Values, Metrics, and Impactful Medical Imaging Research

Brian E. Chapman, PhD

University of Melbourne

2024-05-24

Section 1

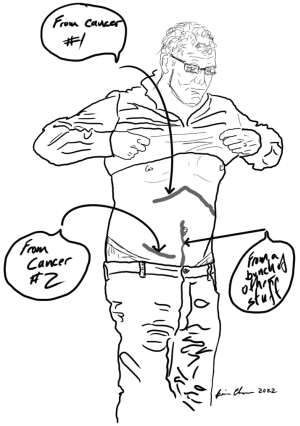
A little bit about myself

Education and experience

- BSEE
  - Senior thesis with ORNL on microwave sintering of ceramics
  - What I really liked was mathematical modeling
  - Started PhD in Plasma Physics
  - ORNL device incident
- But couldn't escape urge to work in healthcare

Why the healthcare urge?

- Long-term patient
  - 4-time cancer survivor
  - Two childhood cancers (1976, 1983)
  - Two adult cancers
  - 12+ surgeries
  - 4 emergencies
  - Advanced peritonitis
  - Gangrenous strangulated bowel
  - Many resulting chronic issues!



My “mobile medical record”

Education and experience

- PhD in medical informatics
  - MR Angiography
- Worked primarily in imaging and NLP in the context of radiology
- Always looking for the blend of my two driving interests: modeling and health

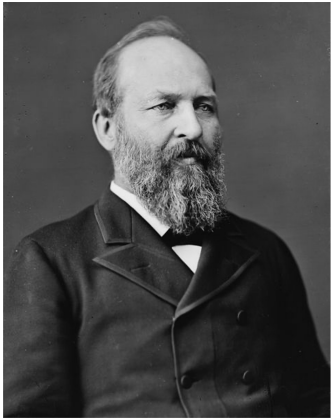
Current position

- Computing and Information Systems at University of Melbourne
- Digital Health Cluster
  - Clinical process mining
  - Digital mental health
  - NLP
  - Ethics and epistemology
- A lot of teaching!
- Creating a collaborative, collegial group

Section 2

Motivation

A tale of two presidents: James Garfield and Ronald Reagan



- Shot July 2, 1881
- Died Sept 19, 1881 (80 days later)

- Shot March 30, 1981
- Died June 5, 2004 (8468 days later)

What was the difference?

- Antibiotics?

What was the difference?

- Antibiotics?
- Improved surgical techniques

What was the difference?

- Antibiotics?
- Improved surgical techniques
- **Medical imaging**

What was the difference?

- Antibiotics?
- Improved surgical techniques
- **Medical imaging**
- Garfield scenario:

What was the difference?

- Antibiotics?
- Improved surgical techniques
- **Medical imaging**
- Garfield scenario:
  - **Where was the bullet?**

The Inventor vs the Physician (Millard and Michael 2017)

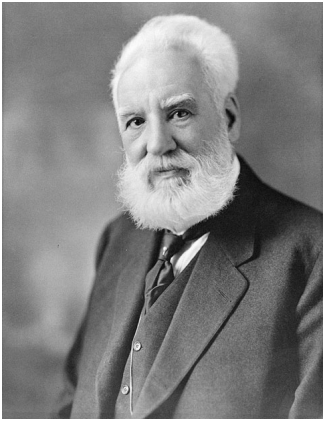


Figure 1: Alexander Graham Bell (Wikipedia)

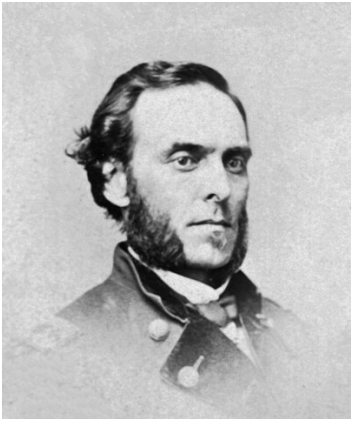


Figure 2: D. Willard Bliss (Wikipedia)

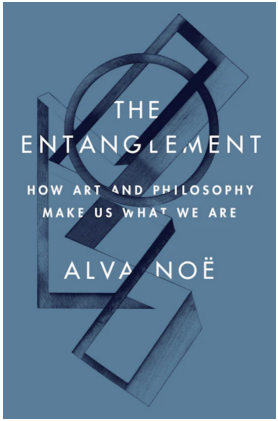
My motivation: early imaging

- 1976: First childhood cancer
  - Vascular imaging: invasive, performed under general anesthesia
    - diagnostic error
  - Ultrasound: had to travel to a different hospital
  - In the absence of reliable soft tissue imaging (e.g. CT, MRI), more expansive surgery was performed (palpating solid organs)
- Contrast this to what would be done in 2024

What do medical imaging researchers aim to do?

