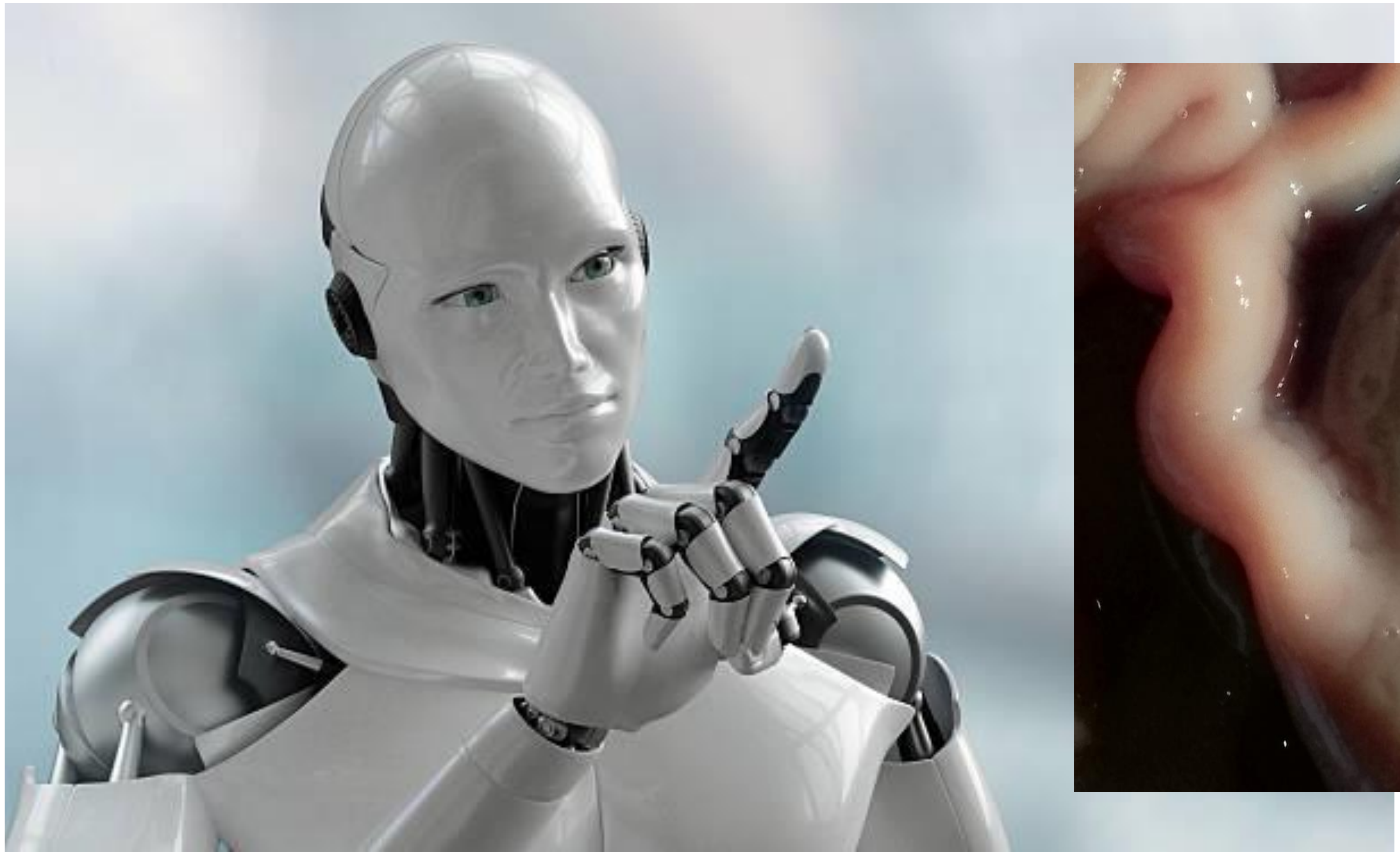
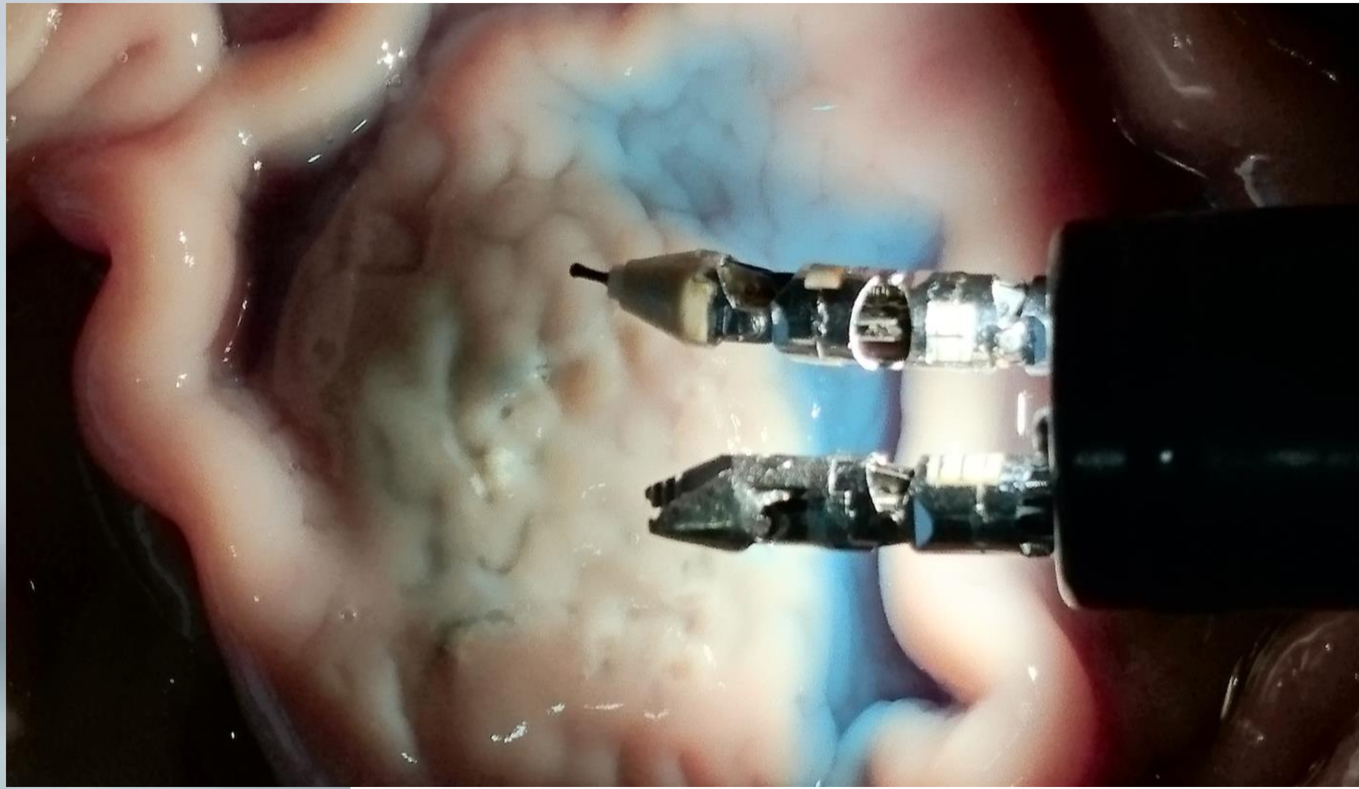
AI in Medicine
Where are we heading?

Joseph Sung MD, PhD
The Chinese University of Hong Kong

AI and Robots are coming in fast





AI will have impact at three different levels

Clinician: Rapid and accurate image interpretation

Health system: improve efficiency and allocation of resources

Patient: Promote personal health by processing own data

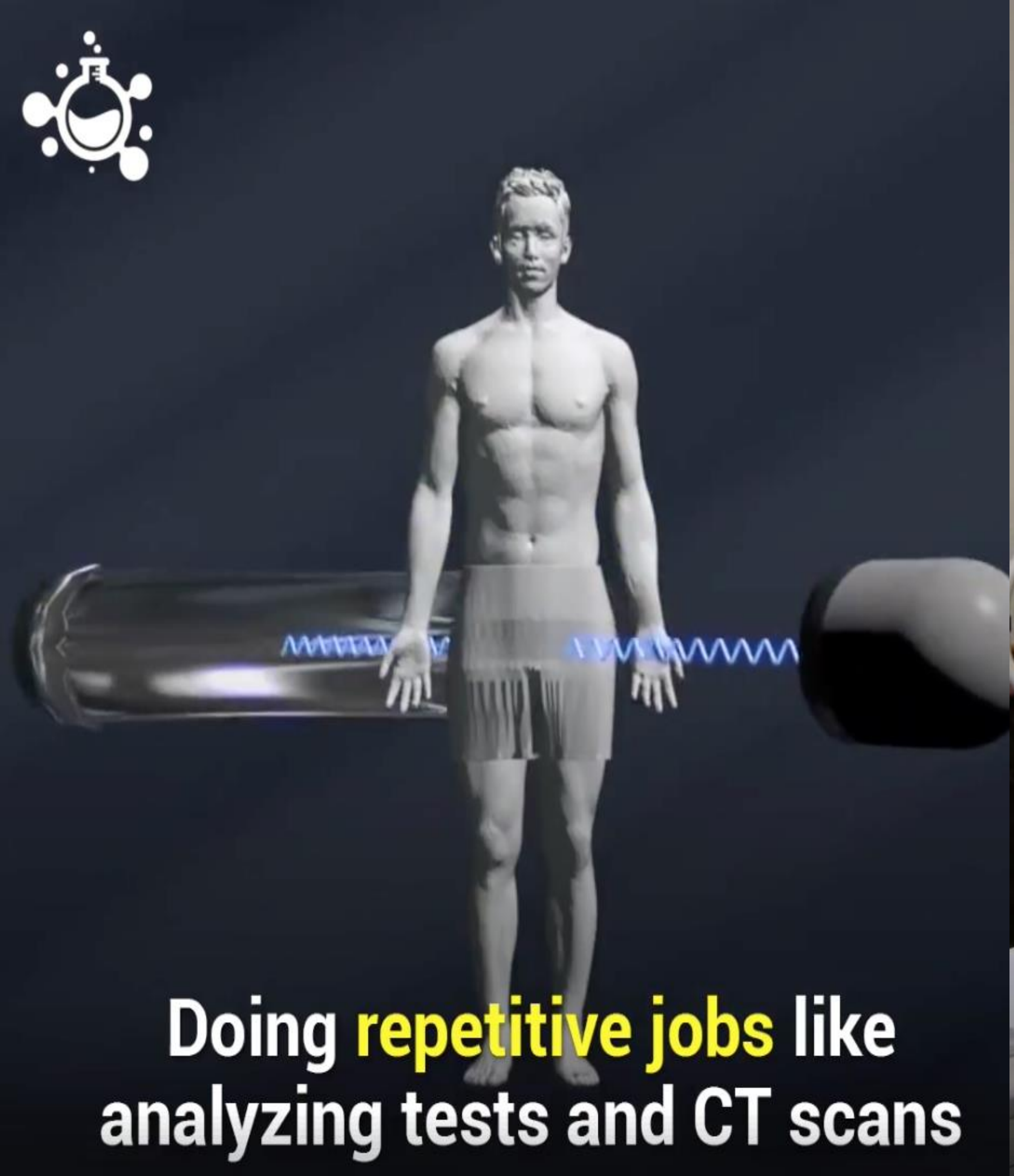
AI will have impact at three different levels

