Continuous Governance of AI-Assisted Assessment in Medical Education: A Practice-Informed Framework

Huong-Tra Ngo, PhD,[1] , Ameer H. Shakur, PhD,[2] , Michael J. Holcomb, MS,[2] ,
Andrew Wong, MD, MS,[1] , Judah Gruen, MD,[2] , Andrew R. Jamieson, PhD,[2]

[1]University of Utah, [2]UT Southwestern Medical Center

Introduction

Artificial intelligence (AI) can support scalable assessment in medical education, but prospective AI-assisted assessment raises governance questions extending beyond model performance.

Research and learner protections, data stewardship, deployment authorization, and poster-deployment monitoring require coordination across existing institutional structures.

At UTSW, we deployed a multimodal AI platform[1] for automated grading of Objective Structured Clinical Examination (OSCE) - a standardized, simulation-based assessment in medical education through structured encounter. We are scaling at 5 other institutions across the UT system[2].

Drawing on this experience, we derived a practice-informed framework for governing the deployment of AI-assisted assessment in medical education.

Participants In Governance

Deployment & Monitoring

- ☑ Learner Representatives
- ☑ Research Team (Principle Investigator)
- ☑ Data Science and Informatics Team
- ☑ Platform Operators

Educational / Operational

- ☑ Medical School Leaderships
- ☑ Educators
- ☑ Simulation / Clinical Skill Leaders
- ☐ Finance / Budget / Operational Leaderships

Research & Institutional Compliance

- ☑ Human Research Protection Program
- ☑ Institutional Review Board
- ☐ Office of Research Integrity
- ☐ Office of Institutional Compliance