- Reduce the burden of human disease and injury by
- Obtaining currently available image information
  - At lower cost
  - At decreased risk
  - At increased quality
- Obtaining image information not previously available

The Entanglement (Noë 2023)



- Technologies create habits of behavior
- Art and philosophy disrupt those habits

Entangled Research

- I would expand Noë's concept of entanglement to include **multidisciplinary research**
  - Each discipline brings its own set of dogmas, traditions, and values
  - Engage in fruitful dialogue **as peers**

Entangled Research: Not always successful

- Bell vs. Bliss
  - Not peers
- Anonymous Nordic computer science researcher: Very difficult for successful medical projects to come out of computer science departments
  - Lack of contextual knowledge or lack of engagement
  - Different values

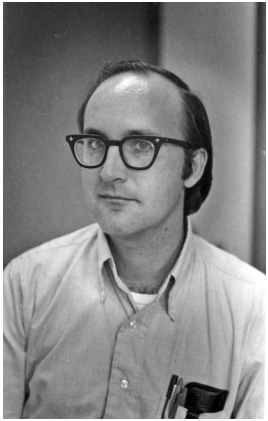
Dissertation: Context for vascular imaging

- Gold standard imaging: catheter angiography
  - invasive, costly
- Catheter angiography evolving techniques
  - digital subtraction
  - cut film
- Developing alternatives:
  - CTA
  - MRA

MRA work

- How to obtain images of sufficient contrast to noise ratio
  - In reasonable time
  - With reasonable levels of artifacts
- The specifics have long lost relevance
- What were the big-picture lessons I took away?

Harry Barrett, Ph.D. (Arizona)



- The quality of an image is **task dependent**
  - Observer performance experiments
  - Other measures are surrogates of varying quality
- Since quality is task dependent, expensive/difficult to generalize
- Imaging research needs to be tied to specific clinical tasks and assessed against those tasks

Fryback and Thornbury (Wisconsin) (Fryback and Thornbury 1991)

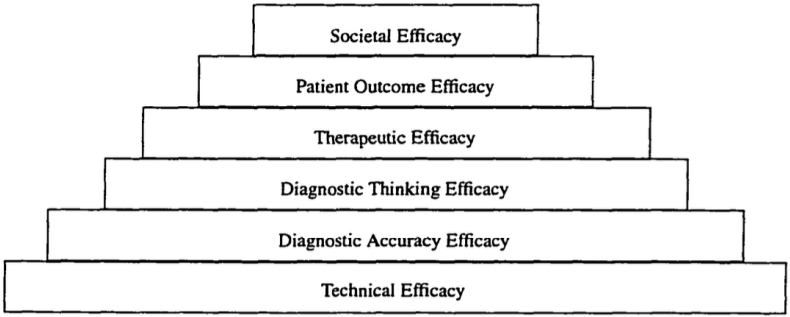
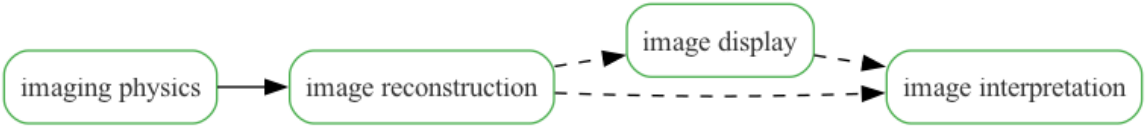


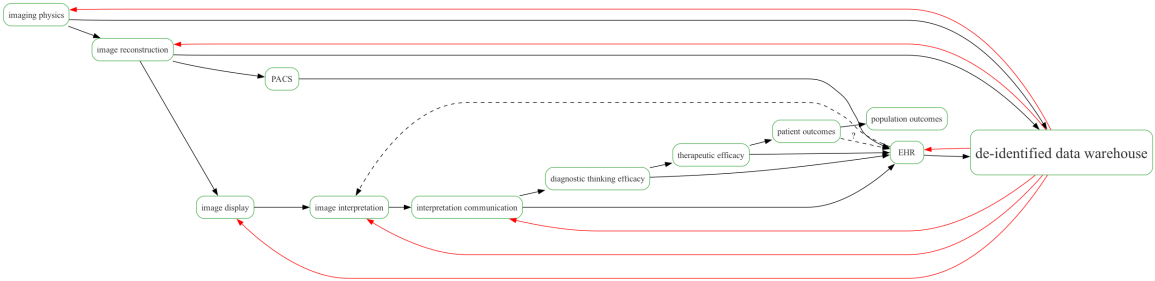
Figure 1.1. Model of diagnostic imaging efficacy adapted from Fryback and Thornbury

Fryback and Thornbury: Implications

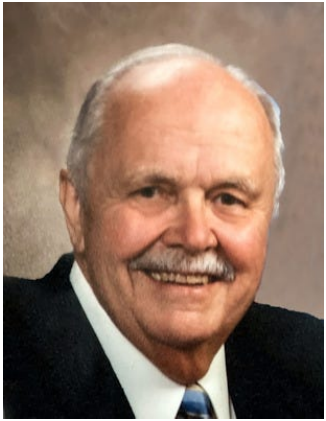


Fryback and Thornbury: Typical research ecosystem

Fryback and Thornbury: Implications (as I see it)