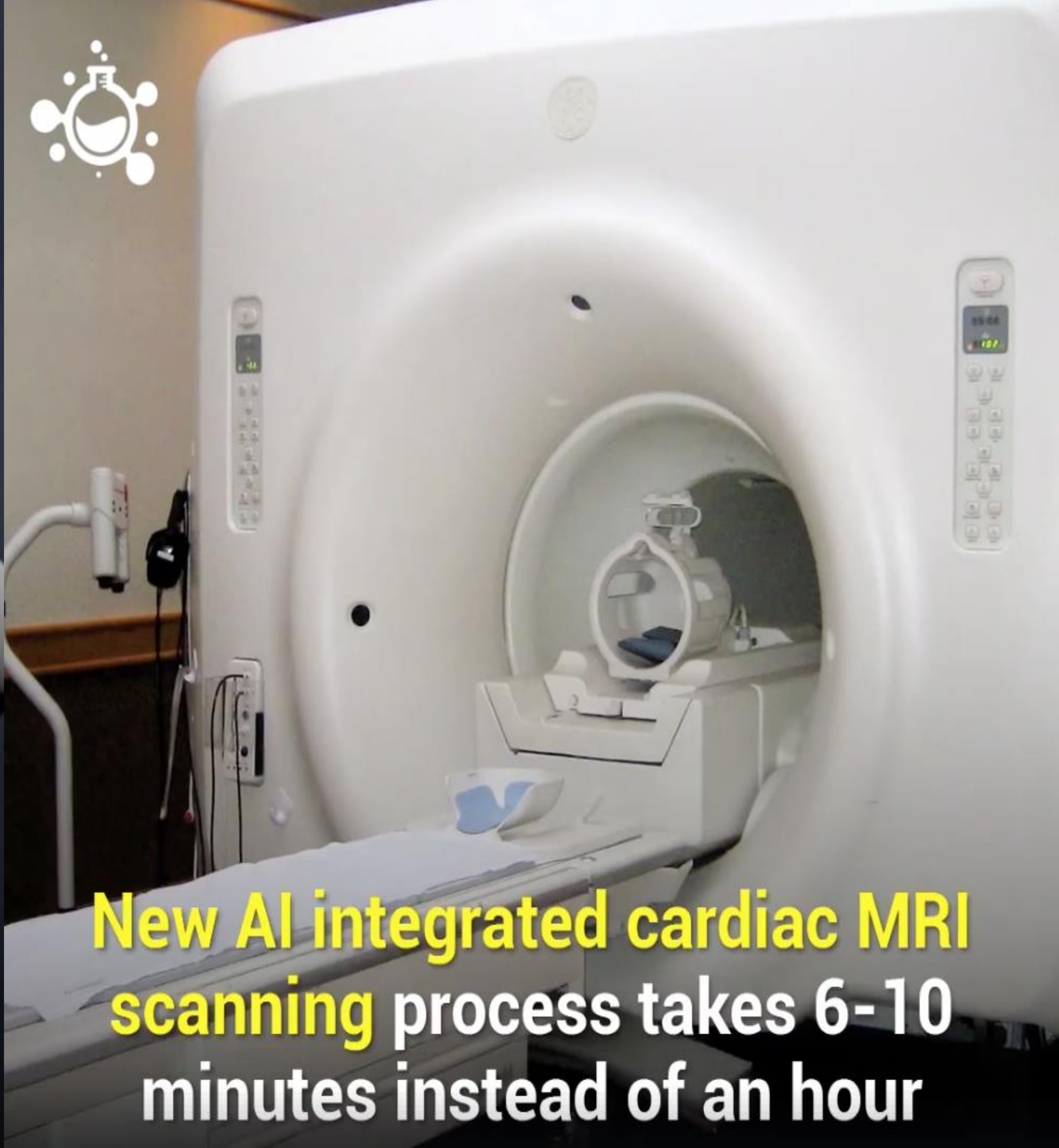
Clinician: Rapid and accurate image interpretation

- Radiology, Pathology, Neurology, Gastroenterology, Cardiology

Health system: improve efficiency and allocation of resources

Patient: Promote personal health by processing own data


Doing repetitive jobs like
analyzing tests and CT scans


New AI integrated cardiac MRI
scanning process takes 6-10
minutes instead of an hour

Radiology/Pathology/Gastroenterology

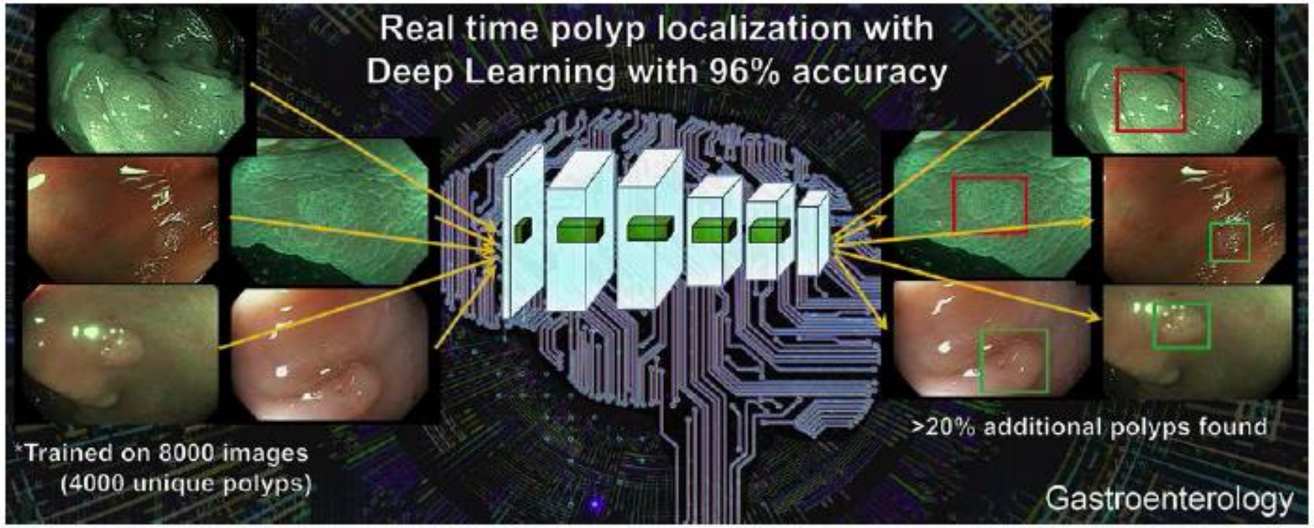
Specialty	Images	Publications
Radiology	CT head in acute neurological events CT brain for intracranial hemorrhage CXR for metastatic lung nodules Mammography for breast cancer screening	Titano et al. Nat Med 2018 Arbabshirani et al. Digital Med 2018 Nam et al. Radiology 2018 Lehman et al. Radiology 2018
Pathology	Breast cancer Lung cancer Brain tumor	Ehteshami et al. JAMA 2017 Coudray et al. Nat Med 2018 Capper et al. Nature 2018
Gastroenterology	Polyps at colonoscopy Capsule endoscopy	Mori et al. Ann Intern Med 2018 Wang et al. Gastroenterol 2019
Cardiology	Echocardiography	Medani et al. Zhang et al.

Gastroenterology 2018;155:1069–1078

Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy

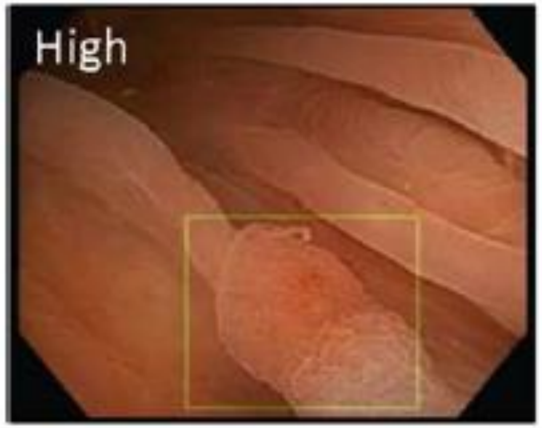
Gregor Urban,[1,2] Priyam Tripathi,[4] Talal Alkayali,[4,5] Mohit Mittal,[4] Farid Jalali,[4,5] William Karnes,[4,5] and Pierre Baldi[1,2,3]

