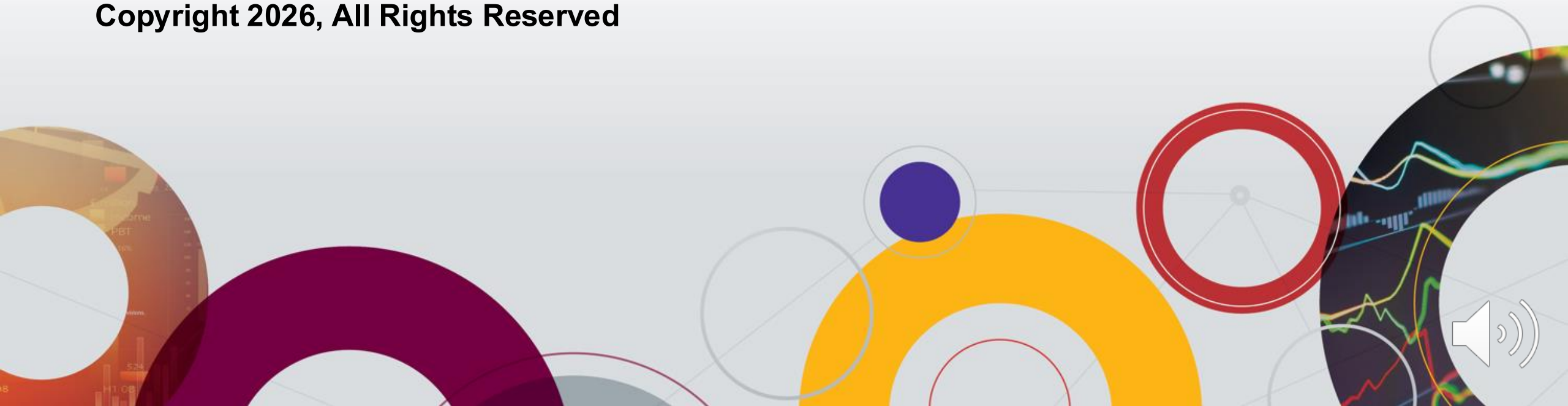


5.3 Predictive Artificial Intelligence

10x10 – Introduction to Biomedical Informatics & Artificial Intelligence
William Hersh, MD, FACMI, FAMIA, FIAHSI
Copyright 2026, All Rights Reserved



Predictive AI

Overview

Applications and results of models

Real-world usage and evaluations

Overview of predictive AI

Resurgence of AI (second era) came about with advances in

- Machine learning, especially deep learning applied to image classification (Hinton, 2006)
 - Led to [2024 Nobel Prize in Physics](#)
- Growth in quantity of data, including electronic health records (EHRs), omics (e.g., genomics), social media, mobile devices, etc.
- Improvements in computers and networks

Initial advances in imaging and other predictive (especially classification) tasks

- Great deal of research showing successful models, less research and real-world implementation of predictive AI
 - In jargon of translational research, much basic science in lab but less research and implementation in clinical settings
- Growing use of generative AI methods in predictive tasks

Biomedical and health applications of predictive AI models and results

Specific domains

- Imaging
- Other pattern recognition
- Clinical prediction from EHR and other data
- Biological processes

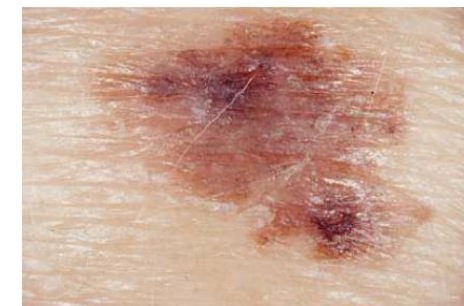
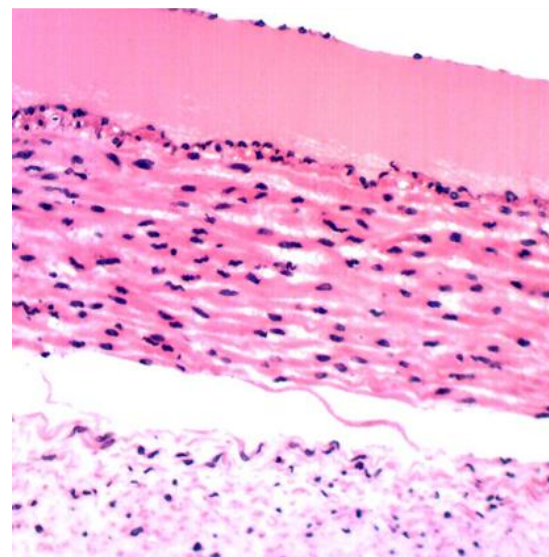
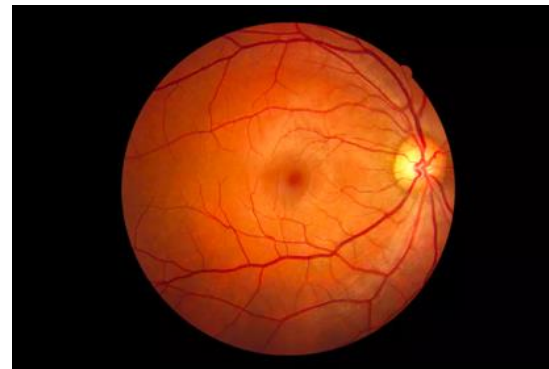
Latent discovery – seeing where humans cannot