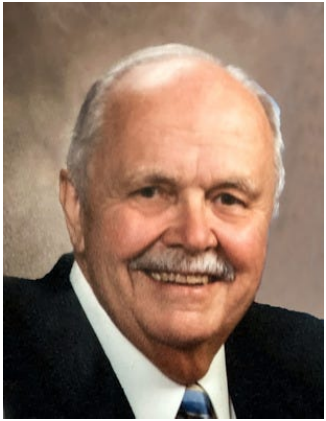


James Potchen, M.D. (Michigan State)



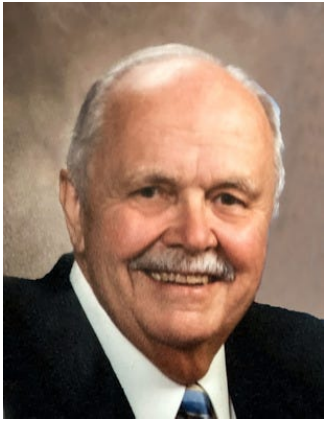
- How do you measure the value of two equally accurate but imperfect imaging tests?

James Potchen, M.D. (Michigan State)



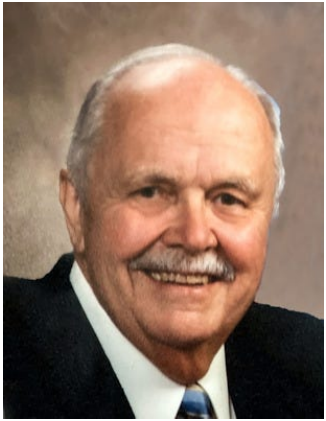
- How do you measure the value of two equally accurate but imperfect imaging tests?
- What if you measured it by how much a patient would pay to avoid either of the two tests?

James Potchen, M.D. (Michigan State)



- How do you measure the value of two equally accurate but imperfect imaging tests?
- What if you measured it by how much a patient would pay to avoid either of the two tests?
- Who is the image valuable for? What are the values?

James Potchen, M.D. (Michigan State)



- How do you measure the value of two equally accurate but imperfect imaging tests?
- What if you measured it by how much a patient would pay to avoid either of the two tests?
- Who is the image valuable for? What are the values?
- **Where is the patient in medical imaging research?**

Alfred North Whitehead

"Plato and Pythagoras stand nearer to modern physical science than does Aristotle. . . .
"The popularity of Aristotelian Logic retarded the advance of physical science throughout the Middle Ages. If only the schoolmen had measured instead of classifying, how much they might have learnt!
"Classification is necessary. But unless you can progress from classification to mathematics, your reasoning will not take you very far."
~Alfred North Whitehead, *Science and the Modern World*

Quantitative imaging

- Anonymous chair of radiology circa 1995: “A radiologist with a ruler is a radiologist in trouble.”

Quantitative imaging

- Anonymous chair of radiology circa 1995: “A radiologist with a ruler is a radiologist in trouble.”
  - Why?

Quantitative imaging

- Anonymous chair of radiology circa 1995: “A radiologist with a ruler is a radiologist in trouble.”
  - Why?
  - Valuing professional judgment (“Anyone can measure!”)

Quantitative imaging

- Anonymous chair of radiology circa 1995: “A radiologist with a ruler is a radiologist in trouble.”
  - Why?
  - Valuing professional judgment (“Anyone can measure!”)
- Prevalent but not universal value.

Quantitative imaging

- Anonymous chair of radiology circa 1995: “A radiologist with a ruler is a radiologist in trouble.”
  - Why?
  - Valuing professional judgment (“Anyone can measure!”)
- Prevalent but not universal value.
  - Vasculitis

Quantitative imaging

Carl Jaffe, M.D.: "No one in clinical [drug] trials takes radiology seriously." (CaBIG, Dec. 2005)

- **The inability to quantitatively monitor therapy**
- The inability to validate findings
- Lack of transparency and data sharing between institutions
- Failure to integrate clinical information in the image assessment
- The unreliability of site interpretations of the imaging studies

Pretty broad re-imagining of radiology

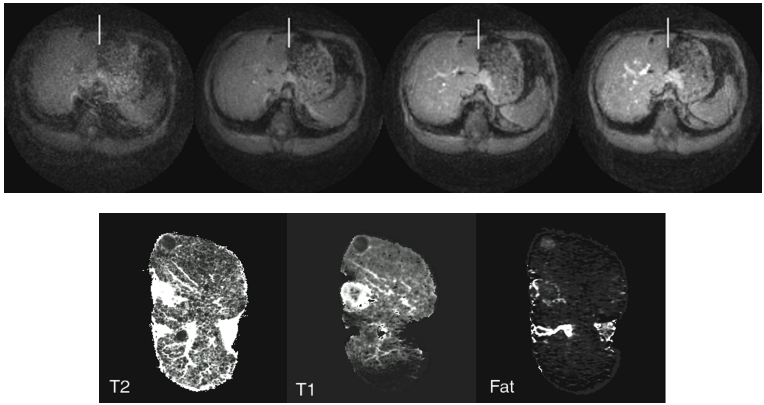
Quantitative imaging: Radiology adapts

Quantitative Imaging Network formed in 2008

Quantitative Imaging: Liver (with W. Marsh)