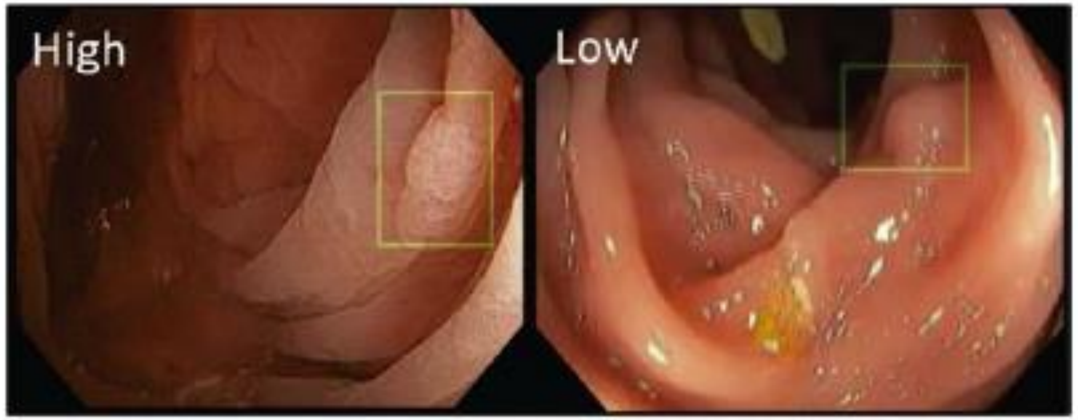
[1]*Department of Computer Science, University of California, Irvine, California;* [2]*Institute for Genomics and Bioinformatics, University of California, Irvine, California;* [3]*Center for Machine Learning and Intelligent Systems, University of California, Irvine, California;* [4]*Department of Medicine, University of California, Irvine, California; and* [5]*H.H. Chao Comprehensive Digestive Disease Center, University of California, Irvine, California*



Ideal automatic polyp detection

- High sensitivity for detection of polyp
- Decreased rate of false positives
- Low latency so that polyps can be tracked and identified in near-real time





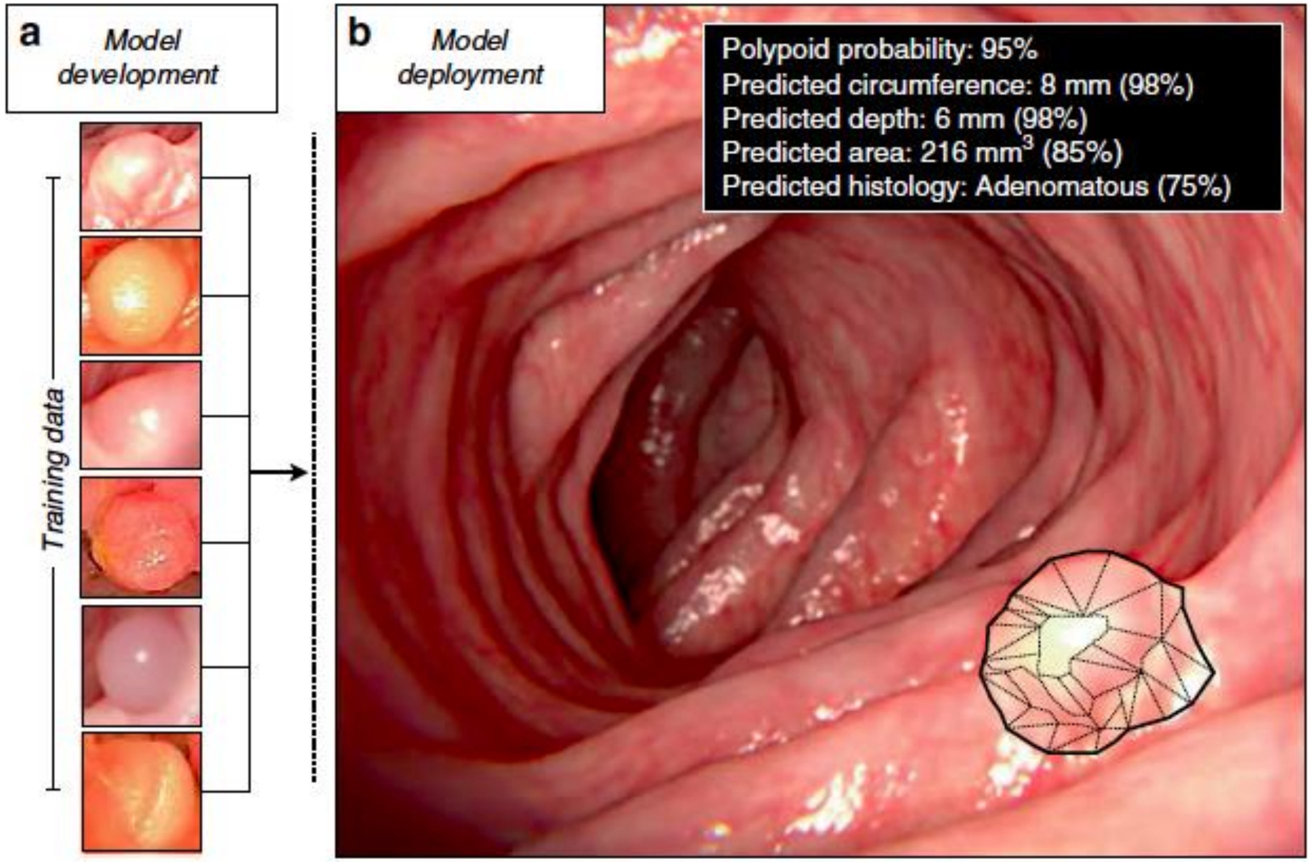
Computer-aided detection

Table 4. Unique Polyps Found and Removed During Colonoscopy, Found by Expert Review, and Found by CNN-Assisted Expert Review of 9 Videos

Polyp size (*mm*)	Original colonoscopist (polyps removed)	Expert review	CNN-assisted review
1–3	12	19	24
4–6	12	13	16
7–9	0	0	1
>10	4	4	4
Total polyps found	28	36	45

NOTE. The VGG-19–based CNN was trained on the 8,641 colonoscopy images and applied to the 9 videos without further adaptation.

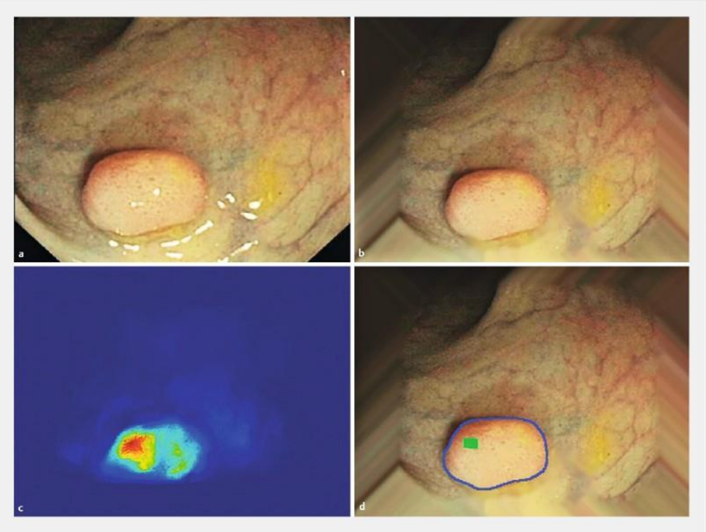
Better than human guessing

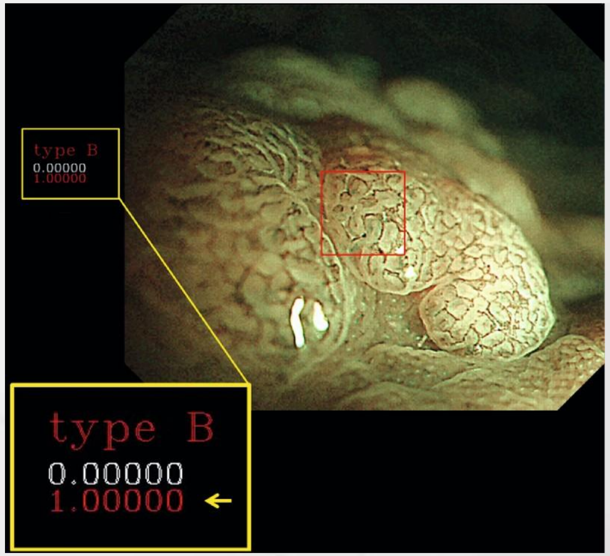


AI in Endoscopy

- **Detection (Macroscopic)**
  - Alert endoscopists using a marker / sound when AI suspect presence of pathology (polyp)

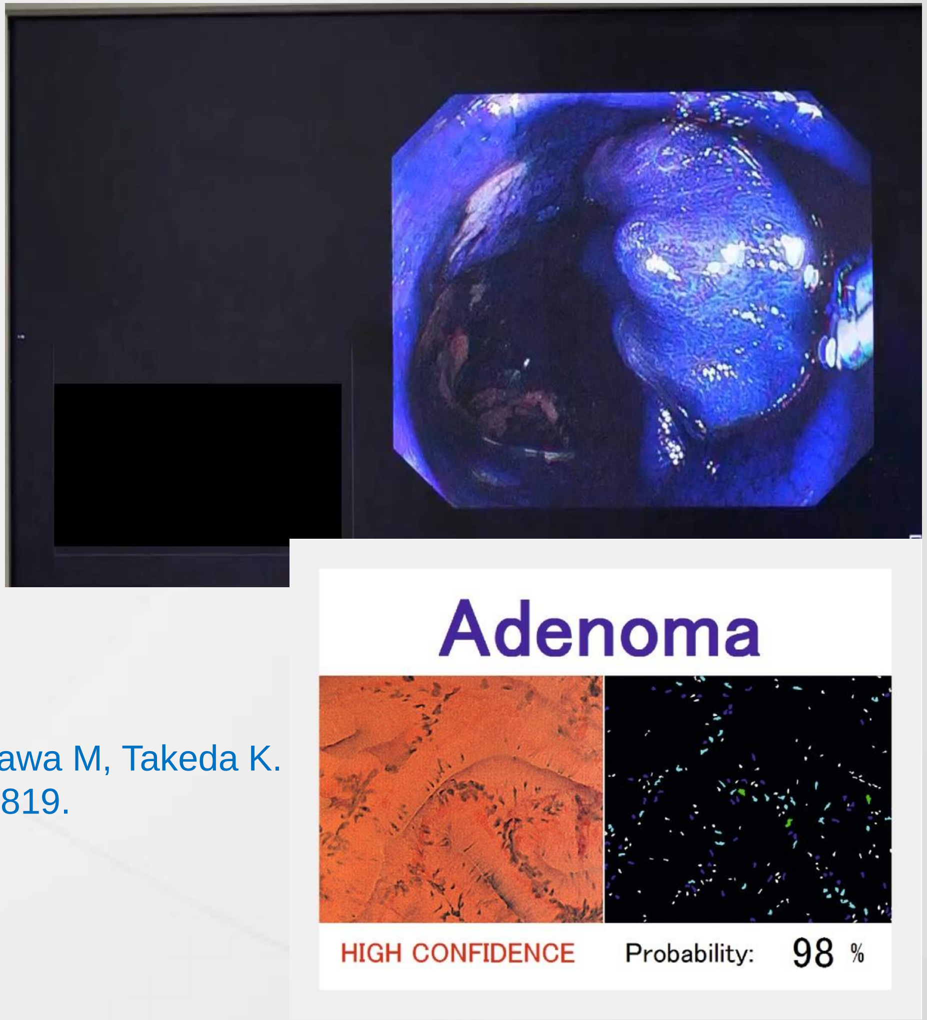
- **Characterization (Magnify)**
  - Predict pathology through classification
  - Selective resection
  - Cost saving