Imaging

Early studies

- Diabetic retinopathy (DR) (Gulshan, 2016; Ting, 2017)
- Histology of cancer (Bejnordi, 2017) and metastases (Veta, 2019)
- Tuberculosis (Lakhani, 2017) and pneumonia (Rajpurkar, 2018)
- Skin cancer (Esteva, 2017; Haenssle, 2018; Tschandi, 2019)

Summarized in systematic review showing deep learning performance approaching or comparable to human experts (Liu, 2019)



Beyond basic image classification

Improved diagnosis

- Ultrasound connected to mobile app assesses gestational age and fetal malpresentation comparably to clinicians (Gomes, 2022)
- Detection of liver disease in echocardiography (Sahashi, 2025)
- Distinguishing benign from malignant vertebral fractures on CT scans comparable to radiologists (Foreman, 2024)

Improved detection of cardiovascular disease or risk

- From low-dose CT for lung cancer screening (Chao, 2021)
- From coronary artery calcium (CAC) scans (Naghavi, 2024)
- Left ventricular structural abnormalities from CXR (Bhave, 2024)

Reconstructing speech from brain imaging

- From fMRI brain recordings 82% accurate (Tang, 2023)
- From other brain imaging (Horikawa, 2025; Kunz, 2025)

Capture of video of patients in emergency department led to better prediction of hospitalization than usual triage data (Ip, 2024)

Other types of pattern recognition

Wave forms – electrocardiograms (ECGs)

- Cardiac arrhythmia detection comparable to cardiologists (Hannun, 2019)
- Interpretation better than conventional algorithm (Smith, 2019; Hughes, 2021)
- Diagnosis and risk stratification of occlusive myocardial infarction (Al-Zaiti, 2023)
- Direct-to-physician reporting of ambulatory ECGs with high negative predictive value (Johnson, 2025)
- Discern sex-related cardiovascular risk (Sau, 2025)

Electroencephalograms (EEGs)

- Clinician-level accuracy for interpretation (Tveit, 2023)

Sounds

- Detecting pathological breath sounds in children (Kevat, 2020; Zhang, 2021)
- Home monitoring of asthma exacerbations (Emeryk, 2023)
- Discerning cough from tuberculosis vs. other causes (Sharma, 2024)

Other pattern recognition (cont.)

Mobile devices – detection of

- Non-melanoma skin cancer (Stafford, 2023)
- FitBit wearable data improved early detection of complications after pediatric appendectomy (Ghomrawi, 2023)
- “Fingernail selfie” for detection and monitoring of anemia (Mannino, 2025)
- Smartwatch prediction of behavioral outbursts in children (Romanowicz, 2025)

Smell

- Graph neural network able to map chemicals to 400 different odorants (Lee, 2023)

Voice as a biomarker (Idrisoglu, 2023; Bensoussan, 2024)

- Pain detection (Borna, 2023)
- Screening for early Alzheimer’s Disease (Fristed, 2022)
- Analysis and prediction of Type 2 diabetes mellitus (Kaufman, 2023)
- Distinguish benign vs. malignant laryngeal cancer from vocal fold lesions (Jenkins, 2025)

Clinical prediction from EHR and other data

Length of stay, mortality, readmission, and diagnosis at two large medical centers (Rajkomar, 2018)

Wide variety of pediatric diagnoses from EHR data at major referral center (Liang, 2019)

Cardiovascular disease

- ML-selected variables outperformed expert-selected variables in predicting patient mortality from coronary artery disease (Steele, 2018)
- Predicted coronary artery disease and its outcomes from EHR data (Forrest, 2023)