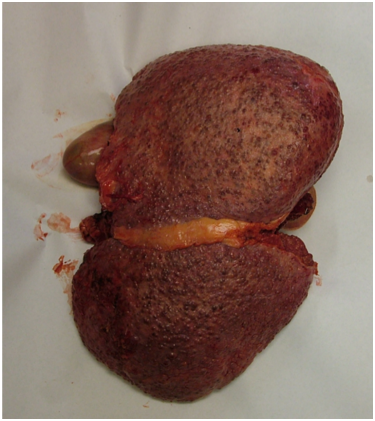
- Early diagnosis of hepatocellular carcinoma (HCC)
- Standard imaging relies on arterial hypervascularity
  - Substantial proportions of patients have metastatic disease by time hypervascularity is present
- Quantify non-vascular MR features
- Detailed correlation with pathology

Quantitative MR (with F. Boada)



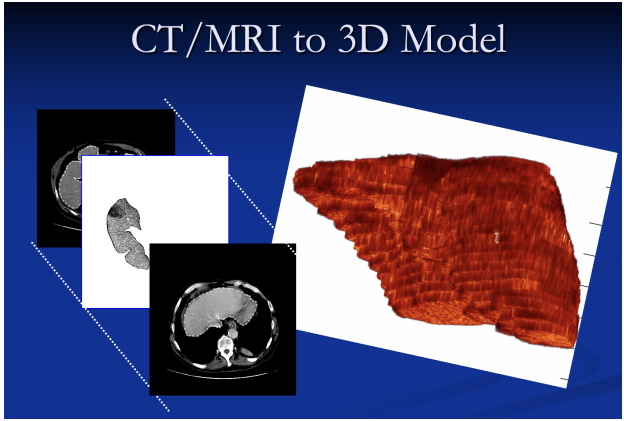
We might have overdone the quantification thing.

Liver pathology correlation (with S. Finkelstein)



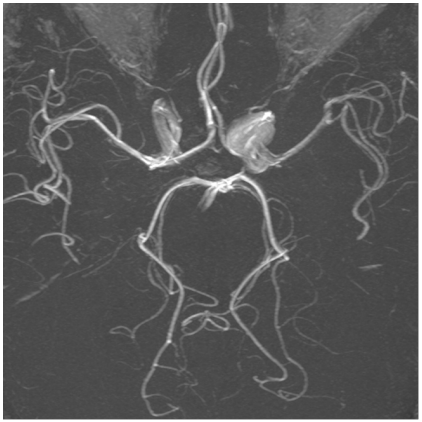
Entanglement: Research, Quality Assurance, Clinical Practice

Quantitative correlation



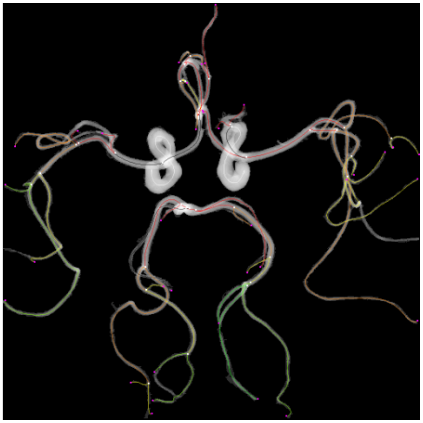
We definitely underestimated the difficulty of this problem!

Quantitative Imaging: Vascular (with D. Parker)

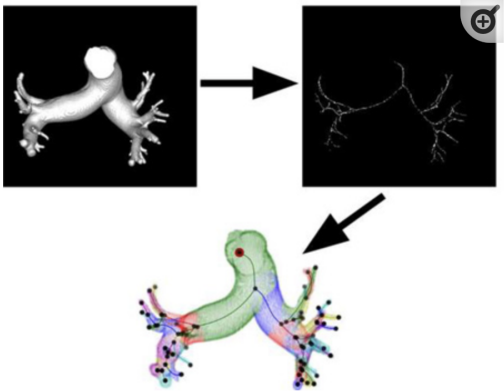


Quantitative Imaging: Vascular

Quantitative Imaging: Vascular



Quantitative Imaging: Vascular (Chapman, Berty, and Schulthies 2015)



Creating an imaging research infrastructure (on a limited budget)

- Prioritize de-identification
  - Data connected with scientific publications should be shareable
  - Minimize patient risk and researcher liability
- Prioritize computational reproducibility
  - Docker/Singularity
  - Kitware's Girder (get data off of student laptops)
- Facilitated substantial increase in research projects by residents and junior doctors
  - "Punk informatics!"