AI in Endoscopy

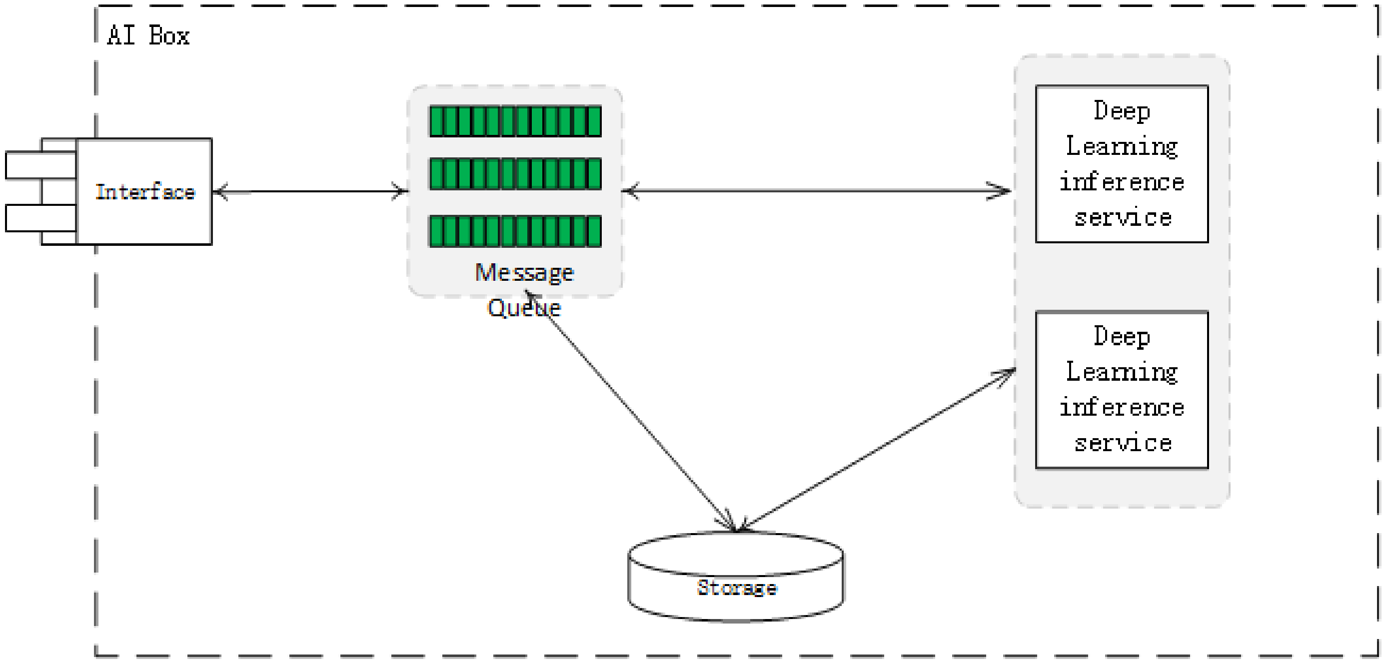
- **Pathology (Real time cytology)**
  - 500x ultramagnification
  - Diagnosis basing on findings from cellular observation
  - Images fixed size for robust image analysis
  - 90% accuracy in identification of adenoma
  - 0.2 second latency



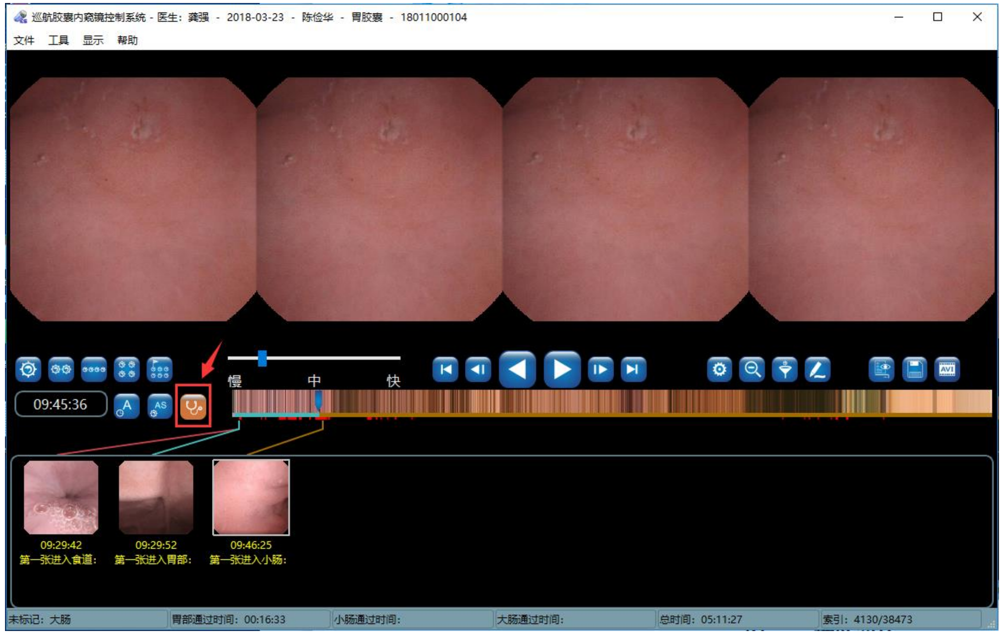
Mori Y, Kudo SE, Berzin TM, Misawa M, Takeda K. Endoscopy. 2017 Aug;49(8):813-819.

AI System for capsule endoscopy

System Design



AI-BOX Capsule Endoscopy Image Reading System Configurations



ESView Software

Data for training and validation

- **Training Data –** for developing AI software
- 158,235 CE images
- from 1970 patients.

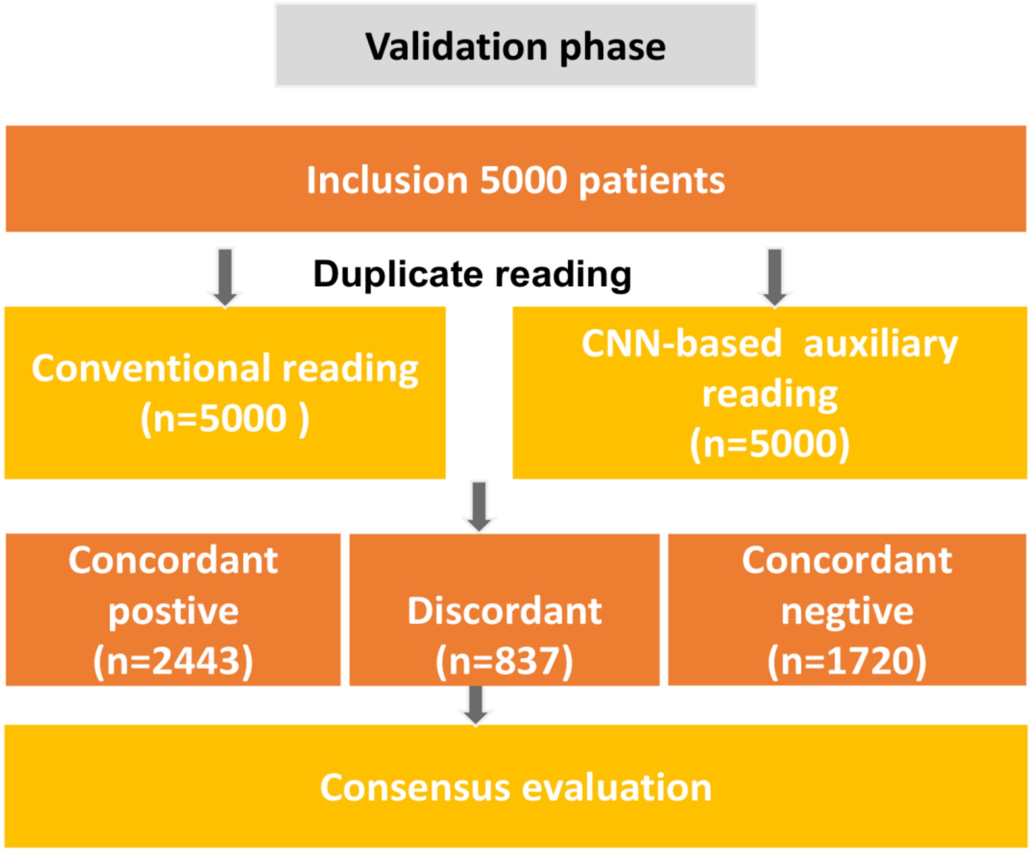
- Validation data - Retrospective Study
- 113,268,334 images
- from 5000 patients

Gastroenterologist-Level Identification of Small-Bowel Diseases and Normal Variants by Capsule Endoscopy Using a Deep-Learning Model

Gastroenterology. 2019 Jun 25. pii: S0016-5085(19)41032-9.

Validation of AI algorithm

- How to handle the discordant
  - When a diagnostic agreement was reached between conventional and CNN-based auxiliary reading, no further evaluation was carried out.
  - In case of a discordant final diagnosis and/or different lesions observed, 20 gastroenterologists sat together, and the images of the patient would be re-evaluated to confirm or reject the discordance.
  - Only were the final consensus diagnoses considered as the reference standard of diagnosis.