Early prediction of other diseases

- End-stage kidney disease (Petousis, 2024)
- Predicted Alzheimer's Disease retrospectively up to 7 years before diagnosis (Tang, 2024)
- Active tuberculosis in HIV (Bartl, 2025)
- Development of diabetic retinopathy (Kim, 2025)

Clinical prediction of sepsis and patient deterioration

Sepsis

- Early RCT found success with Epic sepsis model (Shimabukuro, 2017)
- Subsequent analysis found poor performance of model (Wong, 2021)
- One use led to earlier administration of antibiotics (Tarabichi, 2022)
- More recent studies of models have found
 - Improved prediction (Bhargava, 2024)
 - Improved patient outcomes (Boussina, 2024)

Patient deterioration

- Systematic review of models found many showed some benefits, although did better in simulated than real-world settings (van der Vegt, 2023)
- Comparative evaluation of models in prospective cohort at 7 hospitals found best performance for ML-based and non-ML based models with Epic patient-deterioration model performing worse (Edelson, 2024)
- Early-warning system based on narrative concerns raised by nurses decreased in-hospital mortality and length of stay (Rossetti, 2025)

Clinical prognosis and treatment

Cancer prognosis

- Classifier trained on targeted next-generation sequencing (NGS) data for >36K tumors able to predict cancer of unknown primary with high confidence for 41% of tumors, leading to improved survival for those patients (Moon, 2023)
- Partitioned patients with cancer into distinct symptom severity clusters; patients with higher pretreatment severity were more likely to experience hospitalization and death (Xu, 2023)
- Epigenetic factor expression data for five cancer types successfully separated the patients into poor and better outcome groups (Cheng, 2023)

Treatment

- AI-guided precision parenteral nutrition for neonatal intensive care units (Phongpreecha, 2025)

Biological processes

Genomics

- Predicting clinical outcomes from cancer genomic profiles (Yousefi, 2017)
- Identifying facial phenotypes of genetic disorders (Gurovich, 2019)
- Prioritizing and classifying gene variants in sequencing data (Nicora, 2022)

Drug discovery

- Discovery of existing and new drugs effective as antibiotics (Das, 2021)
- Other discovery of new drugs (Jayatunga, 2022; Wong, 2023)

Protein folding prediction

- Revolutionary success by Google AlphaFold in predicting protein structure from DNA and amino acid sequences (Jumper, 2021; Abramson, 2024)
- Has led to catalog of likely structures of over 200M proteins based on DNA (Varadi, 2022; Bertoni, 2025)
- Led to [2024 Nobel Prize in Chemistry](#)

Latent discovery – seeing where humans cannot (Topol, 2022)

Retinal images

- Age, biological sex, and cardiovascular risk determination from retinal images (Poplin, 2018)
- Race (Coyner, 2023)

Electrocardiograms (ECGs)

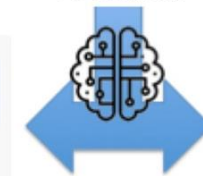
- Age and biological sex determination (Attia, 2019)
- Prediction of future arrhythmias
 - Atrial fibrillation within 31 days from normal sinus rhythm ECGs (Yuan, 2023)
 - Atrial fibrillation from at-home single-lead ECG without arrhythmias (Gadaleta, 2023)
 - Ventricular arrhythmias from single-lead ECG (Fiorina, 2025)
- Predicting in-hospital cardiac arrest from ECG heart-rate variability in ICU (Lee, 2023)
- Predicting non-cardiac diseases
 - Chronic kidney disease (Holmstrom, 2023)
 - Prodromal detection of Parkinson's Disease (Karabayir, 2023)
 - Cirrhosis of liver (Simonetto, 2025)
- Processing of ECG images
 - Left ventricular systolic dysfunction (Sangha, 2023)
 - Detecting risk of future heart failure (Dhingra, 2025)



Using AI techniques, a computer can determine from a 12-lead ECG:



Whether you are male or female with an accuracy of over 90%



Your age, if you're healthy, within 7 years ... And may determine your physiologic age if you have other comorbidities

Latent discovery (cont.)