Current project: Type B Aortic Dissection (TBAD)

- Ruth Lim (radiology)
- Vijay Rajagopal (bioengineering)
- Decision support for type B (descending) aortic dissections
- Maintain medical management or proceed to surgery for uncomplicated TBAD (70%)?
- Intervention risks:
  - Stroke: 4%
  - Renal failure: 11%

Current project: Type B Aortic Dissection (TBAD)

Classic problem of tedious measurements with high inter-rater variability (Willemink et al. 2023)

- Extracted imaging features
  - maximum aortic diameter,
  - relative false lumen circumference,
  - false lumen drainage pattern
  - the number of identifiable intercostal arteries along the dissected aorta
- Intraclass correlation coefficients ranged from 0.04 to 0.68 and 0.5 to 0.89
- Risk (low, medium, high) differences 44.7%

TBAD: Where are the research pain points?

- Computing resources
- Imaging research infrastructure
  - Replicate what I did in Utah
- **Labeling the images**
  - ITKSnap (same tool I've been using for 20 years)

Current project: Type B Aortic Dissection (TBAD)

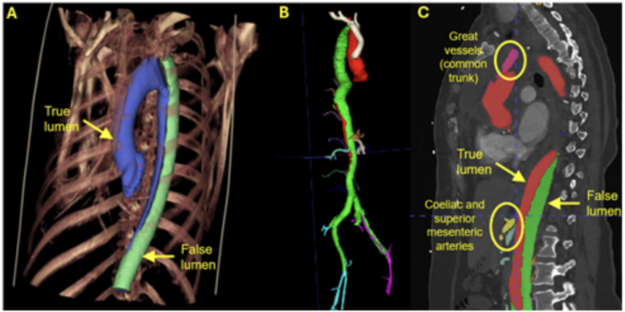
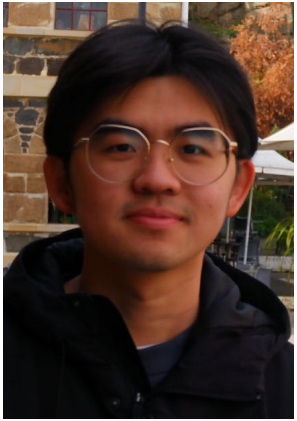


Figure 2. A) 3D rendering of our preliminary model trained on open-source CT data showing automatically segmented true and false lumina of a TBAD patient. B) 3D rendered and C) selected slice from CTA of a residual TBAD patient from Austin Health with ground truth manual labels for model training.

TBAD: Improving segmentation

- Hao Xu
  - MS candidate CIS
- Improve segmentation of thrombus (unbalanced data)
- Limited computation resources
  - No access to university HPC
- No access to hospital data
  - Can't move on to the really interesting questions

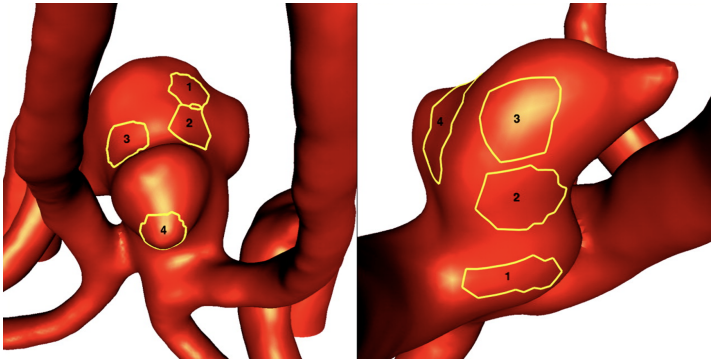


TBAD: Improving segmentation

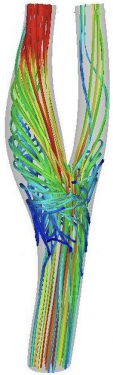
Methods	TL	FL	FLT	True FLT
Default (Yao et al. 2021)	0.78	0.68	0.15	0.25
Single step	0.91	0.88	0.26	0.43
Sequential, multi-task (Wobben et al. 2021)	0.86	0.78	0.32	0.37
Classification → segmentation	0.79	0.67	0.06	0.06
Ensemble (transformer + convolution)	0.91	0.89	0.55	0.43

Really interesting questions: Segmentation to Computation

Moving beyond surrogate measures



Castro, et al, AJNR November 2006 27: 2061-2068



Imperial College London

The Entanglement

How do I use art and philosophy to reimagine medical imaging?

