4.7 ML meets biomedical Informatics

Prof William Hsu (UCLA) @ucalwillhsu
whsu@mednet.ucla.edu
https://hsu-lab.com

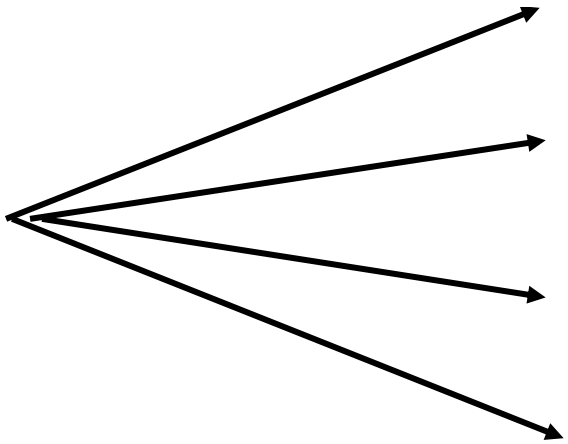
Learning objectives

- State the relationship between machine learning and the field of biomedical informatics
- Describe integration approaches for turning multimodal data into knowledge
- Identify barriers in transforming knowledge into practice

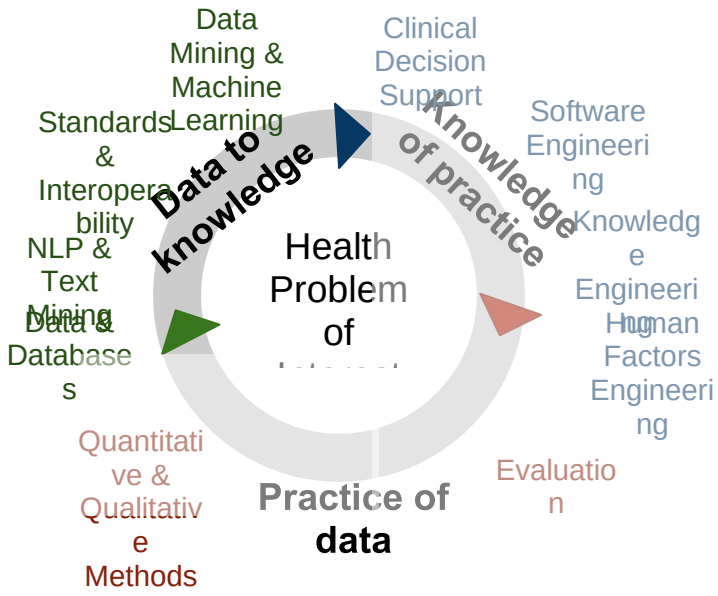
What is biomedical informatics?

- Biomedical informatics (BMI) is the interdisciplinary field that studies and pursues the effective use of biomedical data, information, and knowledge for scientific inquiry, problem solving, and decision making, motivated by efforts to improve human health.
    - IMIA https://imia-medinfo.org/wp/
    - AMIA https://amia.org/
- Goal: augmenting human knowledge
    - Not to replace humans

Subfield of biomedical informatics

<table>
<tr><td>Basic research</td><td colspan="2">  </td><td>Applied research informatics</td></tr>
<tr><td rowspan="4">Biomedical informatics: Methods, technologies, theories</td><td colspan="2" rowspan="4">  </td><td>Bioinformatics
Molecular and Cellular Processes</td></tr>
<tr><td>Imaging informatics
Tissues and Organs</td></tr>
<tr><td>Clinical informatics
Individuals (Patients)</td></tr>
<tr><td>Public health informatics
Populations and Society</td></tr>
</table>

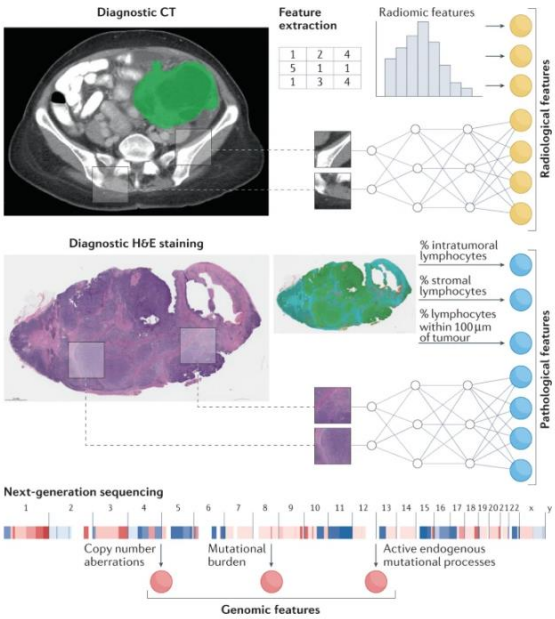
Data <-> Practice cycle



- Chen ES https://link.springer.com/chapter/10.1007/978-3-030-70558-9_2

Towards precision oncology

- Diagnosis process is increasingly multi-modal:
    1. Patient Presents with Symptoms
    2. PCP Orders Imaging Exam
    3. Imaging Study Performed
    4. Radiologist Interpretation Rendered
    5. Suspicious Mass Notes, Biopsy Ordered
    6. Biopsy Taken, Analyzed by Pathology
    7. Pathology Result Rendered
    8. Downstream Care Provided
- **Extracting features from multimodal data**



- Boehm KM et al.
  - https://dx.doi.org/10.1038/s41568-021-00408-3

- Radiomics features

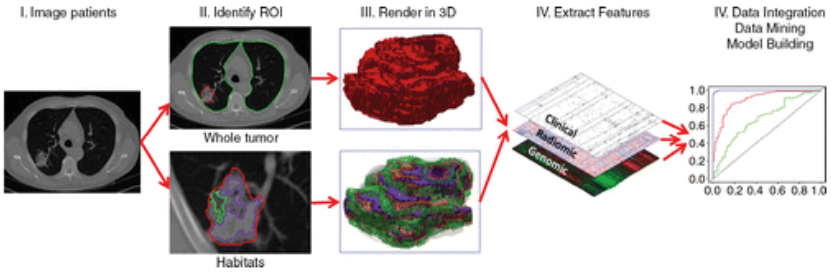


Figure 1: Flowchart shows the process of radiomics and the use of radiomics in decision support. Patient work-up requires information from disparate sources to be combined into a coherent model to describe where the lesion is, what it is, and what it is doing. Radiomics begins with acquisition of high-quality images. From these images, a region of interest *(ROI)* that contains either the whole tumor or subregions (ie, habitats) within the tumor can be identified. These are segmented with operator edits and are eventually rendered in three dimensions *(3D)*. Quantitative features are extracted from these rendered volumes to generate a report, which is placed in a database along with other data, such as clinical and genomic data. These data are then mined to develop diagnostic, predictive, or prognostic models for outcomes of interest.

- Gilies RJ et al., 2015
  - https://pubs.rsna.org/doi/full/10.1148/radiol.2015151169
- Families of radiomic features:

Table 2 Feature families and required image processing. For each feature family, the number of features in the document, the required input of a morphological (morph.) and/or intensity (int.) ROI mask, as well as the requirement of image discretisation (discr.) is provided. [a] The entire image volume should be available when computing local intensity features. [b] Image discretisation for the intensity-volume histogram is performed with finer discretisation than required for e.g. textural features.