Validation of AI algorithm

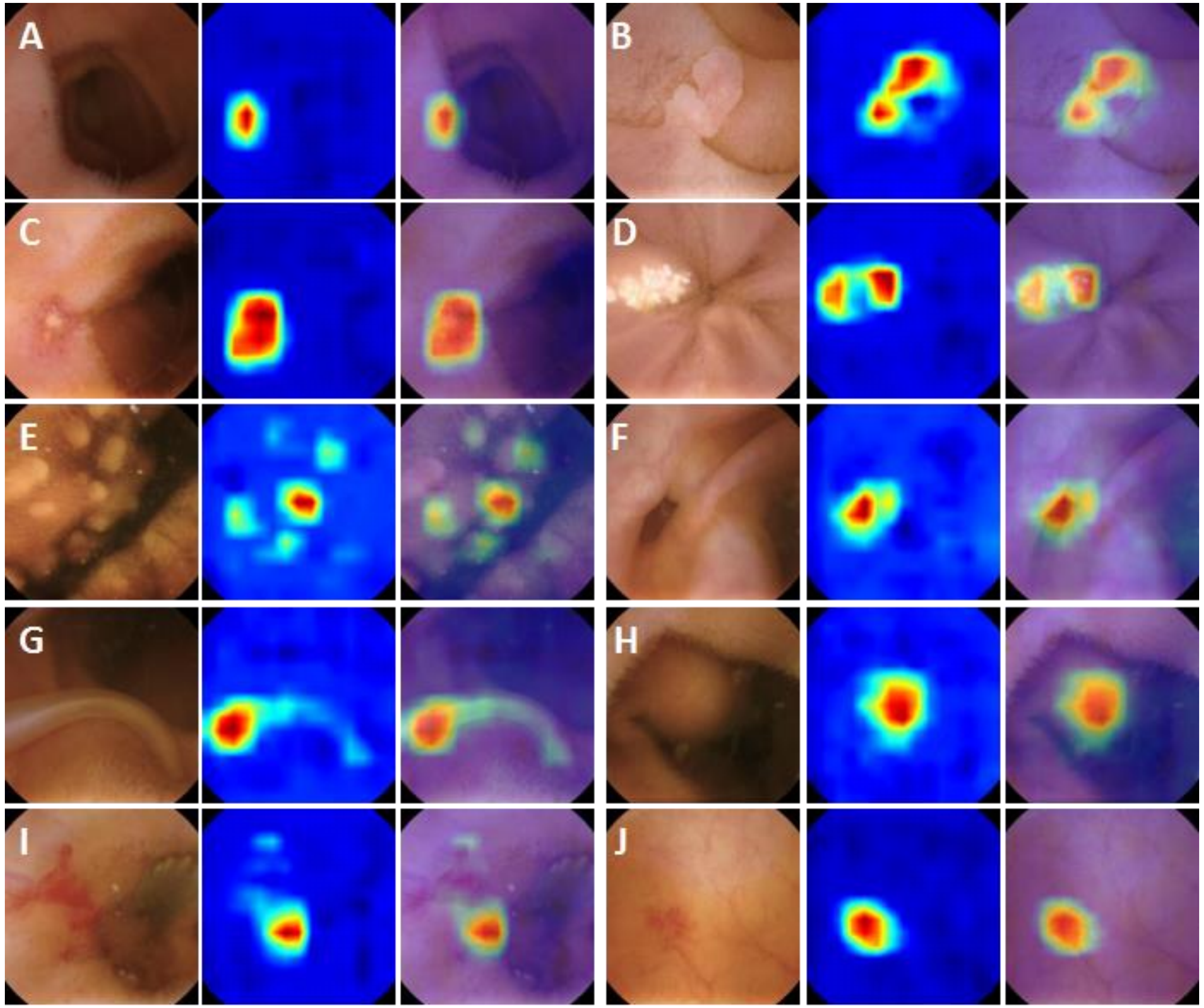


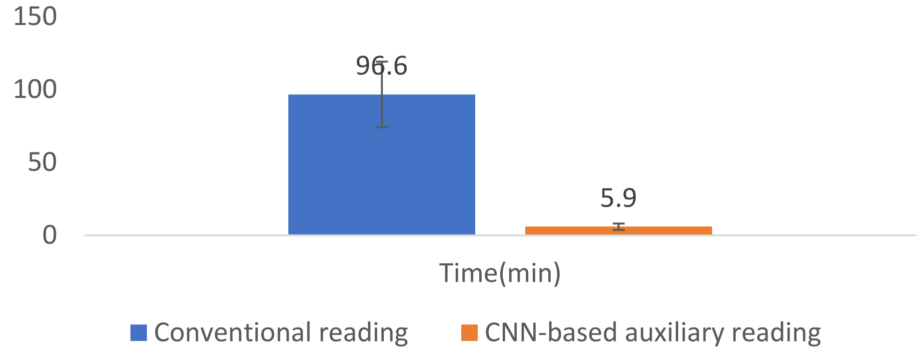
Image of Different Lesions in Small Bowel

Every lesions include three images: the left one is original image shot by NaviCam™ system; the middle one is heatmap by AI; the right one is merge of two images.

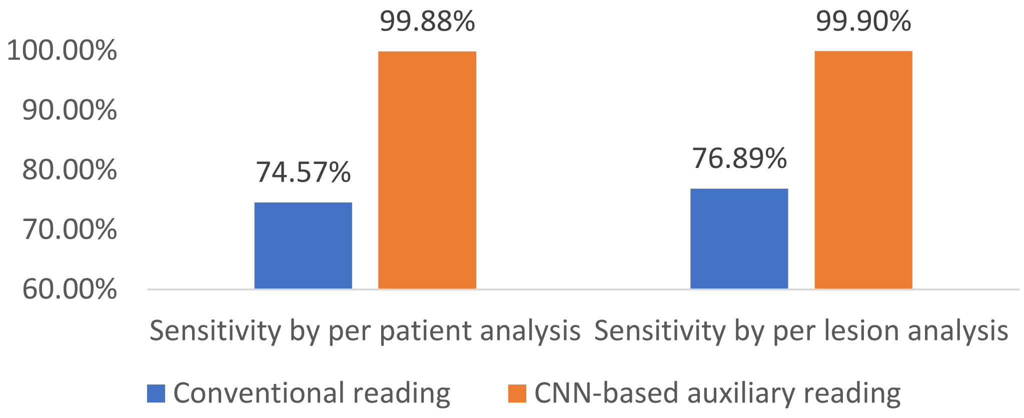
A, inflammation; B, polypus; C, ulcer; D, lymphangiectasia; E, lymphatic follicular hyperplasia; F, diverticulum; G, parasite; H, protrusive lesion; I, bleeding; J, vascular disease.