Medical knowledge from RCT is "impoverished"

"Often it is not too much of a leap to infer that the intervention causes the positive outcome." **But the resulting knowledge is rather impoverished: it is knowledge of what works, without knowledge of how it works (or why it does not work), or how to make it work better. It is knowledge of effects without knowledge of underlying mechanisms.**" (Solomon 2015, 117)

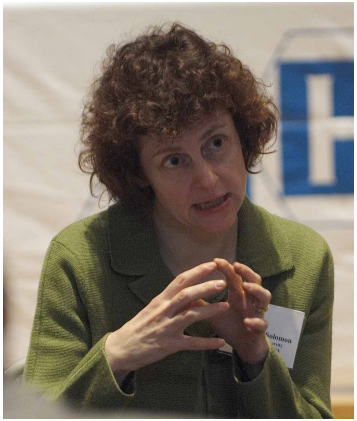


Figure 3: Miriam Solomon, Wikipedia

Medical Knowledge is “impoverished”

“Statistical fog” (Nicholas Rescher)

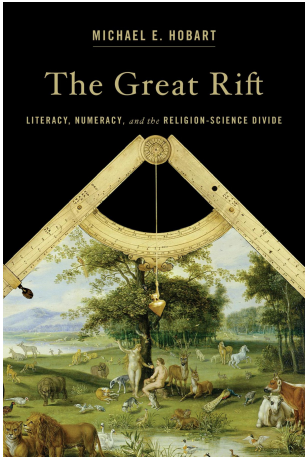
Figure 4: Wikipedia

How to move out of the fog?

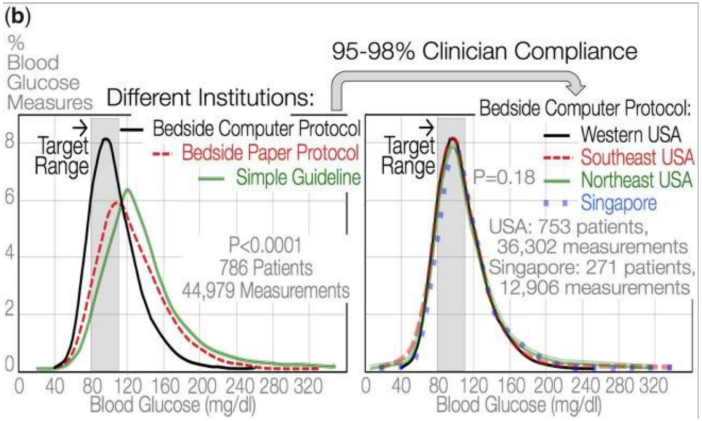
Be willing to prioritize mechanistic explanations

- "[M]any consider a good (or 'gold standard') diagnosis to be one that proposes a causal explanation. . . " (Lalumera and Fanti 2021)
- Valuing pluralistic sources of knowledge

Dangerous rifts



Improved mechanistic reasoning: Glucose management (Morris et al. 2021)



Improved mechanistic reasoning: Data assimilation

Research | September 01, 2020

Print

Data Assimilation in Medicine

By David Albers, Cecilia Diniz Behn, and George Hripcsak

Medicine is an ongoing forecasting process with sparse, inaccurate data. Practitioners rarely employ physiologic knowledge in actual medical practice because of limited data, imperfect models, and a disconnect between measurable variables and potential treatments. Nevertheless, the medical community is currently clamoring for effective employment of *precision medicine*, which involves providing the right treatment to the right patient; physiology can inspire the solution. We aim to apply current physiologic models—e.g., mechanistic ordinary differential equation (ODE) models of physiology—to medical data to produce useful forecasts. This approach requires fundamental changes to the way in which researchers merge models and data.

Consider the use of inference methods to estimate model states and parameters and evaluate the performance of predictions relative to actionable decisions. One can handle this coupling naturally via data assimilation (DA), a collection of deterministic and stochastic inference methods that estimate and forecast states and parameters of ODE-like models with individual patient data [7]. However, real patient data—which are often sparse and nonstationary—are particularly difficult to use, thus making both inference and model evaluation highly complex. Application to medicine is therefore not possible without model validation and verification that is anchored to clinical consequence. Mathematical innovations must address these realities, and the pursuit of methods that generate actionable knowledge within clinical-specified constraints provides an exciting and deep well of new problems.

Science Translational Medicine

Current Issue First release papers

HOME > SCIENCE TRANSLATIONAL MEDICINE > VOL. 11, NO. 473 > SEX DIFFERENCES IN GBM REVEALED BY ANALYSIS OF PATIENT IMAGING, TRANSCRIPTOM...

RESEARCH ARTICLE | CANCER

Sex differences in GBM revealed by analysis of patient imaging, transcriptome, and survival data

WEI YANG, NICOLE M. WARRINGTON, SARA J. TAYLOR, PAULA WHITMIRE, EDUARDO CARRASCO, KYLE W. SINGLETON, NINGYING WU, JUSTIN D. LATHIA, MICHAEL E. BERENS, ALBERT H. KIM, JILL S. BARNHOLTZ-SLOAN, KRISTIN R. SWANSON, JINGQIN LUO, AND JOSHUA B. RUBIN

fewer Authors Info & Affiliations

SCIENCE TRANSLATIONAL MEDICINE · 2 Jan 2019 · Vol 11, Issue 473 · DOI: 10.1126/scitranslmed.aao5253

“Medicine is the management of uncertainty” [1]

What happens when we mismanage uncertainty?