		ROI mask		
Feature family	**count**	**morph.**	**int.**	**discr.**
morphology	29	✓	✓	×
local intensity	2	×	✓ [a]	×
intensity-based statistics	18	×	✓	×
intensity histogram	23	×	✓	✓
intensity-volume histogram	5	×	✓	✓ [b]
grey level co-occurrence matrix	25	×	✓	✓
grey level run length matrix	16	×	✓	✓
grey level size zone matrix	16	×	✓	✓
grey level distance zone matrix	16	✓	✓	✓
neighbourhood grey tone difference matrix	5	×	✓	✓
neighbouring grey level dependence matrix	17	×	✓	✓

    - More info: https://ibsi.readthedocs.io/en/latest/03_Image_features.html
    - Required input of a morphological and/ or intensity ROI mask
    - Requirement of image discretization (discr.)
- Textures:

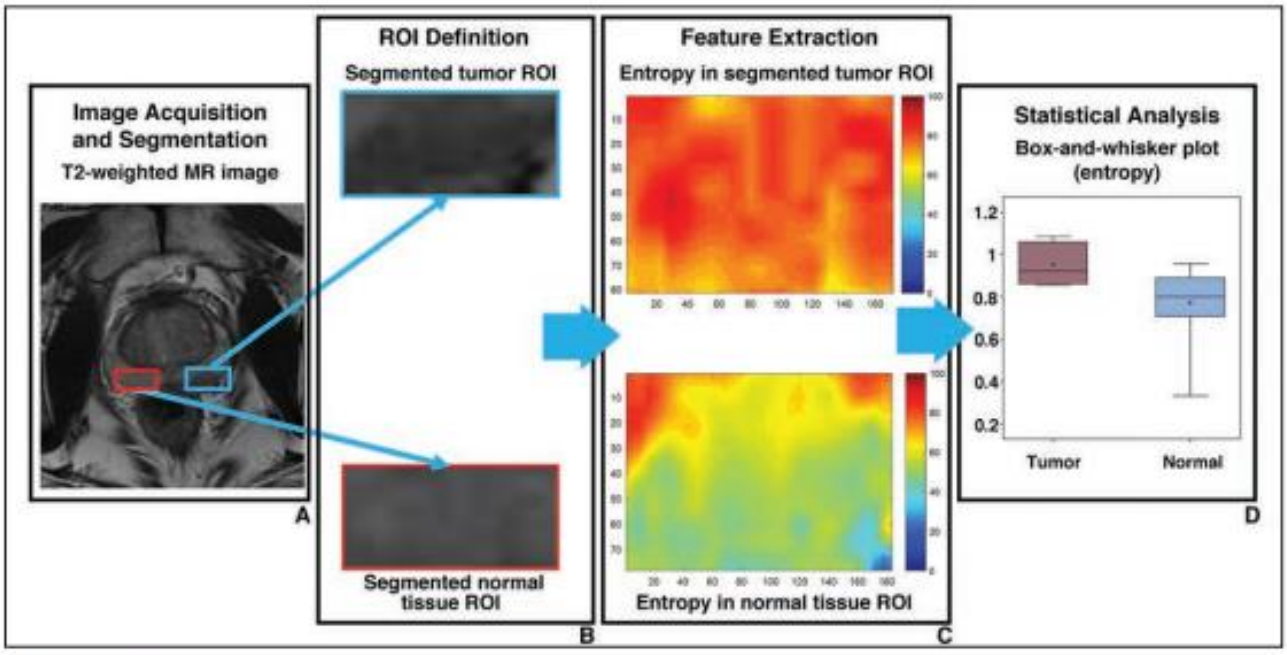
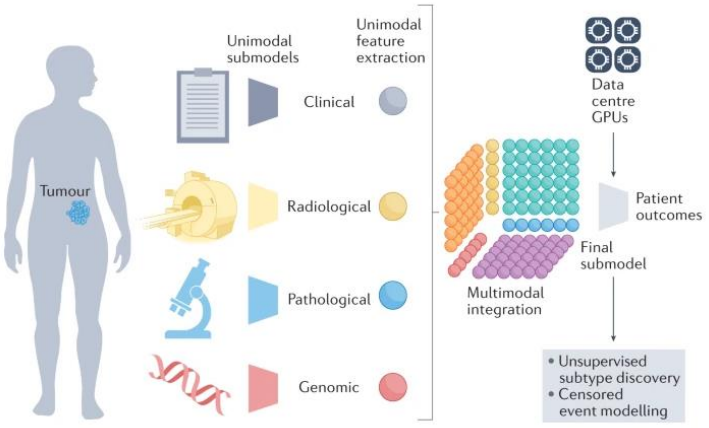


Fig. 2—Schematic of typical radiomics workflow showing four basic modules.
A and **B**, 72-year-old man with elevated prostate specific antigen level. T2-weighted axial MR image (**A**) of prostate reveals hypointense focal area in left posterior peripheral zone that was biopsy-proven focus of prostate adenocarcinoma. Two ROIs are marked on image: one on normal tissue (*red rectangle*, **A** and **B**) and one on tumor (*blue rectangle*, **A** and **B**) that underwent subsequent feature extraction.
C, Feature extraction images of same patient shown in **A** and **B** present entropy measurements within ROIs on tumor (*top*) and normal tissue (*bottom*). *x*- and *y*-axes represent image coordinates of segmented ROI image.
D, Box-and-whisker plot shows findings of statistical analysis performed after feature extraction to assess performance of radiomic metrics. Three horizontal lines shown from top to bottom in boxes denote three quartiles (75th percentile, 50th percentile, and 25th percentile, respectively) and top and bottom whiskers denote maximum and minimum values, respectively. Plus signs within boxes denote mean values.

- https://www.ajronline.org/doi/pdf/10.2214/AJR.18.20624

- Extracting features from multimodal data: biomedical data are multi-scale, multi-modal

Fig. 2: Multimodal models integrate features across modalities.



Submodels extract unimodal features from each data modality. Next, a multimodal integration step generates intermodal features – a tensor fusion network (TFN) is indicated here[86]. A final submodel infers patient outcomes. GPU, graphics processing unit.

