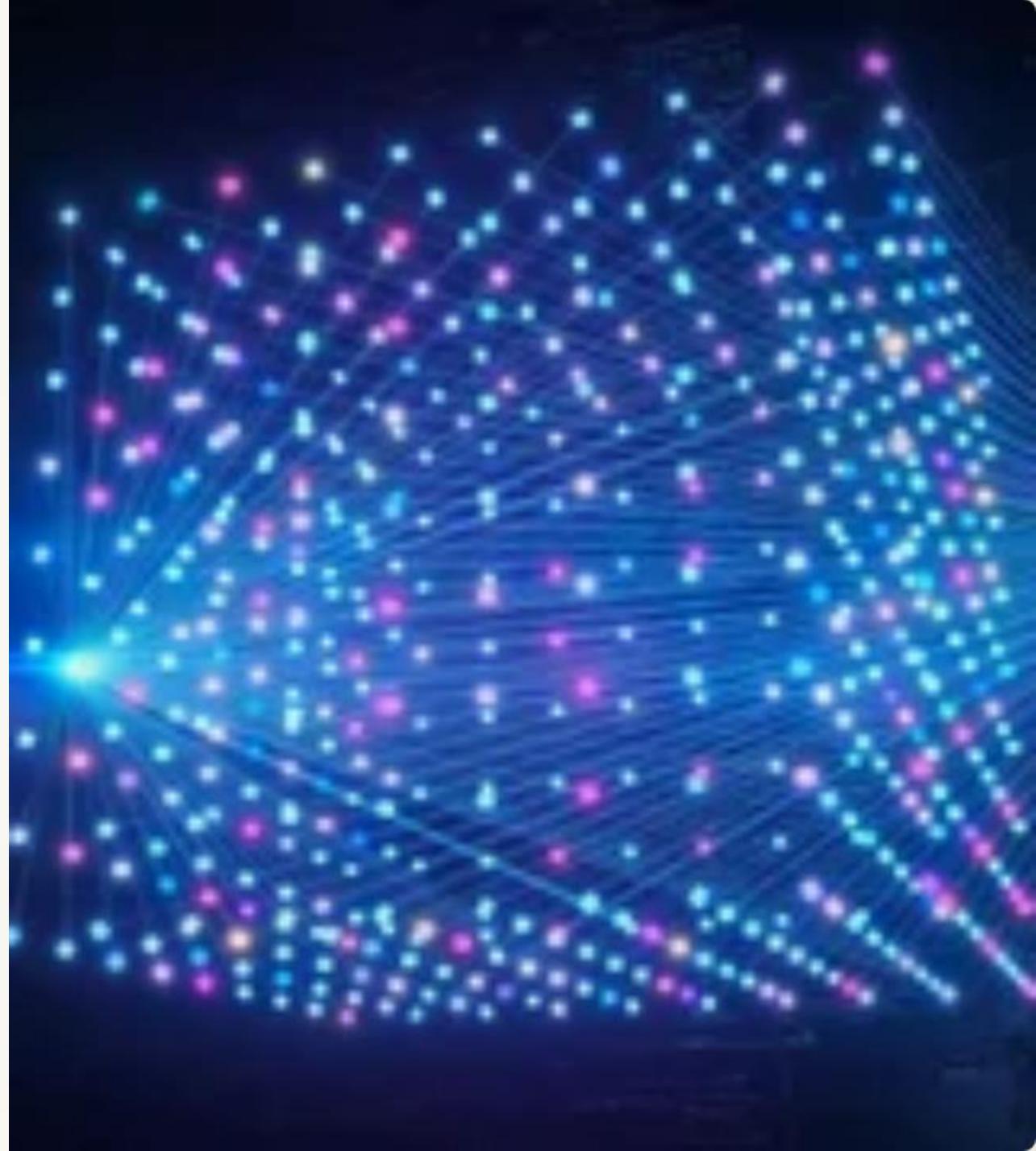
Clinical Successes

❤ Cardiology (CHAMPION)

AI analysis of remote wearable sensors proved revolutionary in early detection and active patient management.

- Reduced heart failure patient volume by **33.1%**

- Improved vital medication adherence by **20-30%**

- Provided interpretable early warnings for rapid intervention.

⊞ Imaging Innovations

Hybrid frameworks like SpikeNet are setting new, highly transparent benchmarks in complex medical image analysis.

