AI for Human Health: Achieving the Promise While Avoiding the Perils

Gregory D. Hager, Director

Johns Hopkins University
Baltimore, MD USA
Johns Hopkins Malone Center for Engineering in Healthcare

**Mission:** To catalyze and accelerate the development, translation, and deployment of research-based innovations that advance the effectiveness and efficiency of health care.

**AI, Machine Learning and Data Analytics**
Develop tools for analytics and machine learning that advance decision support and improve prediction and process modeling for health care.

**Systems Optimization**
Create better systems to streamline hospital operations, tackling problems such as emergency room crowding and scheduling issues.

**Technology & Devices**
Pioneer new assistive technologies and medical devices for patients and care providers.

**User-Centered Design**
Design future end-user technologies, including mobile, web, household, rehabilitation, and robotics technologies.
Our North Star: Medicine is About People Taking Care of People

How is your arm feeling today?
Successful Technology Enhances Care

I just got your images back and I saw that .....
Successful Technology Enhances Care

Medical Images
Lab tests
Patient Vitals
AI

My Analytics tell me ....??
Dear Medicine: Welcome to Life In an Exponential World

PubMed Articles with Keyword Machine Learning


https://medicalfuturist.com/fda-approvals-for-algorithms-in-medicine/
Opportunities Abound – Not All Created Equal

Learning-based algorithms had been shown to accurately
forecast the onset of septic shock,\(^1\)
classify skin lesions with dermatologist-level accuracy,\(^2\)
identify diabetic retinopathy during routine primary care visits,\(^3\)
screen for breast cancer better than human radiologists by a fairly large margin,\(^4\)
differentiate the emergency severity index,\(^5\)
simplify interventional workflows,\(^6\)
enable redesign of infusion centers\(^7\) .... and much much more ....

Peril #1: Let Us Not Forget History

- Expert Systems: Mycin, Dendral

- Use expert knowledge to establish a logical (later probabilistic) chain of reasoning from inputs to outputs

- Viewed as a physician’s assistant (augmenting, not replacing)

- Ultimately hit a wall
  - Not scalable
  - Too many corner cases
  - Hard to engineer
  - Got the easy stuff right, but it was easy ...

- Early results inspire, but it takes effort to find the real value proposition amid the hype
Finding a Value Proposition: Reducing ED Crowding

Machine Learning to triage patients significantly reduced time in ER by reducing uncertainty to admit, treat, or send home

Levin et al., Machine-Learning-Based Electronic Triage More Accurately Differentiates Patients With Respect to Clinical Outcomes Compared With the Emergency Severity Index, Annals of Emergency Medicine, 2018
Finding a Value Proposition: Reducing ED Crowding

**ED Objective**

58 – 82 min decrease \(\downarrow\) in time to critically ill
- Time to ICU transfer
- Time to emergent surgery

20 – 30 min decrease \(\downarrow\) door-to-disposition*
- Site 1: \(\sim\)29 min for admits | \(\sim\)3 min discharges
- Site 2: \(\sim\)10 min for admits | \(\sim\)11 min discharges

8,000 – 11,000 bed-hours gained
$520k - $730k in cost-savings

**Improving Triage**

Increased \(\uparrow\) early identification of critical care
- 15% \(\uparrow\) (e.g., 67% to 80%) identification as high-acuity and expedited
- No real change in low-acuity hospital admits
  - Despite increases in low-acuity designations

**Nurse Agreement with TriageGO**
- 45-55% prior to implementation
- 65-70% early implementation
- 80-90% in steady-state

Machine Learning to triage patients significantly reduced time in ER by reducing uncertainty to admit, treat, or send home
Peril #2: Correlation is Not Causality

- Machine learning is solving a **regression problem**
- Exponential data growth + exponential model size + exponential experiments = **exponential chance for irreproducible results**
Peril #2: Correlation is Not Causality

For COVID-19:
- 33 diagnostic models were identified for detecting COVID-19,
- 75 other diagnostic models based on medical images,
- 10 diagnostic models for severity classification,
- 107 prognostic models for predicting progression

Proposed models are poorly reported and at high risk of bias ... their predictions could be unreliable when applied in daily practice

Two prediction models ...were identified as being of higher quality

(Wynats et al., BMJ, 2020)
We Will Learn to Transcend the Limitations of Data

From development to deployment: dataset shift, causality, and shift-stable models in health AI

ADARSH SUBBASWAMY
Department of Computer Science, Johns Hopkins University, 160 Malone Hall, 3400 N. Charles Street, Baltimore, MD, USA

SUCHI SARIA*
Department of Computer Science: Department of Applied Math & Statistics, and Department of Health Policy & Management, Johns Hopkins University, 160 Malone Hall, 3400 N. Charles Street, Baltimore, MD, USA
ssaria@cs.jhu.edu

“In domains like healthcare and education where safety is crucial, building models that don’t fail in unanticipated ways due to small shifts in underlying datasets is critical.”
Peril #3: Forgetting Medicine is About People Taking Care of People

INNOVATION

AI Will Change Radiology, but It Won’t Replace Radiologists

by Thomas H. Davenport and Keith J. Dreyer, DO

MARCH 27, 2018  Harvard Business Review
DeepCAT: Deep Computer-Aided Triage of Mammograms to Enhance Efficiency

Of 595 testing images, DeepCAT would theoretically discard 315 images (53%) as normal!

None contained a malignant mass.

Yi et al., *Journal of Digital Imaging*, 2021
AI Will Change Medicine: But not by replacing doctors or nurses
... but by changing the way they do their job!

Less effort on needless work
Better decisions informed by analytics
More time for what matters: taking care of patients