“Some of these potential negative outcomes that can arise from mismanaged diagnostic uncertainty in the clinical setting include the **epistemic (in the form of overdiagnosis or underdiagnosis), the pragmatic (in the form of increased financial and time costs), and the ethical (in the form of increased patient suffering).**” (Lalumera and Fanti 2021)

[1] Ted Shortliffe

What are all the sources of uncertainty in medical imaging?

And how might these guide research questions?

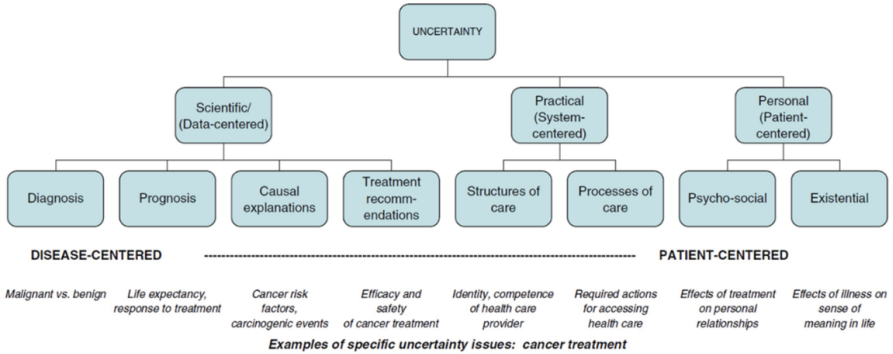


Fig. 2.2 **Different conceptions and sources of uncertainty from Han et al. (2011)**

2

Reducing uncertainty in medical imaging

Fig. 2.3 Five basic issues and questions to ask in order to reduce diagnostic uncertainty. (Adapted from Hofmann (2018))

3

[3] Lalumera and Fanti (2021)

Section 3

Parting thoughts

Fred Brooks' greatest contribution" (Kelly 2010)

Wired: What do you consider your greatest technological achievement?
Brooks: The most important single decision I ever made was to change the IBM 360 series from a 6-bit byte to an 8-bit byte, thereby enabling the use of lowercase letters. That change propagated everywhere.

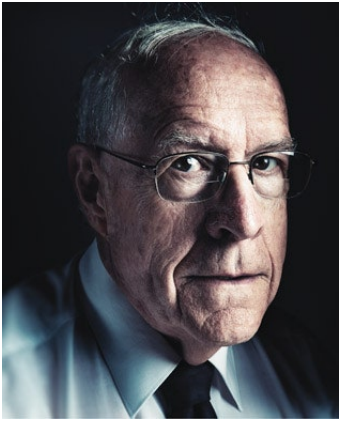


Figure 5: Wired

Fred Brooks on Science vs. Design

the scientist builds in order to study; the engineer studies in order to build. (Brooks 1996)

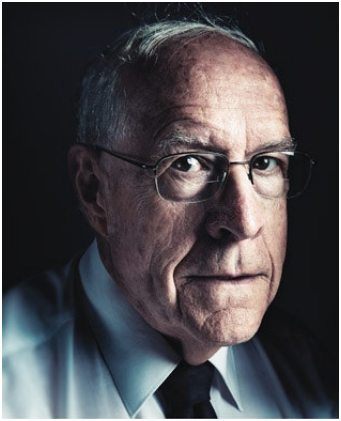


Figure 6: Wired

Fred Brooks on Science vs. Design

"[S]ciences legitimately take the discovery of facts and laws as a proper end in itself. A new fact, a new law is an accomplishment, worthy of publication. If we confuse ourselves with scientists, we come to take the invention (and publication) of endless varieties of computers, algorithms, and languages as a proper end. **But in design, in contrast with science, novelty in itself has no merit.** If we recognize our artifacts as tools, we test them by their usefulness and their costs, not their novelty." (Brooks 1996)

Fred Brooks on Science vs. Design

"If we perceive our role aright, we then see more clearly the proper criterion for success: **a toolmaker succeeds as, and only as, the users of his tool succeed with his aid.**"

Chuck Mistretta: Subtraction!

"Can you patent subtraction?"

"The DSA technique has been distributed worldwide and is still the gold standard against which the image quality of new angiographic techniques is measured. **Patent royalties from DSA presently rank second among all inventions in University of Wisconsin history.**"

Conclusions

- Values drive behavior, including research

Conclusions

- Values drive behavior, including research
- Our values are shaped by our habits

Conclusions

- Values drive behavior, including research
- Our values are shaped by our habits
- Entangled research: Disrupting our habits is important for transformative research

Conclusions

- Values drive behavior, including research
- Our values are shaped by our habits
- Entangled research: Disrupting our habits is important for transformative research
  - Quantification!