- Integrates advanced Spiking Neural Networks.

- Achieves an exceptional **97.1%** diagnostic accuracy.

- Delivers high explanation fidelity alongside extreme efficiency.

Legal & Ethical Frameworks

The "Liability Sink"

Clinicians risk absorbing outsized legal blame if forced to oversee systems without having genuine, transparent control over the algorithmic output.

Regulatory Demands

The EU AI Act (Article 14) legally mandates explicit human oversight, making "black-box" systems functionally non-compliant in high-risk scenarios.

FUTURE-AI Guidelines

The 2025 consensus emphasizes six core pillars: Fairness, Universality, Traceability, Usability, Robustness, and Explainability.

Educational Imperatives

The Current Gap: Most medical students and faculty self-identify as novice AI users.

TIER 2
Evaluative Proficiency

Teaches clinicians how to properly interpret and critically evaluate XAI outputs for patient diagnostics.

FUTURE
Institutional Shift

Integrating AI literacy as a core, competitive differentiator and mandatory component in medical curricula.

TIER 1
Foundational Skills

Focuses on safe usage practices and raising awareness about automation bias among early medical trainees.

TIER 3
Advanced Skills

Equips leaders with the requisite knowledge for complex AI system governance and institutional deployment.

Conclusion & Future

◎ Summary

True clinical utility strictly requires balancing predictive power with inherent, "glass-box" interpretability to protect patient autonomy and clinician liability.

💬 The Future of XAI

Moving away from static heatmaps and migrating towards establishing "genuine dialogue" between the clinician and the algorithmic system.

🤝 Ultimate Goal

Positioning AI as a collaborative partner equipped with social capabilities to interact seamlessly and transparently with the clinical team.
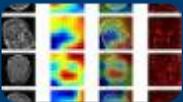
Questions & Discussion

Thank you for your attention. I am now open to your questions.

- ❓ How do we manage the accuracy vs. interpretability trade-off in low vs. high-risk settings?

- ❓ What is the rigorous threshold for informed consent when using AI?

- ❓ How do we actively prevent "automation bias" among incoming medical trainees?

mohammed.aatir.shaikh@.eeu.edu.ge

References

1. **Ghassemi, M., Oakden-Rayner, L., & Beam, A. L.** (2021). The false hope of current approaches to explainable artificial intelligence in health care. *The Lancet Digital Health*, 3(11), e745-e750. https://doi.org/10.1016/S2589-7500(21)00208-9

2. **Lecuyer, H., et al.** (2025). ConSurv: A Concept Bottleneck Model for Transparent Survival Analysis in Oncology. *bioRxiv*. https://doi.org/10.1101/2025.04.11.648380

3. **Lekadir, K., et al.** (2025). FUTURE-AI: Guiding principles and consensus recommendations for trustworthy artificial intelligence in medical imaging. *The BMJ*, 388, e081554. https://doi.org/10.1136/bmj-2024-081554

4. **Räz, T., Pahud De Mortanges, A., & Reyes, M.** (2025). AI in medicine: Challenges of integrating XAI into the future clinical routine. *Frontiers in Radiology*, 5, 1627169. https://doi.org/10.3389/fradi.2025.1627169

5. **Giorgetti, C., Giorgetti, A., & Boscolo-Berto, R.** (2025). Establishing new boundaries for medical liability: The role of AI as a decision-maker. *Advances in Clinical and Experimental Medicine*, 34(10), 1601–1606. https://doi.org/10.17219/acem/208596

6. **SCAI.** (2025, May 2). New Studies Highlight Potential of Artificial Intelligence to Improve Heart Failure Detection and Treatment. *Society for Cardiovascular Angiography & Interventions*. Retrieved from scai.org.

7. **Amann, J., et al.** (2020). Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Medical Informatics and Decision Making*, 20, 310.

https://www.frontiersin.org/files/Articles/1535478/xml-images/fonc-15-1535478-g007.webp

Source: www.frontiersin.org

https://www.ibm.com/adobe/dynamicmedia/deliver/dm-aid--09453567-91b3-495b-ab61-97b6c8eb3615/adobestock-1511873223.jpg?preferwebp=true&width=320

Source: www.ibm.com

https://img.freepik.com/premium-photo/female-doctor-interacting-with-futuristic-medical-interface-holographic-display-data-analysis_174533-96627.jpg

Source: www.freepik.com