- Boehm KM et al.
  - https://dx.doi.org/10.1038/s41568-021-00408-3

Basic types of multimodal integration

- Feature fusion:

    - Combine features across sources to generate a single feature vector for classification
- Decision fusion:
    - Classify features from each source independently then aggregate the results
- More information: https://ieeexplore.ieee.org/document/5192987
- An example of general approach:
    - https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9684949

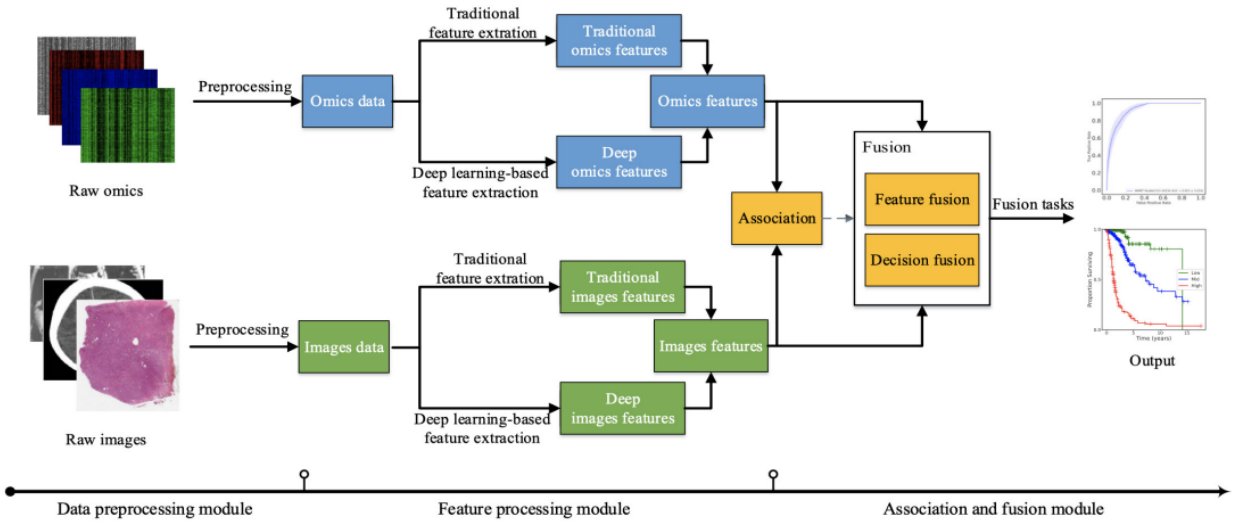


Fig. 1. **Flowchart of omics-imaging fusion.**

Fusion level	Fusion type	Fusion method
Feature fusion	Fusion through association	Canonical correlation analysis
		Independent component analysis
		Nonnegative matrix factorization
	Direct fusion	Vector concatenation
		Sparse representation
		Multiple kernel learning
		Deep neural network
		Graph-based network

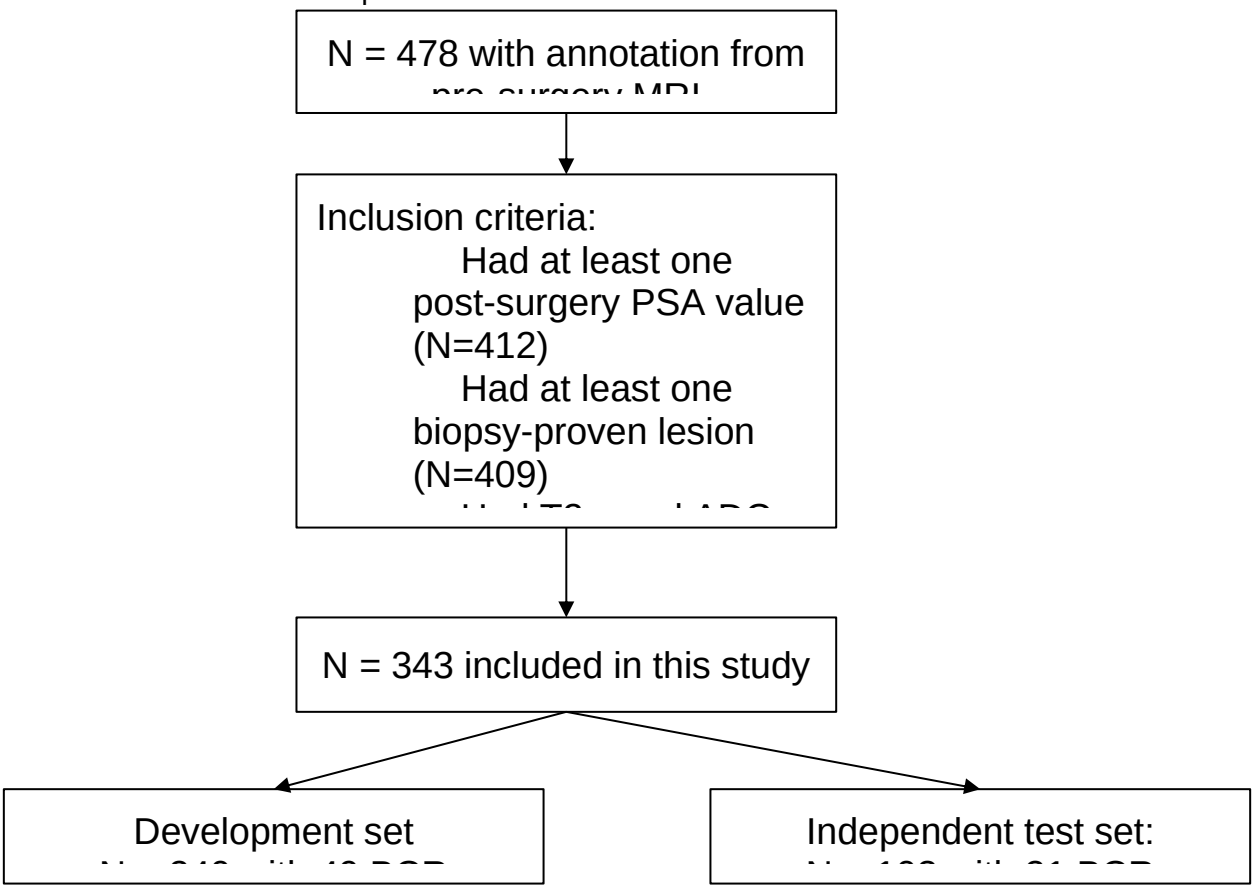
Decision fusion	Ensemble learning	Random forest

- Example: Prostate cancer (PCa) aggressiveness
    - Identifying patients who are likely to experience biochemical recurrence (BCR) after radical prostatectomy by fusing radiomic features extracted from the lesions shown in T2w and ADC
    - Existing works:

Name	Data	Treatment	Imaging	Radiomics	Feature fusion
Gnep et al., 2017 https://onlinelibrary.wiley.com/doi/10.1002/jmri.25335	N=47 1 center	RTx	T2w, ADC	144 intensity, Texture features	Direct concatenation
Shiradkar et al., 2018 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6222024/	N=120 2 centers	RTx or RP	T2w, ADC	600 intensity, Texture features	Direct concatenation
Dinis Fernandes et al., 2018 https://phiro.science/article/S2405-6316(17)30085-4/fulltext	N=120 1 center	RTx	T2w	254 texture features	NA
Zhong et al., 2020 https://www.frontiersin.org/articles/10.3389/fonc.2020.00731/full	N=91 1 center	RTx	T1, T2, DWI	1536 deep features	Direct concatenation
Bourbonne et al., 2020 https://www.mdpi.com/2072-6694/12/4/814	N=195 2 centers	RP	T2w, ADC	27,376 intensity, Shape features	Direct concatenation

Li et al., 2020 https://www.thelancet.com/journals/ebiom/article/PIIS2352-3964(20)30539-9/fulltext	N=198 4 centers	RP	T2w, ADC	200 intensity, Shape features	Direct concatenation