Conclusions

- Values drive behavior, including research
- Our values are shaped by our habits
- Entangled research: Disrupting our habits is important for transformative research
  - Quantification!
- Leverage pluralistic sources of disruption

Conclusions

- Values drive behavior, including research
- Our values are shaped by our habits
- Entangled research: Disrupting our habits is important for transformative research
  - Quantification!
- Leverage pluralistic sources of disruption
- Ideally, medical imaging research needs to...

Conclusions

- Values drive behavior, including research
- Our values are shaped by our habits
- Entangled research: Disrupting our habits is important for transformative research
  - Quantification!
- Leverage pluralistic sources of disruption
- Ideally, medical imaging research needs to...
  - Value impact as well as novelty

Conclusions

- Values drive behavior, including research
- Our values are shaped by our habits
- Entangled research: Disrupting our habits is important for transformative research
  - Quantification!
- Leverage pluralistic sources of disruption
- Ideally, medical imaging research needs to...
  - Value impact as well as novelty
  - Not all impactful work needs to be novel

Conclusions

- Values drive behavior, including research
- Our values are shaped by our habits
- Entangled research: Disrupting our habits is important for transformative research
  - Quantification!
- Leverage pluralistic sources of disruption
- Ideally, medical imaging research needs to...
  - Value impact as well as novelty
  - Not all impactful work needs to be novel
  - Be conducted in the context of the entire patient/population data ecosystem

Conclusions

- Values drive behavior, including research
- Our values are shaped by our habits
- Entangled research: Disrupting our habits is important for transformative research
  - Quantification!
- Leverage pluralistic sources of disruption
- Ideally, medical imaging research needs to...
  - Value impact as well as novelty
  - Not all impactful work needs to be novel
  - Be conducted in the context of the entire patient/population data ecosystem
- Ultimately, it's all about the patient

References I

Brooks, Frederick P. 1996. “The Computer Scientist as Toolsmith II.” *Commun. ACM* 39 (3): 61–68. https://doi.org/10.1145/227234.227243.

Chapman, B. E., H. P. Berty, and S. L. Schulthies. 2015. “Automated generation of directed graphs from vascular segmentations.” *J Biomed Inform* 56 (August): 395–405.

Fryback, D. G., and J. R. Thornbury. 1991. “The efficacy of diagnostic imaging.” *Med Decis Making* 11 (2): 88–94.

Kelly, Kevin. 2010. “Master Planner: Fred Brooks Shows How to Design Anything.” *Wired*. https://www.wired.com/2010/07/ff-fred-brooks/.

Lalumera, E., and S. Fanti. 2021. *Philosophy of Advanced Medical Imaging*. SpringerBriefs in Ethics. Springer International Publishing. https://books.google.com/books?id=ZvggEAAAQBAJ.

Millard, C., and P. Michael. 2017. *Destiny of the Republic: A Tale of Madness, Medicine and the Murder of a President*. Penguin Random House Audio Publishing Group. https://books.google.com/books?id=pOSNEAAAQBAJ.

References II

Morris, A. H., B. Stagg, M. Lanspa, J. Orme, T. P. Clemmer, L. K. Weaver, F. Thomas, et al. 2021. "Enabling a learning healthcare system with automated computer protocols that produce replicable and personalized clinician actions." *J Am Med Inform Assoc* 28 (6): 1330–44.

Noë, A. 2023. *The Entanglement: How Art and Philosophy Make Us What We Are*. Princeton University Press. https://books.google.com/books?id=gpCkEAAAQBAJ.

Solomon, M. 2015. *Making Medical Knowledge*. Oxford University Press. https://books.google.com/books?id=___dsBwAAQBAJ.

Willemink, M. J., D. Mastrodicasa, M. H. Madani, M. Codari, L. L. Chepelev, G. Mistelbauer, K. Hanneman, et al. 2023. "Inter-observer variability of expert-derived morphologic risk predictors in aortic dissection." *Eur Radiol* 33 (2): 1102–11.

References III

Wobben, Liana D., Marina Codari, Gabriel Mistelbauer, Antonio Pepe, Kai Higashigaito, Lewis D. Hahn, Domenico Mastrodicasa, et al. 2021. "Deep Learning-Based 3D Segmentation of True Lumen, False Lumen, and False Lumen Thrombosis in Type-b Aortic Dissection." In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 3912–15. https://doi.org/10.1109/EMBC46164.2021.9631067.

Yao, Z., W. Xie, J. Zhang, Y. Dong, H. Qiu, H. Yuan, Q. Jia, et al. 2021. "ImageTBAD: A 3D Computed Tomography Angiography Image Dataset for Automatic Segmentation of Type-B Aortic Dissection." *Front Physiol* 12: 732711.