CXR

- Recognition of race via CXR (Gichoya, 2022)
- Cardiac function and valvular heart diseases (Ueda, 2023)
- Diabetes mellitus (Pyrros, 2023)
- Correlation with chronological age in healthy cohorts and, for various chronic diseases, difference between estimated age and chronological age (Mitsuyama, 2023)
- Cardiac risk as accurately as common models, e.g., atherosclerotic cardiovascular disease (ASCVD) (Weiss, 2024)

Mammography

- Predict cardiovascular events from mammographic features (Barracough, 2025)

EHR data

- Past ICD-10 codes and lab results predicted future diagnoses in office visits (Mukherjee, 2023)

Usage and results of predictive AI in real world

Usage in US hospitals (Chang, 2025)

- 71% reported usage integrated within EHR
- Higher for medium, large, urban, system-affiliated, and non-critical access hospitals than small, rural, independent, and critical access hospitals
- Most common to predict health trajectories or risks for inpatients; fastest growing uses were to simplify billing and facilitate scheduling
- Most hospitals evaluated predictive AI models for accuracy and bias and conducted post-implementation evaluation or monitoring for most but not all uses
- Higher use for hospitals using Epic EHR

FDA approval of AI/ML devices (Sivakumar, 2025)

- 950 by end of 2024, 723 (76%) in radiology
- Of radiology devices
 - 33 (5%) underwent prospective testing
 - 56 (8%) included human-in-the-loop
 - 208 (29%) incorporated clinical testing

How well does predictive AI work in real world?

Diabetic retinopathy (DR)

- At primary care clinics, sensitivity 87.2%, specificity 90.7%, imageability rate 96.1% (Abràmoff, 2018)
- In rural India, sensitivity 88.9%, specificity 92.2%, comparable to manual grading (Gulshan, 2019)
- In Scotland, able to maintain sensitivity while improving specificity, saving time and reducing need for manual grading (Fleming, 2023)
- Implementing systems in rural Thailand and India found challenges related to equipment operation, workflows, image quality, and attention to human factors (Widner, 2023)

Radiology imaging

- Overall reading times shortened when radiologists referred to AI but some abnormalities detected could lengthen reading times (Shin, 2023)
- Large-scale study of 140 radiologists over 15 CXR tasks found heterogeneous benefits across spectrum of radiologist experience and performance, with inaccurate AI causing most amount of decline in radiologist performance across all levels (Yu, 2024)
- In mammography, increased early detection of clinically relevant cancer and reduced screen-reading workload without increasing false positives (Hernström, 2025) or negatively affecting recall rate (Eisemann, 2025)

Lessons from computer-aided detection (CAdE) of polyps in colonoscopy

Systematic reviews found increased detection of adenomas but not advanced adenomas and higher rates of unnecessary removal of nonneoplastic polyps (Hassan, 2023; Lou, 2023)

Follow-on studies

- RCT showed no significant difference in advanced colorectal neoplasia detection rate (34.8% with intervention vs. 34.6% for controls) or mean number of advanced colorectal neoplasias detected per colonoscopy (Mangas-Sanjuan, 2023)
- Observational study found reduction in adenoma detection rate from use of CAdE (Budzyń, 2025), raising concerns of “de-skilling” (Berzin, 2025)

Many challenges in implementing AI in real world, which may impact RCT results, but must build evidence base to support use (Shung, 2023)

Other examples of accurate and valid models that do not result in improved care

30-day hospital readmissions – in US, 18% of all hospital admissions

- RCT on a hospital readmissions initiative showed no improvement (Donzé, 2023)
- Accompanying editorial noted “just because we can predict something does not mean we can necessarily do something about it” (Wachter, 2023)

RCT to assess whether use of previously validated hospital-acquired venous thromboembolism (HA-VTE) prognostic model, together with pediatric hematologist review, could reduce pediatric inpatient rates of HA-VTE (Walker, 2023)

- No difference for intervention group randomized to use model
- Reluctance to use model by primary physicians – used only 26% of time
- For children in intervention arm, model mostly not used (Wilson, 2023)

Pragmatic trial of predictive algorithm and practice facilitators in outpatient practices did not reduce hospitalization rate in chronic kidney disease over 1 year compared with usual care (Vazquez, 2024)

Systematic review of use of predictive AI (Han, 2024)

Inclusion criteria

- Non-linear computational model (i.e., ANNs)
- Intervention integrated into clinical practice
- Results published as full-text article in peer-reviewed, English-language journal