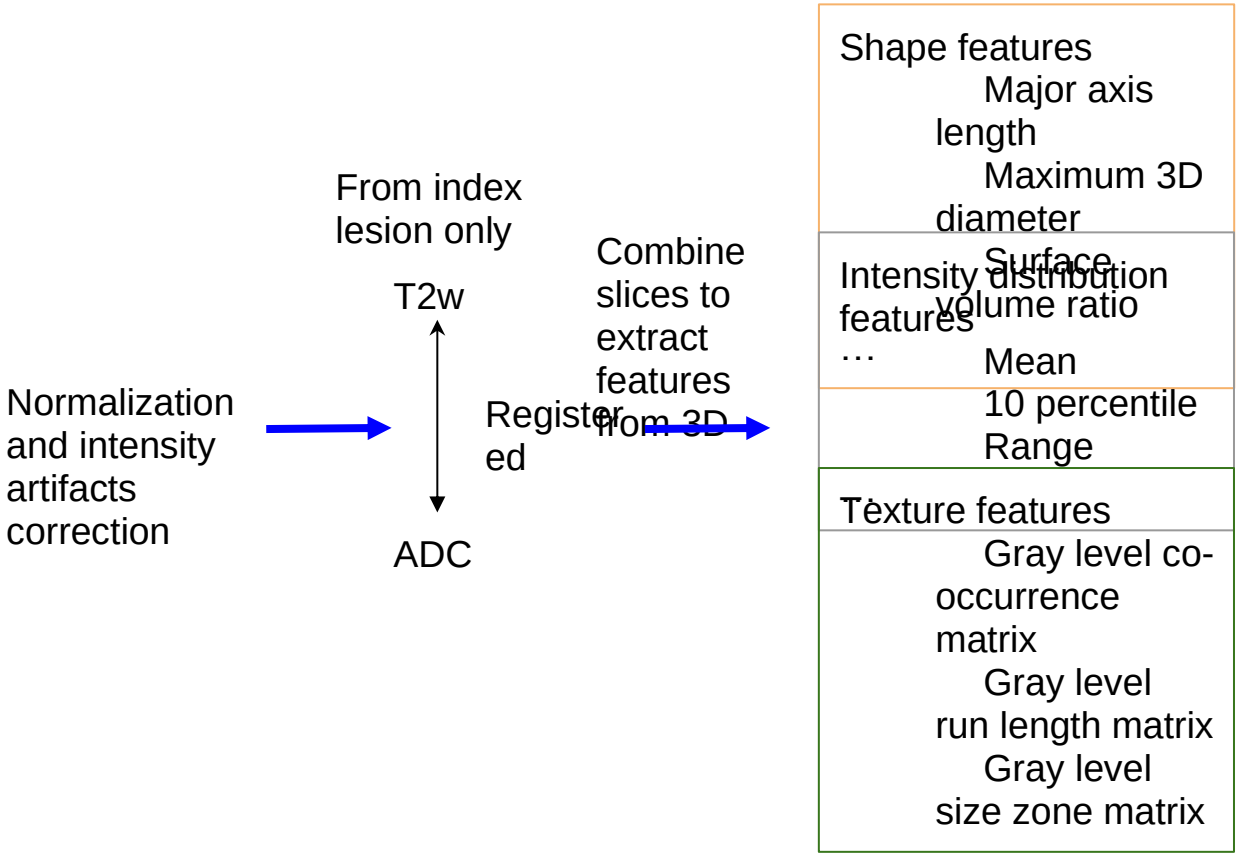
- Hypothesis
  - Fused radiomic features from T2w and ADC are more prognostic of BCR than:
    - 1) T2w radiomic features alone
    - 2) ADC radiomic features alone
    - 3) Direct feature combination from T2w and ADC
    - 4) Clinical assessments of T 2 score, T 2 shape, T 2 signal, average ADC values
- Dataset and patient selection

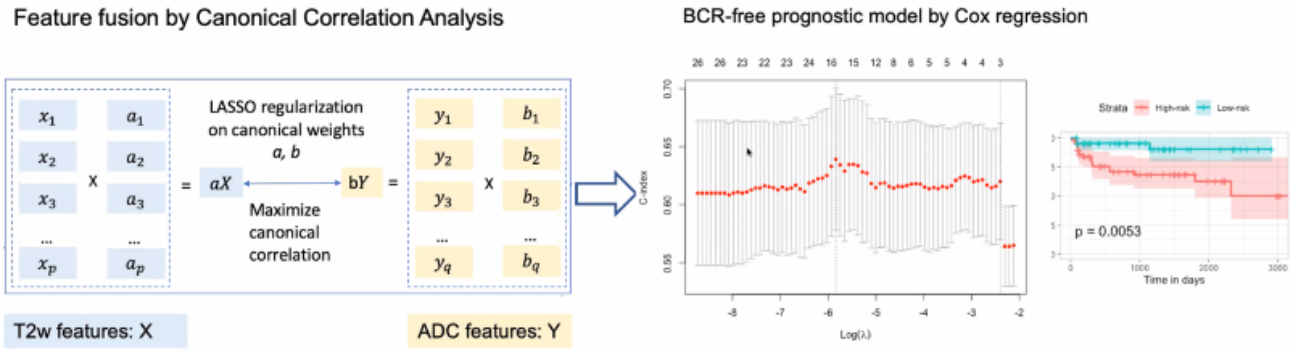


  - Median follow up for all 343 cases:
    - 27.9 (0.2-110.1) months
  - Median follow up for 70 BCR cases:
    - 12.1 (0.-92.9) months

- Median follow up for 273 non-BCR cases:
  - 31.4 (1.9-110.1) months

- Radiomics feature extraction

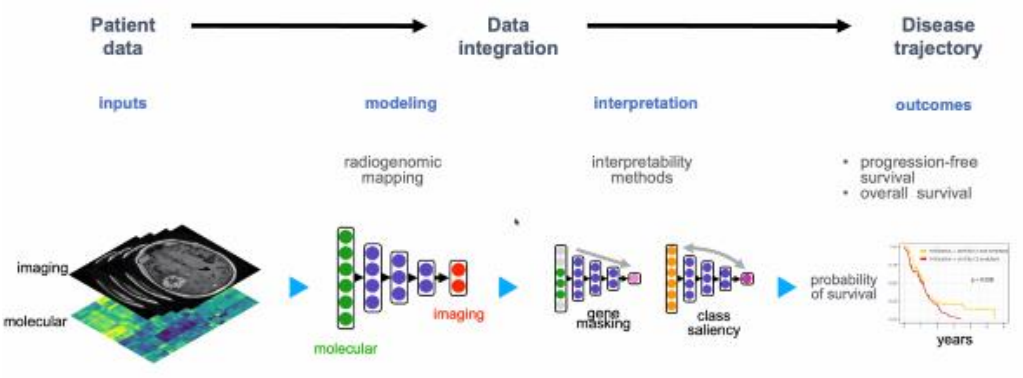




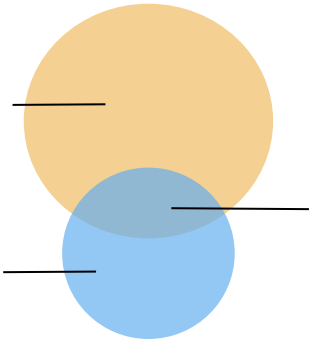
- Results
  - CCA-fused features outperform clinical, radiomic T2, Radiomic ADC, and Concatenate radiomic T2 and ADC
  - On test set, able to separate high risks and low risks groups