Improved diagnosis at a shorter time

Significant Reading Time Saving (from >1hr to <6 mins

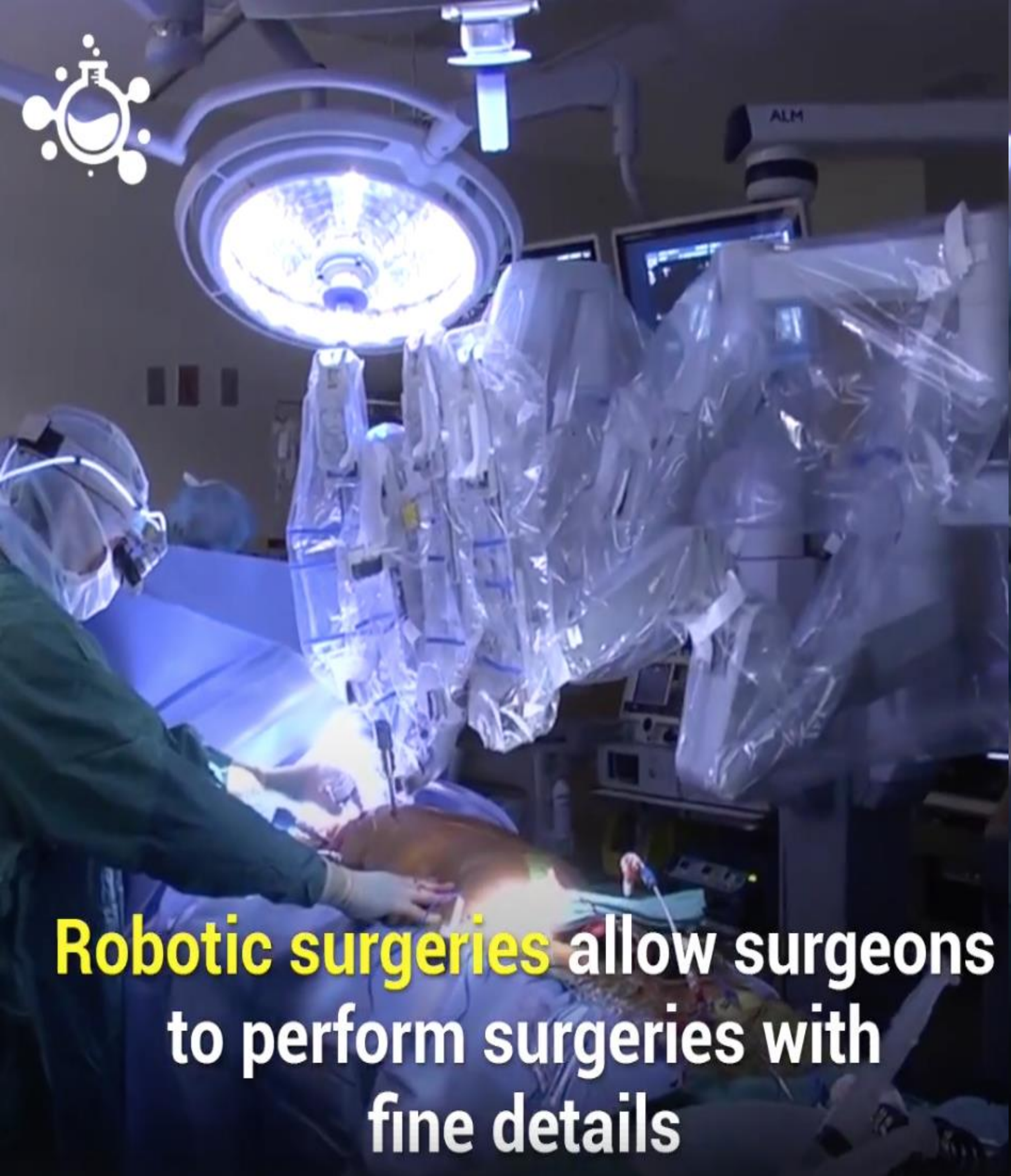


Significant Sensitivity Increase



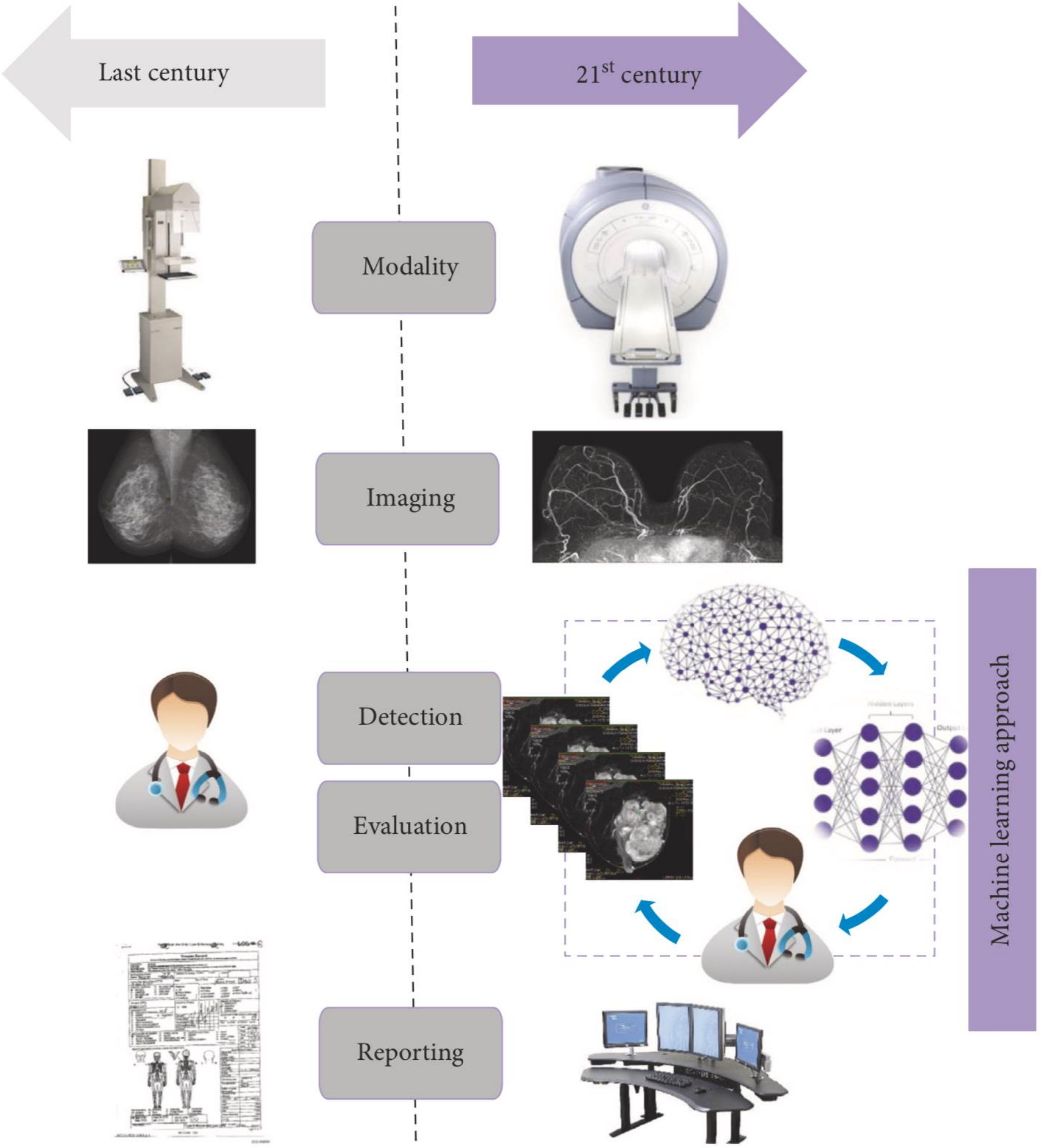
Higher Lesion Detection Rate Across Board

Type of intestine lesions	Conventional reading	CNN-based auxiliary reading
Total	3154	4144
Inflammation	1577	1663
lymphangiectasia	373	770
Ulcer	365	372
Polypus	204	285
Protrusive lesion	126	238
Lymphatic follicullar hyperplasia	120	263
Vascular disease	128	178
Bleeding	31	40
Parasite	16	16
Diverticulum	6	6
Others	213	320


ALM
Robotic surgeries allow surgeons to perform surgeries with fine details

And in tight spaces with less tremors than would be possible by human hands alone

Mammography for Breast Cancer Screening



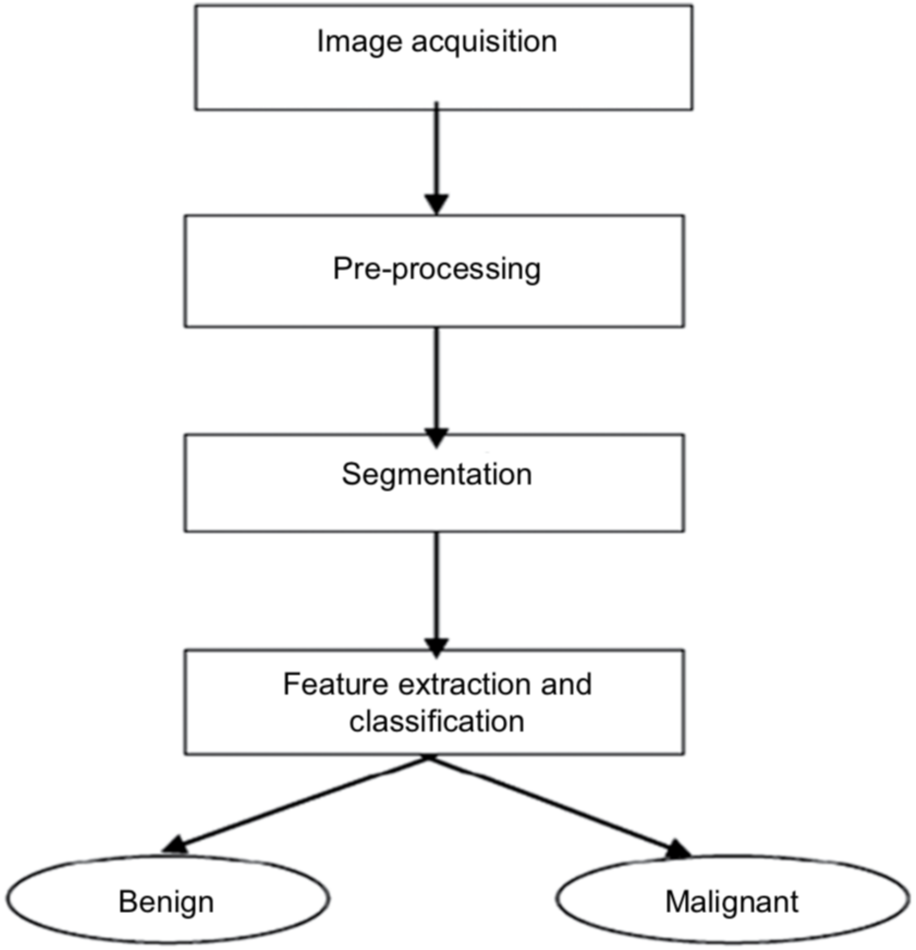
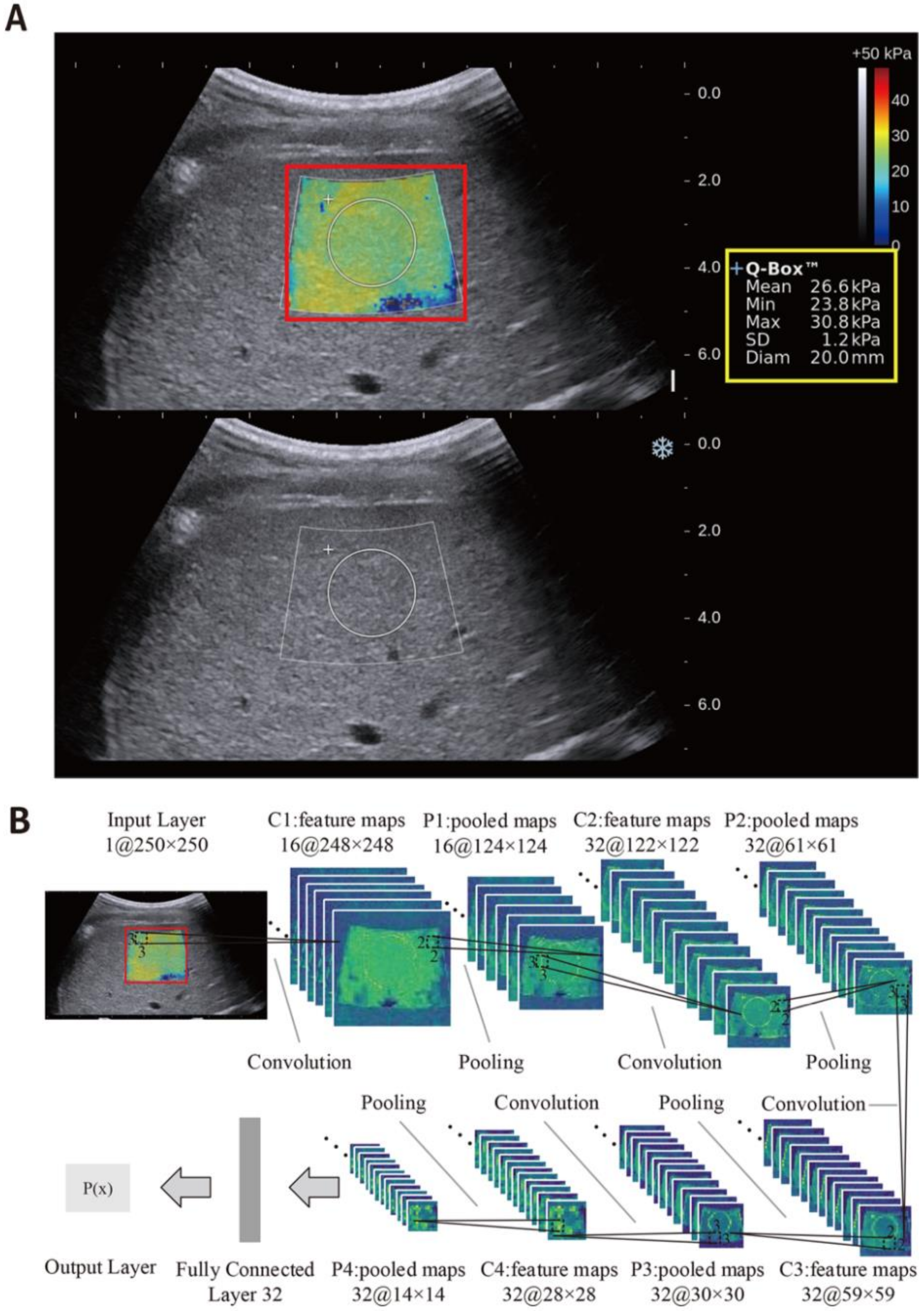


