Figure 2. Conceptual framework for AI applications in biomedical domain

Learning strategy	Learning goal		Examples	
Supervised learning	Classification	Predict a categorical variable	**Prostate cancer classification in biopsy WSIs** Pathologists improved their performance by AI assistance that classified biopsies and pinpointed suspicious areas	**Cardiovascular disease classification on heart sound recordings** The CardioXNet algorithm classifies phonocardiograms into five classes of cardiac auscultation, with high accuracy on challenge datasets
	Regression	Predict a continuous variable	**Joint space width measure in osteoarthritis knee radiographs** The FDA-approved clinical support tool uses deep learning to determine the minimum joint space width, while also detecting osteoarthritis	**Chemotherapy dose recommendation with CURATE.AI** The AI platform predicts the optimal dosing in a combination regimen, leading to increased efficacy and patient safety.
Unsupervised learning	Clustering	Identify similar groups of observations in unlabeled data	**Identification of wheezing phenotypes in asthmatic children** Latent class analysis of birth cohort data identified wheezing types, consistent with another cohort's study, while also detecting a new type	**Dietary pattern analysis using principal component analysis** PCA identified five dietary patterns based on questionnaire and health data, some with significantly increased risk for Crohn's disease and ulcerative colitis
	Dimensionality reduction	Transform data from a high-dimensional into a low-dimensional space	**Multi-omics dimension reduction for ovarian cancer analysis** DL is used to integrate and compress genomics, transcriptomics and epigenomics into latent representations for improved downstream analysis	**Text analysis on electronic health records data** Using DL to create low-dimensional latent vectors from large-scale and diverse text data, which can be used for further clustering to find disease subtypes Abc
	Generation	Generate new data examples based on patterns in learning data	**Accelerated CS-MRI reconstruction using generative AI** RefineGAN uses generative adversarial networks for faster and more accurate reconstruction of CS-MRI, outperforming current benchmarks	**Histological cancer image generation for data size increase** Synthetic images were visually indistinguishable from real images, assessed by trained pathologies, while also improving a cancer classification model
Semi-supervised learning	*All of the above*	Can be all of the above	**Ventricular hypertrophy detection based on partly labelled ECGs** A semi-supervised GAN architecture uses mostly unlabelled and a small portion of labelled training data, delivering still 92% detection accuracy	**Semi-supervised cancer detection on pathological images** Semi-supervised learning achieves similar diagnostic performance to fully labeled datasets on multiple cancer types, reducing costly annotations
Reinforcement learning	Decision-making	Continuous improvement of a task to make increasingly better decisions	**Robot-assisted guidewire navigation for coronary intervention** The autonomous guidewire navigation can reach all target locations in arteries, reaching over 98% accuracy in 2D and 3D setting	**Clinical support for adaptive radiotherapy** Use individual patient dose response to improve clinical decision-making, showing 10% improvement potential compared to unaided clinical practice.

Data modality Tabular Time series Image Audio Abc Text