Example: relating imaging to gene expression

- https://academic.oup.com/bioinformatics/article/36/11/3537/5758261



- Data:



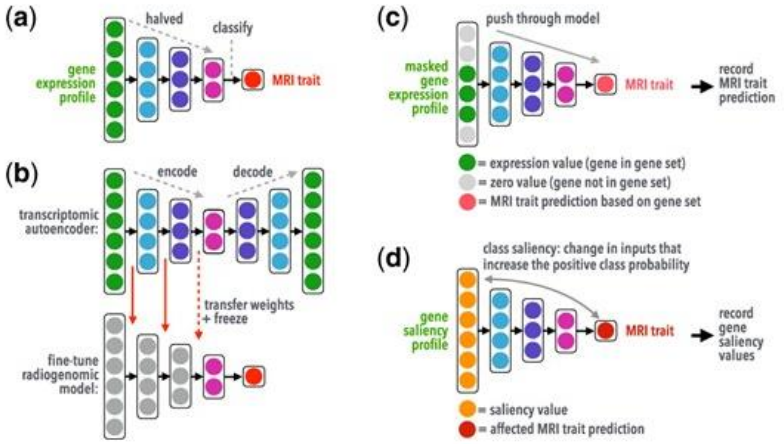
TCGA
n = 528 GBM patients
Microarray
Untreated, primary tumours
Multiple image modalities

TCIA
n = 262 GBM patients

n = 109 paired, radiogenomic samples
Gene expression profile and pre-operative MR imaging

- Radiogenomic neural network:

Fig. 2.



Open in new tab Download slide

Illustration showing(**a**) the radiogenomic neural network's architecture, (**b**) transfer learning using a deep transcriptomic autoencoder, and interpretation methods using (**c**) gene masking and (**d**) gene saliency. Pretrained weights learned in the autoencoder were transferred to a radiogenomic model, where weights were frozen (non-trainable, long red arrows) and/or fine-tuned (trainable, dashed red arrow) during radiogenomic training. (Color version of this figure is available at *Bioinformatics* online.)

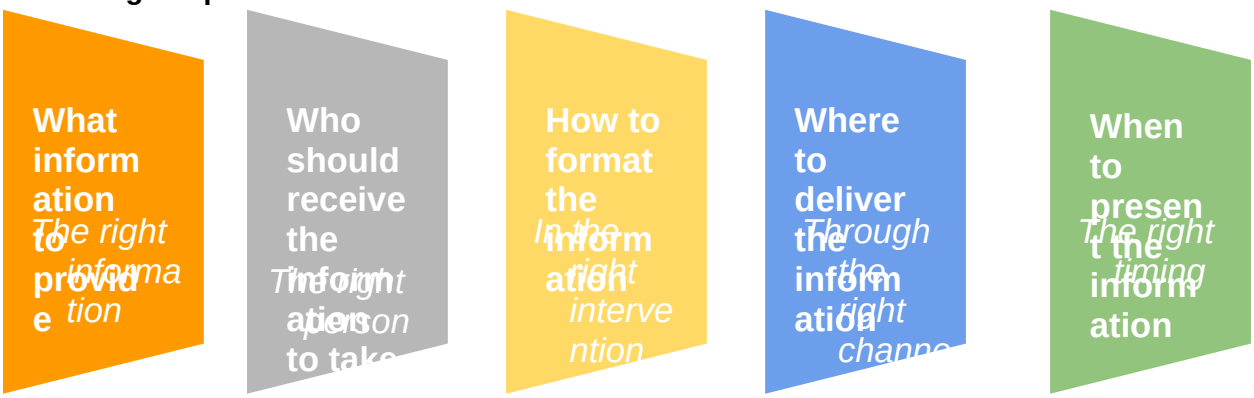
- Measures performance of model when only using genes from a gene set
- Use AUC as strength of association
- Attach weights to genes based on their importance in predicting a class label
- Use of GSEA to determine what gene sets (pathways, biological processes etc.) were at the top of ranked genes
- Model interpretation

Transcriptomic drivers

Theme	Gene set (collection,[a] query[b])	AUC	AP	See also
Enhancing				
Growth/death	Growth (G, *PTEN*)	0.86	0.84	
	Sensory organ development (G, *EGFR*, *KCNK3*)	0.85	0.84	Gutman *et al.* (2013); Jamshidi *et al.* (2014)
Immune system	IL2/STAT5 signaling (H)	0.77	0.76	
	Complement system (H)	0.79	0.75	
	Activation of immune response (G, *PTEN*)	0.90	0.89	
	Leukocyte and lymphocyte activation (G, *PIK3R1*)	0.85[c]	0.84[c]	
	Immune effector process (G, *PIK3CA*)	0.87	0.84	
Hormones	Early and late responses to estrogen (H)	0.79[c]	0.73[c]	
	Response to steroid hormone (G, *RBI*)	0.88	0.88	
	Reg. of hormone levels (G, *PARK2*)	0.87	0.84	
ECM	Related to ECM proteins (C, ECM)	0.81[c]	0.75[c]	Diehn *et al.* (2008)
	Apical junction (H)	0.80	0.75	
Vasculature	Heme metabolism (H)	0.77	0.65	
	Vasculature and heart development (G, *LTBP1*)	0.80[c]	0.79[c]	Gutman *et al.* (2013)
Kinase activity	Multiple (G, *EGFR*, *LTBP1*, *KCNK3*)	0.87[c]	0.85[c]	Gutman *et al.* (2013); Jamshidi *et al.* (2014)

- Identified the most predictive gene sets
- Gene sets are related to growth, vasculature, immune system processes, and involved EGFR
- Gene sets align with previously identified associations

Knowledge to practice



Algorithm results are only useful if they are accessible

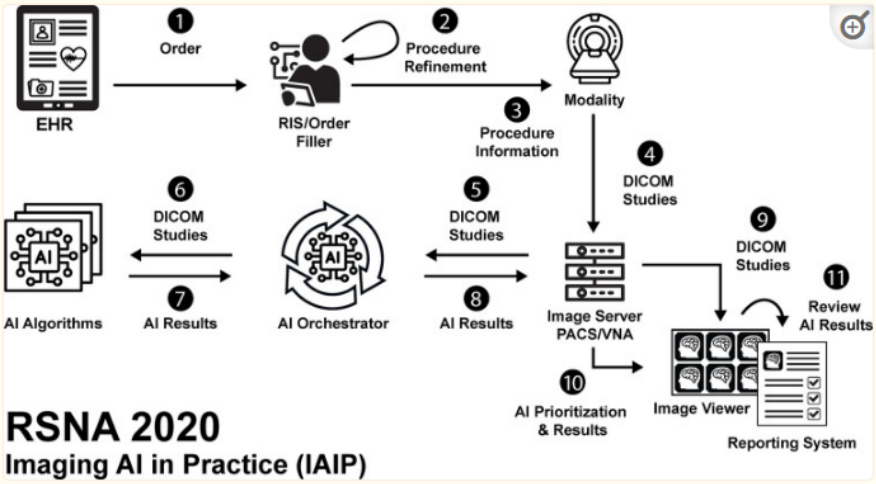


Figure 1:

Image acquisition workflow diagram. **1**, An order is placed in the electronic health record (EHR) and transmitted via Health Level 7 version 2 or Fast Healthcare Interoperability Resources to the radiology information system (RIS) using the Integrating the Healthcare Enterprise Order Filler profile. **2**, The exact procedure may be refined or protocolled in the RIS prior to **3**, being communicated to the modality via Digital Imaging and Communications in Medicine (DICOM) Modality Worklist. **4**, The images are acquired by the scanner and sent with associated metadata in DICOM format to the image server, which may then **5**, forward the DICOM images to an AI orchestrator. **6**, The AI orchestrator then de-identifies the images if necessary and sends them to the appropriate AI algorithm(s). **7**, The AI algorithms process the images and return the AI results to the AI orchestrator, which then **8**, associates the AI results with the appropriate imaging study and sends the results back to the image server. **9**, **10**, The image server will forward the DICOM studies and AI results to the image viewer, where **11**, the studies and associated AI results are reviewed, and a report is generated. AI = artificial intelligence, PACS = picture archiving and communication system, RSNA = Radiological Society of North America, VNA = vendor-neutral archive.

- https://pubs.rsna.org/doi/epdf/10.1148/ryai.2021210152

Interaction between algorithm and human factors

Table 1 Eight human factors and ergonomics principles for healthcare AI

Situation awareness	Design options need to consider how AI can support, rather than erode, people's situation awareness.
Workload	The impact of AI on workload needs to be assessed because AI can both reduce as well as increase workload in certain situations.
Automation bias	Strategies need to be considered to guard against people relying uncritically on the AI, for example, the use of explanation and training.
Explanation and trust	AI applications should explain their behaviour and allow users to query it in order to reduce automation bias and to support trust.
Human–AI teaming	AI applications should be capable of good teamworking behaviours to support shared mental models and situation awareness.
Training	People require opportunities to practise and retain their skill sets when AI is introduced, and they need to have a baseline understanding of how the AI works. Attention needs to be given to the design of effective training that is accessible and flexible. Staff should be provided with protected time to undertake training during their work hours.
Relationships between people	The impact on relationships needs to be considered, for example, whether staff will be working away from the patient as more and more AI is introduced.
Ethical issues	AI in healthcare raises ethical challenges including fairness and bias in AI models, protection of privacy, respect for autonomy, realisation of benefits and minimisation of harm.

- https://doi.org/10.1136%2Fbmjhci-2021-100516

Holistic view of an algorithm

Table 1 Checklist of points to consider when assessing a commercial AI solution in radiology

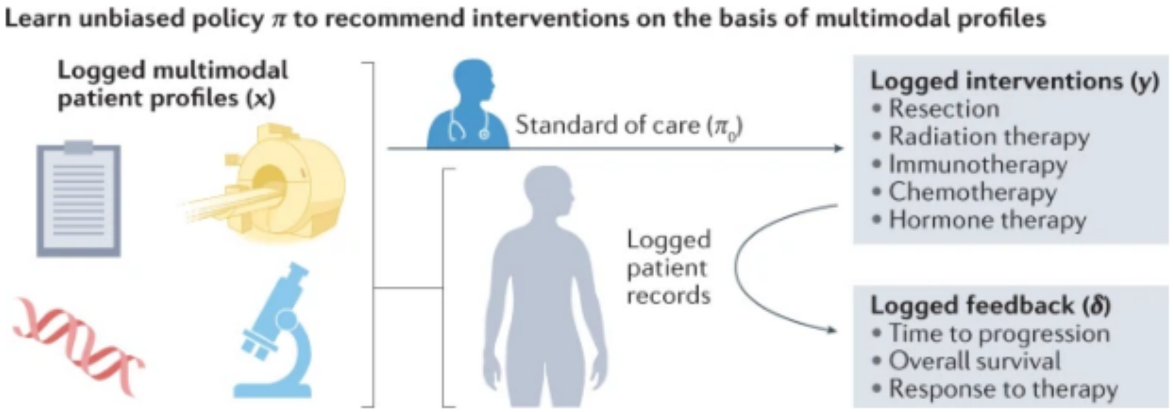
1. Relevance	**1.1. What problem is the application intended to solve, and who is the application designed for?** *Define the scope of application; end-users; research vs. clinical use; usage as double reader, triage, other; outputs (diagnosis, prognosis, quantitative data, other), indications and contra-indications* **1.2 .What are the potential benefits, and for whom?** *Consider benefits for patients, radiologists/referring clinicians, institution, society* **1.3. What are the risks associated with the use of the AI system?** *Consider risks of misdiagnosis (including legal costs), of negative impact on workflow, of negative impact on quality of training*
2. Performance and validation	**2.1. Are the algorithm's design specifications clear?** *Check robustness to variability of acquisition parameters; identify features (radiomics) or network architecture (deep learning) used* **2.2. How was the algorithm trained?** *Assess population characteristics and acquisition techniques used, labeling process, confounding factors, and operating point selection* **2.3. How has performance been evaluated?** *Check proper partitioning of training/validation/testing data, representativeness and open availability of data. Assess human benchmarks, application scope during evaluation, source of clinical validation* **2.4. Have the developers identified and accounted for potential sources of bias in their algorithm?** *Assess training data collection, bias evaluation, stratification analyses* **2.5. Is the algorithm fixed or adapting as new data comes in?** *Check whether user feedback is incorporated, if regulatory approval is maintained, and if results are comparable with previous versions. **
3. Usability and integration	**3.1. How can the application be integrated into your clinical workflow?** *Consider integration with your information technology (IT) platform, check for compliance with ISO usability standards, consider issues related to practical management of the software* **3.2. How exactly does the application impact the workflow?** *Identify modifications to bring to your current workflow, identify roles in the new workflow (physicians and non-physicians)* **3.3. What are the requirements in terms of information technology (IT) infrastructure?** *Consider on-premise vs. cloud solutions. Identify requirements in terms of hardware and network performance, consider network security issues* **3.4. Interoperability - How can the data be exported for research and other purposes?** *Check whether the export formats are suitable* **3.5. Will the data be accessible to non-radiologists (referring physicians, patients)?** *Check whether the form of the output is suitable for communication with patients/referring physicians* **3.6. Are the AI model's results interpretable?** *Check whether and which interpretability tools (i.e. visualization) are used*
4. Regulatory and legal aspects	**4.1. Does the AI application comply with the local medical device regulations?** *Check whether the manufacturer obtained regulatory approval from the country where the application will be used (CE, FDA, UKCA, MDSAP, or other local guidance), and for which risk class* **4.2. Does the AI application comply with the data protection regulations?** *Check whether the manufacturer complies with local data protection regulations and provides contractual clauses protecting patient's data*
5. Financial and support services considerations	**5.1. What is the licensing model?** *Assess one-time fee vs. subscription models, total costs, scalability* **5.2. How are user training and follow-up handled?** *Check whether training sessions are included and at which conditions further training can be obtained* **5.3. How is the maintenance of the product ensured?** *Check whether regular maintenance is included, assess the procedure during downtime and for repair* **5.4. How will potential malfunctions or erroneous results be handled?** *Assess the procedure in the event of malfunction and post market surveillance and follow-up*

* Note that at the time of writing of these guidelines, no adaptative AI application exists on the market.