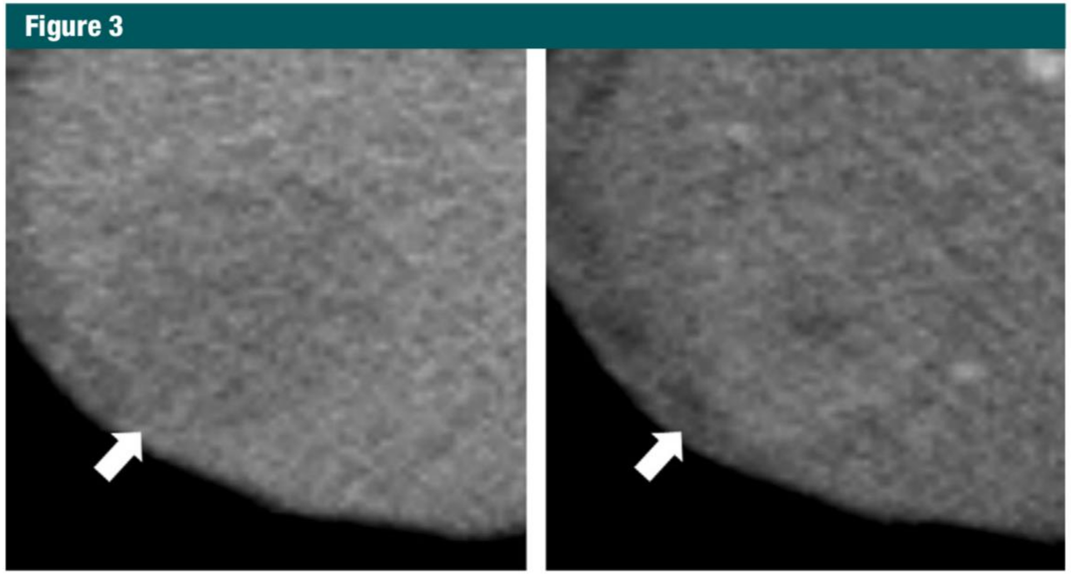
Figure 2 Stages of cancer detection by image processing.
Note: Data from Pradeep et al[16] and Lin et al.[19]

Prediction of Liver Fibrosis

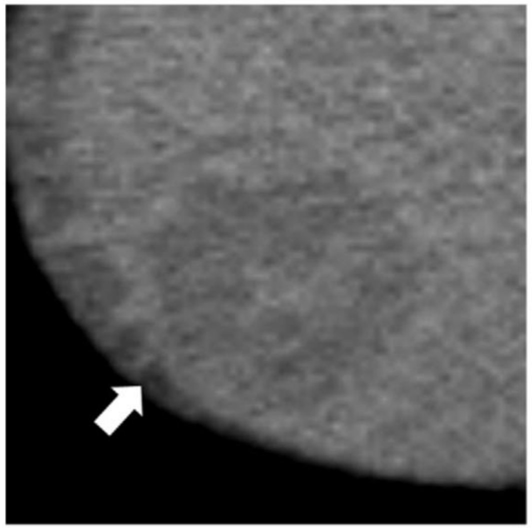


Diagnosis of liver cancer

Figure 3



a. b.



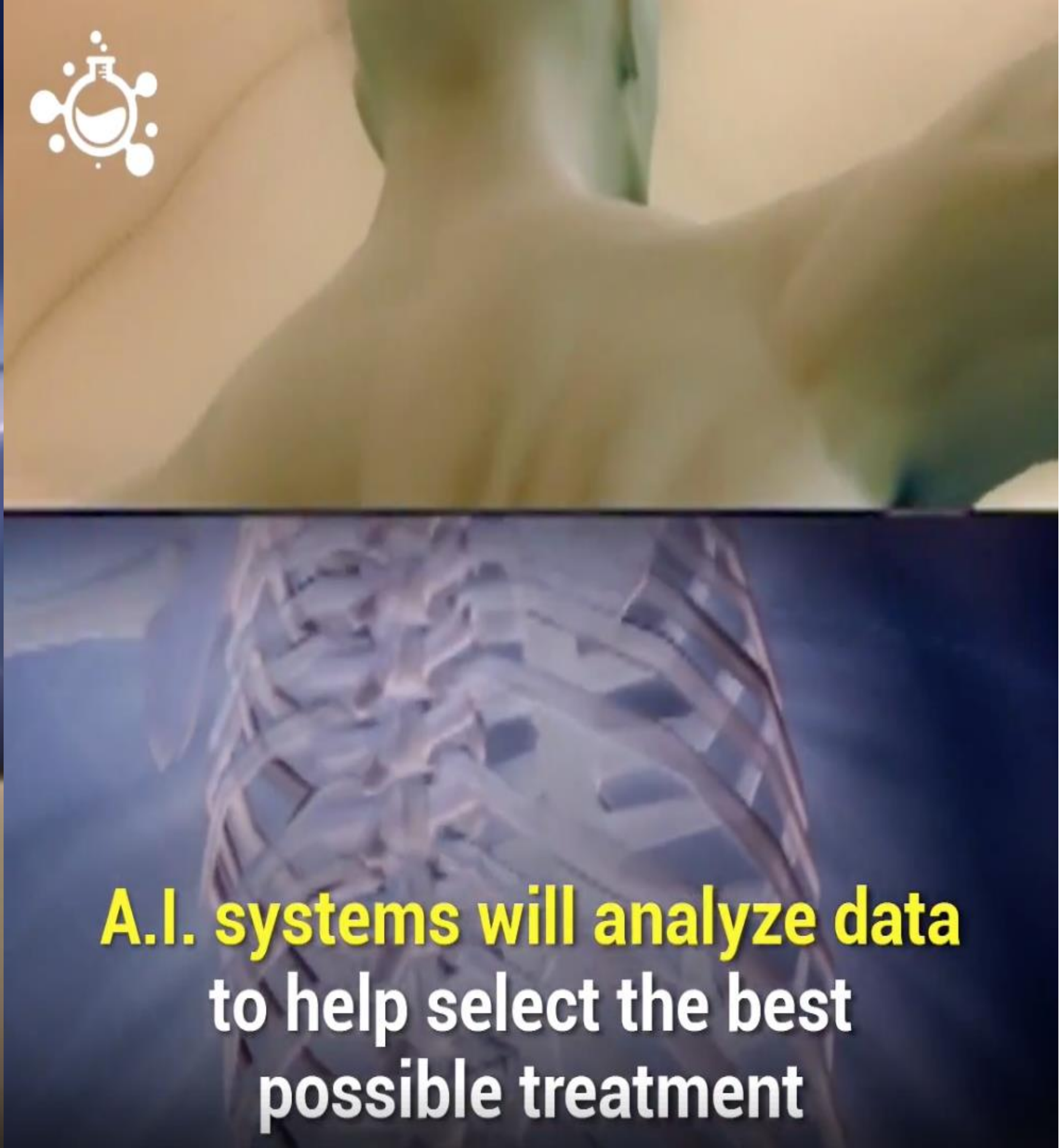
c.

Figure 3: **(a)** Unenhanced, **(b)** arterial phase, and **(c)** delayed phase CT image sets in an 82-year-old woman with a classic HCC (26-mm diameter; arrow) that was diagnosed by histopathologic analysis after surgery. The mass was correctly diagnosed as category A by using models triphasic and art/del. It was misdiagnosed as category B, C, and C by using models unenhanced, arterial, and delayed, respectively.

correctly diagnosed by using the model triphasic is shown in Figure 3.

The median area under the receiver operating characteristic curve among five sessions for discriminating categories A–B from C–E by using model


Individually customized
treatment is the next frontier
in medicine


A.I. systems will analyze data
to help select the best
possible treatment

Pathway Genomics pathway.com

Public list: digital health (3434)

MOSAIC			
820	Momentum	670	-50
	Market	950	
-20 1-month	Money	910	-10

Dashboard | Performance

Pathway Genomics, founded in
with actionable and accurate ge
and wellness. The company's n
intelligence and deep learning w
personalized health and wellne

Company (Alive / Active)

PATHWAY GENOMICS®

OUR TESTS | ORDERING & BILLING | ABOUT | LOG IN

LIQUID BIOPSY

A blood-based non-invasive test for the detection of circulating tumor DNA (ctDNA).

- CANCERINTERCEPT™ DETECT
- CANCERINTERCEPT™ MONITOR

HEREDITARY CANCER

Understand your family history of cancer, and learn about future potential cancer risks.