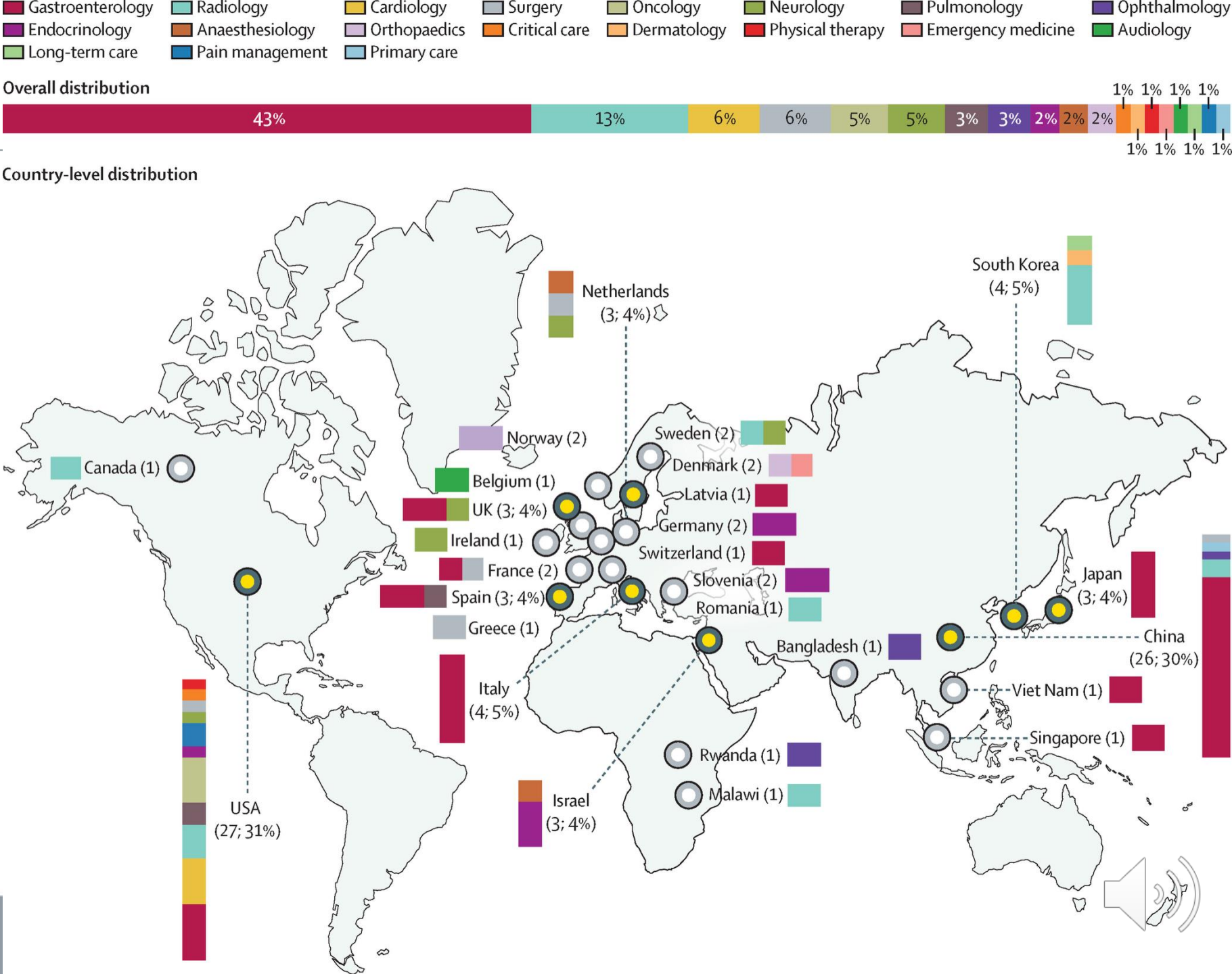
Exclusion criteria

- Linear risk scores (e.g., logistic regression)
- Secondary studies, abstracts, or interventions not integrated into clinical practice

Reported via PRISMA extension for scoping reviews

86 unique RCTs identified

Specialties and countries



Endpoints and group comparisons

	Statistically significant improvement	No statistically significant effect	Showed non-inferiority	Statistically significant deterioration	Total
Care management	15	1	2	..	18
Clinical decision making	6	1	..	..	7
Diagnostic yield or performance	34	10	1	1	46
Patient behaviour and symptoms	10	3	2	..	15
Total	65	15	5	1	86

Data are n.

Table 1: Primary endpoints and types for randomised controlled trials of artificial intelligence in clinical practice

	Statistically significant improvement	No statistically significant effect	Showed non-inferiority	Statistically significant deterioration	Total
AI vs clinician	3	1	3	1	8
AI vs routine care	16	4	..	..	20
AI-assisted clinician vs unassisted clinician	46	10	2	..	58
Total	65	15	5	1	86

Data are n. AI=artificial intelligence.

Table 2: Primary endpoint results and group comparisons for randomised controlled trials of AI in clinical practice