- https://link.springer.com/article/10.1007/s00330-020-07684-x

Generating new data for future improvement

Fig. 5: Recommender systems could learn from retrospective data to assist in clinical decision-making.



Logged health-care data comprise multimodal patient contexts x, interventions y based on the standard of care (π_0) and feedback δ based on the outcome of the intervention. Learning from such data is challenging because of the lack of two-arm design and the biased data based on the changing standard of care. Counterfactual recommender systems learn theoretically guaranteed unbiased policies from these data. The validated policy π can then be applied prospectively to support physicians' management decisions.

- Boehm KM et al.
  - https://dx.doi.org/10.1038/s41568-021-00408-3

Getting started: datasets

- IVY GAP Atlas Project
  - https://glioblastoma.alleninstitute.org/
- Cancer Research Data Commons
  - https://datacommons.cancer.gov/
- Alzheimer's Disease Neuroimaging Initiative
  - https://adni.loni.usc.edu/
- UCLA Integrated Diagnostics Shared Resource
  - https://idx.medsch.ucla.edu/

Caveats

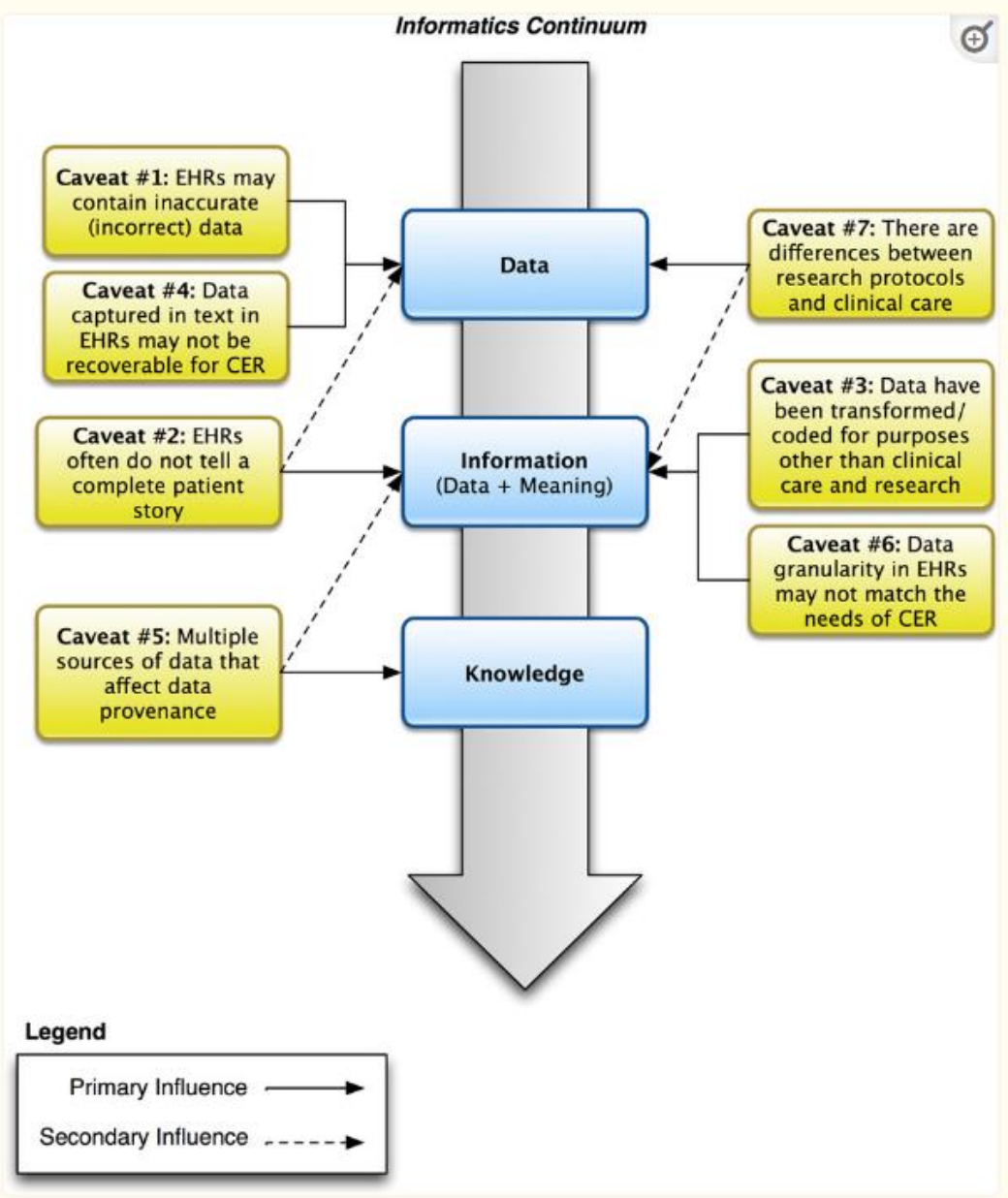
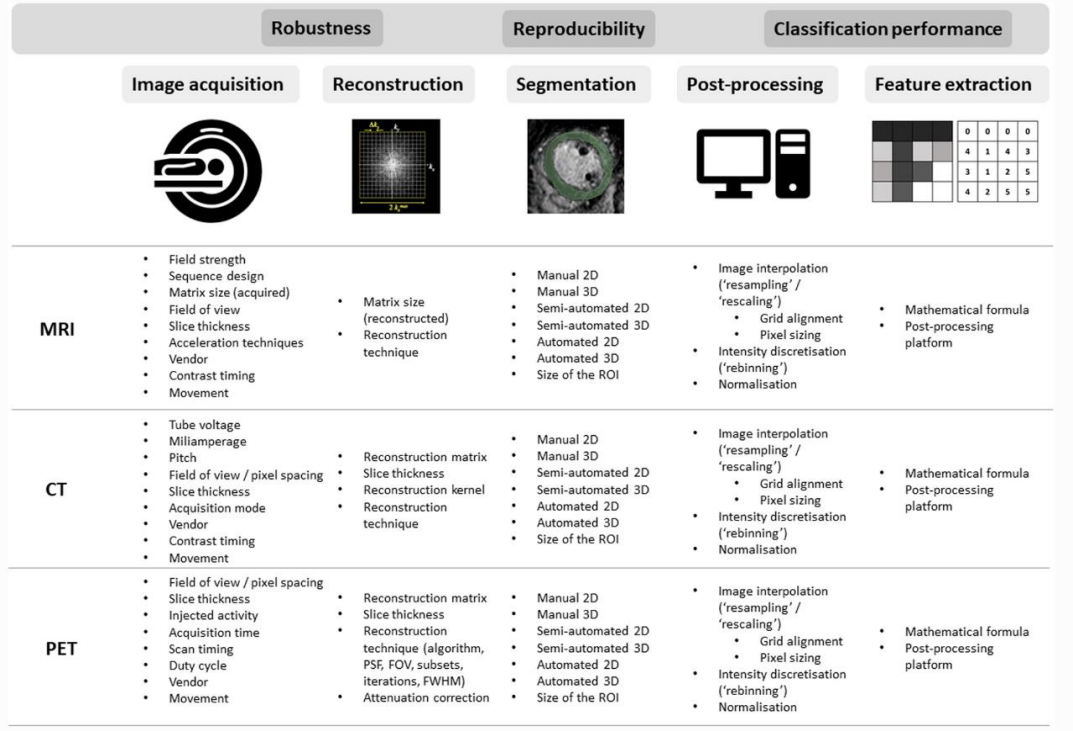