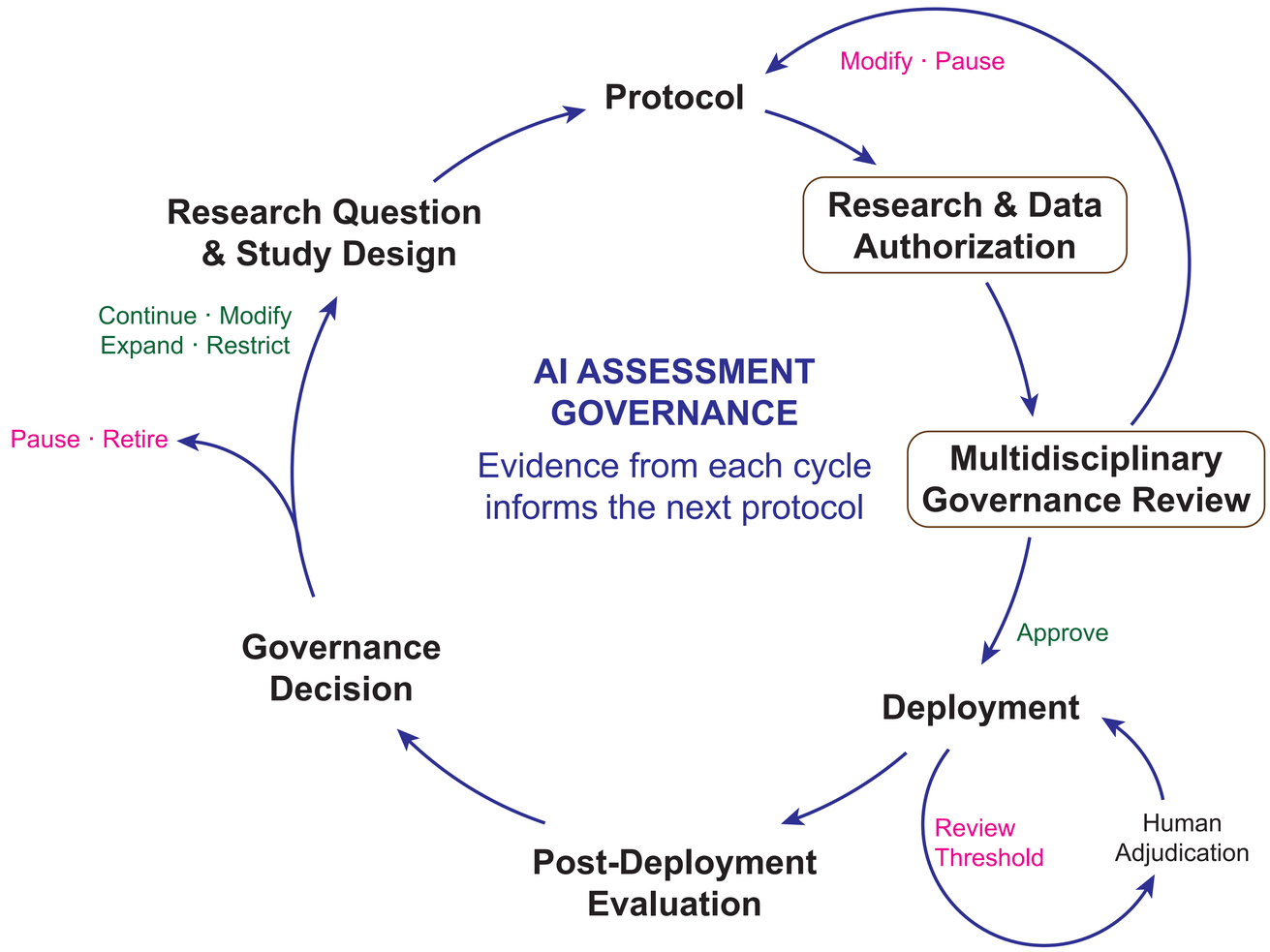
Data / Privacy / Security

- ☐ Enterprise IT
- ☐ Privacy
- ☐ Cybersecurity
- ☑ Data Steward

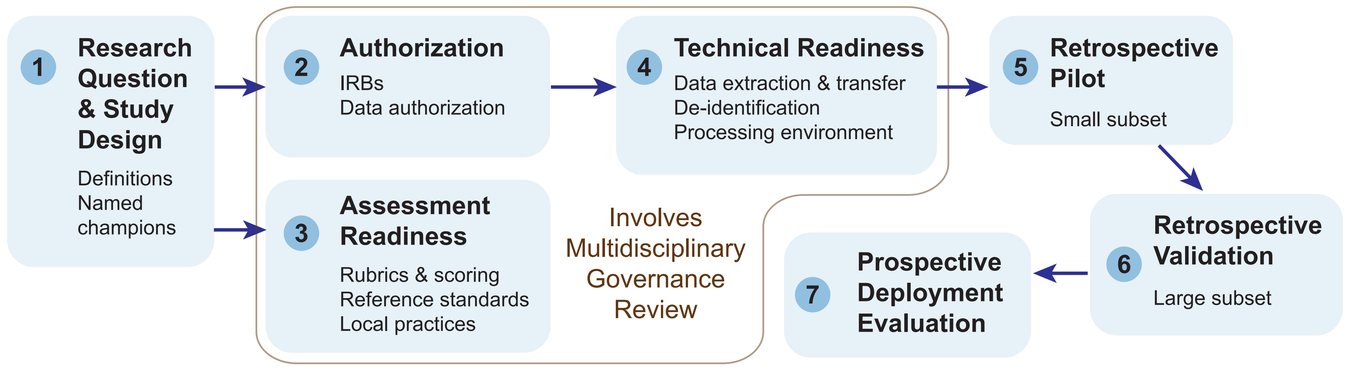
Legal / Agreements

- ☐ Legal Counsel
- ☐ Contracts Administration
- ☐ Technology Transfer
- ☐ Procurement

✓ Local first deployment



From Single-Site Evaluation to Multi-Site Deployment



UT REAL Partner Progress [2]

	1	2	3	4	5	6	7
UT Southwestern Medical Center	●	●	●	●	●	●	●
UT Health Houston - McGovern Medical School	●	●	●	◐			
UT Tyler School of Medicine	●	●	●	●	●	●	
UT Rio Grande Valley School of Medicine	●	●	●	◐			
UT Health San Antonio - Long School of Medicine	●	●	●	◐	●		
UTMB Galveston - John Sealy School of Medicine	●	◐	◐				

● Ready ◐ In Progress

Pre-Assessment Checklist

Governance Review Content

- **Inventory of the alternatives**: reliability (score distribution, correlations), failures, successes, resource cost
- **Implementation**: workflow (identified champions, adoption), deployment set-up (data inventory, study domain, model selection), feedback
- **Performance**: agreement, adjudication preference, pass/fail uncertainty estimate.
- **Impact**: learner, resource planning (effort, cost)
- **Decision points**: timeline planning, review thresholds.

Key Lessons

- AI-assessment governance is iterative.
- Responsibilities are distributed between many champions.
- Human-review and escalation criteria should be pre-specified.
- Governance decisions incorporate implementation ease, AI system performance, impact on learners and operation resources, and trade-offs of the alternatives.
- Model, rubrics, workflow changes should prompt consideration of revalidation and governance review.
- AI-enabled management can support shared awareness by making relevant information and decisions visible to stake holders.

Acknowledgement

Funded by UT REAL Health AI Pilot Program from the UT System and by the IMPACT Program from University of Utah.

References

[1] Ngo, H.-T., Shakur, A. H., Holcomb, M., Kang, S., et al. (2026). Prospective deployment of multimodal AI grading for medical student OSCEs [Preprint]. Research Square.
[2] https://ut-real-ai-project-maples.com/

Have questions? Please email tra.ngo@utah.edu