Figure 2

Relationship between central tenets of biomedical informatics and data-centric caveats pertaining to the design and conduct of comparative effectiveness research. Solid lines indicate the primary influencers of the caveats, while the secondary influencers are indicated with dashed lines.

- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3748381/

Factors that influence radiomics stability



	Robustness		Reproducibility		Classification performance
	Image acquisition	Reconstruction	Segmentation	Post-processing	Feature extraction

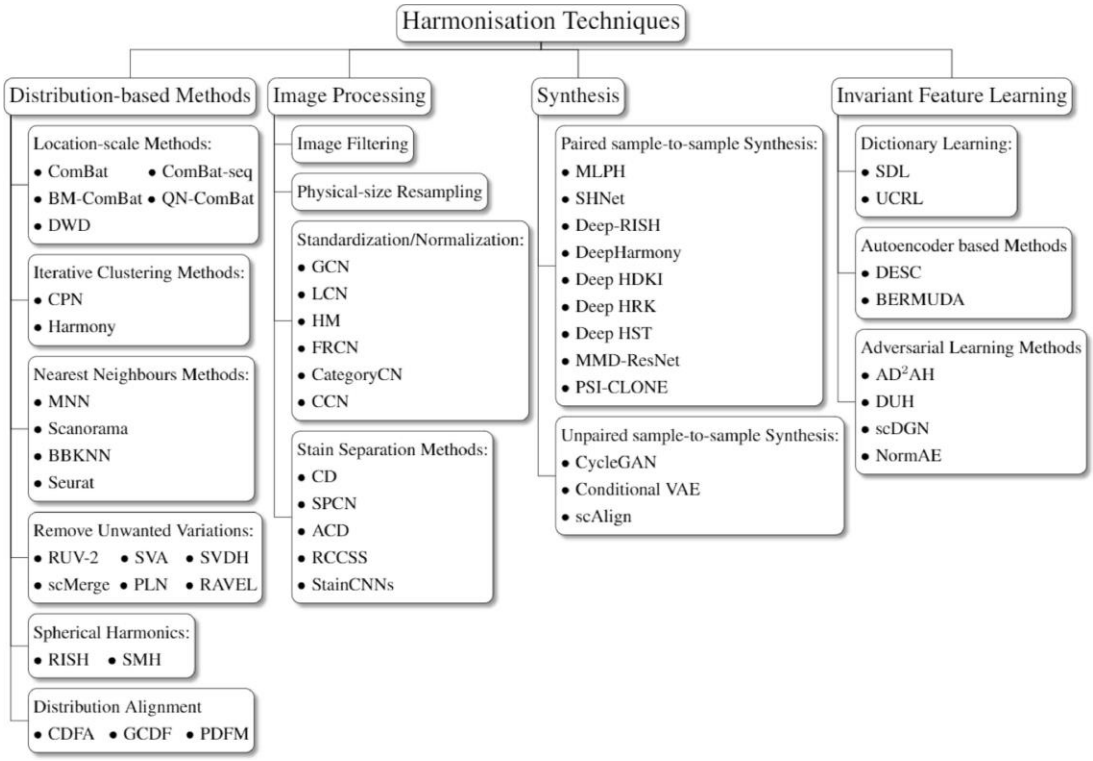
MRI	• Field strength
	• Sequence design
	• Matrix size (acquired)
	• Field of view
	• Slice thickness
	• Acceleration techniques
	• Vendor
	• Contrast timing
	• Movement
		• Matrix size (reconstructed)
		• Reconstruction technique
			• Manual 2D
			• Manual 3D
			• Semi-automated 2D
			• Semi-automated 3D
			• Automated 2D
			• Automated 3D
			• Size of the ROI
				• Image interpolation ('resampling' / 'rescaling')
				  • Grid alignment
				  • Pixel sizing
				• Intensity discretisation ('rebinning')
				• Normalisation
					• Mathematical formula
					• Post-processing platform

CT	• Tube voltage
	• Miliamperage
	• Pitch
	• Field of view / pixel spacing
	• Slice thickness
	• Acquisition mode
	• Vendor
	• Contrast timing
	• Movement
		• Reconstruction matrix
		• Slice thickness
		• Reconstruction kernel
		• Reconstruction technique
			• Manual 2D
			• Manual 3D
			• Semi-automated 2D
			• Semi-automated 3D
			• Automated 2D
			• Automated 3D
			• Size of the ROI
				• Image interpolation ('resampling' / 'rescaling')
				  • Grid alignment
				  • Pixel sizing
				• Intensity discretisation ('rebinning')
				• Normalisation
					• Mathematical formula
					• Post-processing platform

PET	• Field of view / pixel spacing
	• Slice thickness
	• Injected activity
	• Acquisition time
	• Scan timing
	• Duty cycle
	• Vendor
	• Movement
		• Reconstruction matrix
		• Slice thickness
		• Reconstruction technique (algorithm, PSF, FOV, subsets, iterations, FWHM)
		• Attenuation correction
			• Manual 2D
			• Manual 3D
			• Semi-automated 2D
			• Semi-automated 3D
			• Automated 2D
			• Automated 3D
			• Size of the ROI
				• Image interpolation ('resampling' / 'rescaling')
				  • Grid alignment
				  • Pixel sizing
				• Intensity discretisation ('rebinning')
				• Normalisation
					• Mathematical formula
					• Post-processing platform

- https://insightsimaging.springeropen.com/articles/10.1186/s13244-020-00887-2

Methods for normalization



- https://www.sciencedirect.com/science/article/pii/S156625352200015X

Takeaways

- Biomedical informatics encompasses the entire cycle from data to knowledge to practice to new data
- Machine learning, complemented by frameworks for model training, rigorous validation, and deployment with human factors, is critical to effectively harness data
- The case for multimodal data fusion
- Different modalities carry complementary (but sometimes conflicting) information
- Incorporating information across biological scales and modalities can lead to improved predictions of prognosis and treatment response
- Challenges and opportunities
- Availability of multimodal datasets
- Further investigation into how multi-modal features relate, model interpretability