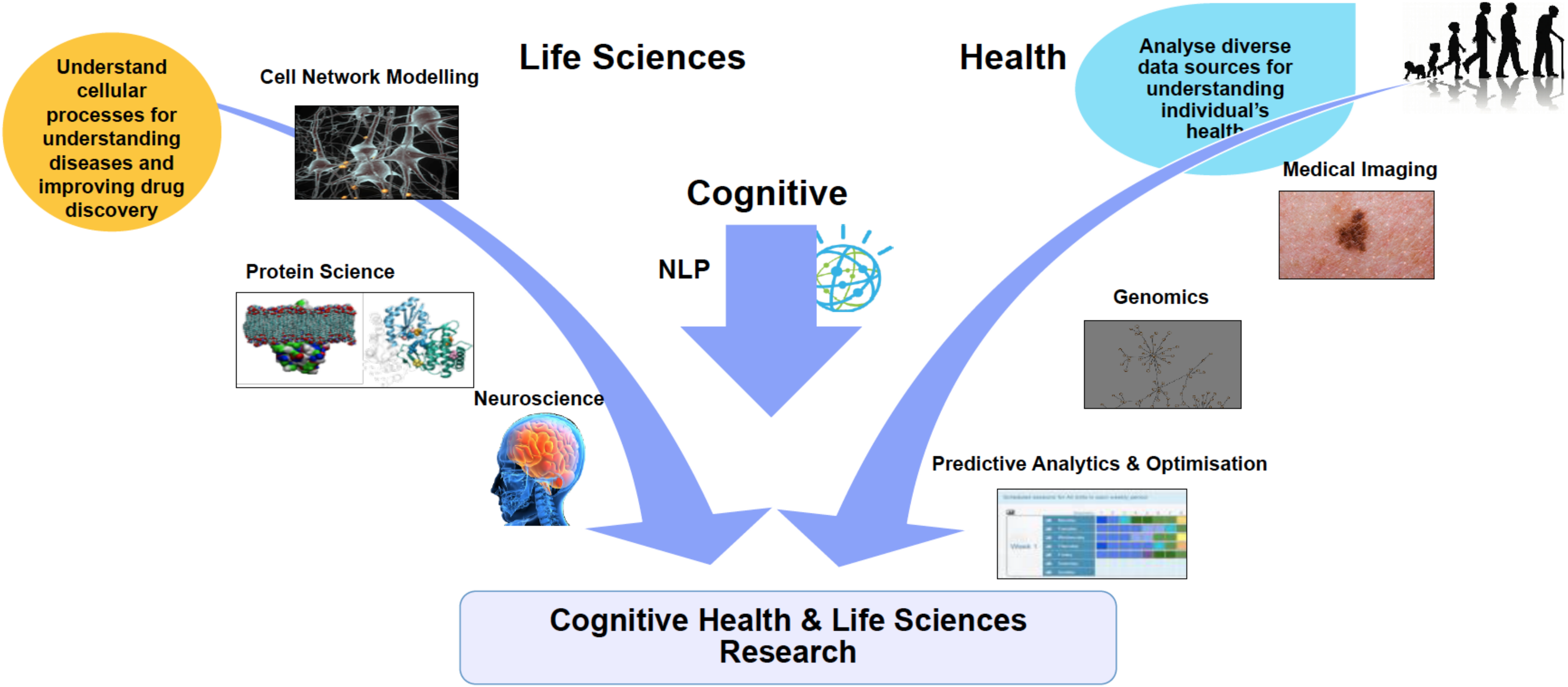
- BREASTTRUE® HIGH RISK PANEL
- BRCATRUE®
- BRCATRUE® ASHKENAZI JEWISH (3-SITE) / HISPANIC (8-SITE)
- COLOTRUE®
- LYNCH SYNDROMETRUE™

CARRIER SCREENING

Understanding potential hereditary risks for future children.

- CARRIER STATUS DNA INSIGHT®

Driving technology innovation through the convergence of health, life sciences & cognitive research



AI will have impact at three different levels

Clinician: Rapid and accurate image interpretation

Health system: improve efficiency and allocation of resources

- Prediction of key outcome- mortality and morbidity

Patient: Promote personal health by processing own data

Machine- and Deep-learning algorithm predicting clinical outcome

Prediction	N	AUC	Publication
In-hospital mortality, unplanned readmission, prolonged LOS	216,221	0.93, 0.75, 0.85	Rajkomar et al
All cause 3-12 month mortality	221,284	0.93	Avati et al
Developing diseases	704,587	range	Miotto et al
Alzheimer's Disease	273	0.91	Cleret de Langavant et al
Mortality after cancer chemotherapy	26,946	0.94	Elfiky et al
Disease onset for 133 conditions	298,000	range	Razavian et al
Suicide	5,543	0.84	Walsh et al

Is it possible to connect to eMR?

Risk estimation for hepatocellular carcinoma in chronic hepatitis B (REACH-B): development and validation of a predictive score

What is the consequence of wrong prediction?

Hwai-I Yang, Man-Fung Yuen, Henry Lik-Yuen Chan, Kwang-Hyub Han, Pei-Jer Chen, Do-Young Kim, Sang-Hoon Ahn, Chien-Jen Chen, Vincent Wai-Sun Wong, Wai-Kay Seto, for the REACH-B Working Group

Lancet Oncol 2011; 12: 568–74

Published **Online** April 15, 2011 DOI:10.1016/S1470-2045(11)70077-8

See **Comment** page 517

Genomics Research Centre, Academia Sinica, Taipei, Taiwan (H-I Yang PhD, Prof C-J Chen ScD); **Molecular and Genomic Epidemiology Research Centre, China Medical University Hospital, Taichung, Taiwan** (H-I Yang); **Department of Medicine, University of Hong Kong, Hong Kong** (Prof M-F Yuen MD, W-K Seto MBBS); **Department of Medicine and Therapeutics and Institute of Digestive Disease, The Chinese University of Hong Kong, Hong Kong** (Prof H L-Y Chan MD, V W-S Wong MD); **Department of Internal Medicine, Institute of Gastroenterology, Yonsei University College of Medicine,**

Summary

Background Therapy for chronic hepatitis B reduces the risk of progressing to hepatocellular carcinoma (HCC); however, there is no suitable and accurate means to assess risk. This study aimed to develop and validate a simple scoring system to predict HCC risk in patients with chronic hepatitis B.

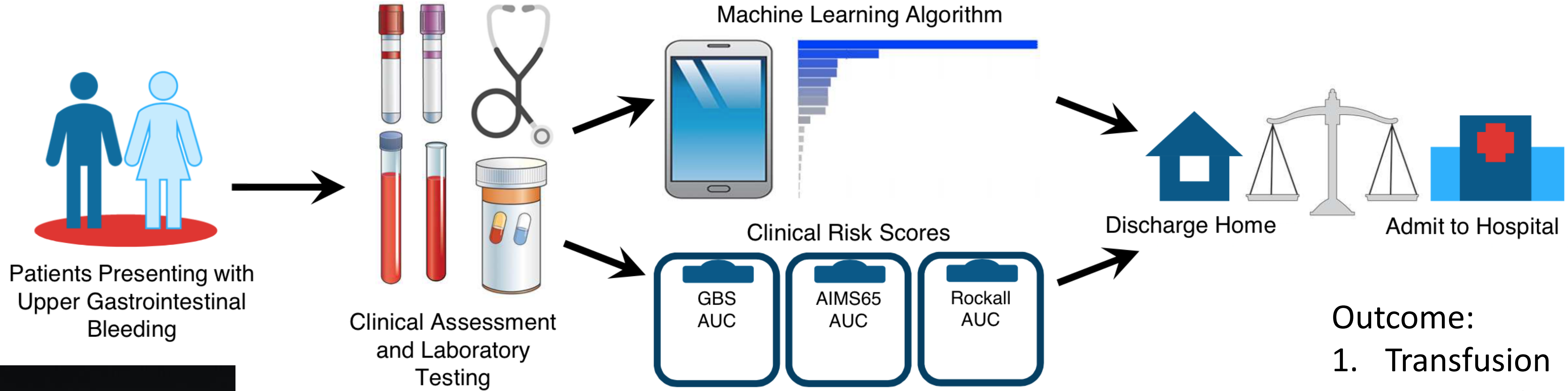
Methods The development cohort consisted of 3584 patients without cirrhosis from the community-based Taiwanese REVEAL-HBV study (of whom 131 developed HCC during follow-up), and a validation cohort of 1505 patients from three hospitals in Hong Kong and South Korea (of whom 111 developed HCC during follow-up). We used Cox multivariate proportional hazards model to predict risk of HCC at 3, 5, and 10 years. Variables included in the risk score were sex, age, serum alanine aminotransferase concentration, HBeAg status, and serum HBV DNA level. We calculated the area under receiver operating curve (AUROC) and calibration of predicted and observed HCC risk.

Findings A 17-point risk score was developed, with HCC risk ranging from 0·0% to 23·6% at 3 years, 0·0% to 47·4% at 5 years, and 0·0% to 81·6% at 10 years for patients with the lowest and highest HCC risk, respectively. AUROCs to predict risk were 0·811 (95% CI 0·790–0·831) at 3 years, 0·796 (0·775–0·816) at 5 years, and 0·769 (0·747–0·790) at 10 years in the validation cohort, and 0·902 (0·884–0·918), 0·783 (0·759–0·806), and 0·806 (0·783–0·828), respectively, after exclusion of 277 patients in the validation cohort with cirrhosis. Predicted risk was well calibrated with Kaplan-Meier observed HCC risk.

Interpretation A simple-to-use risk score that uses baseline clinical variables was developed and validated. The score accurately estimates the risk of developing HCC at 3, 5, and 10 years in patients with chronic hepatitis B. Clinicians can use this score to assess risk of HCC in patients with chronic hepatitis B and subsequently make evidence-based decisions about their clinical management.

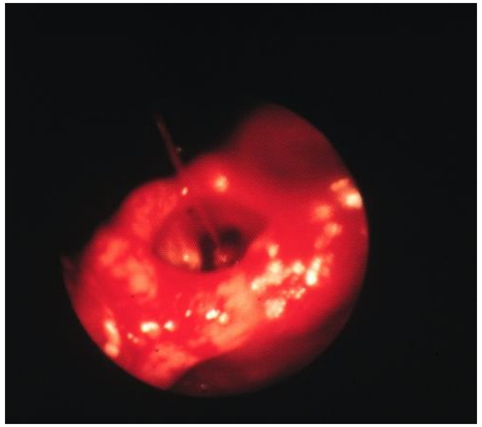
Validation of a machine learning model that outperforms clinical risk scoring system for UGIB

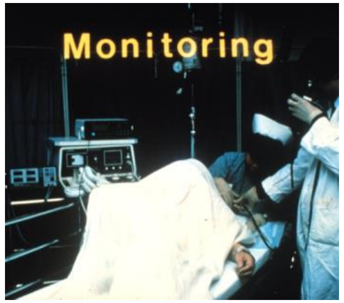
Shung et al. Gastroenterol 2019 (in press)

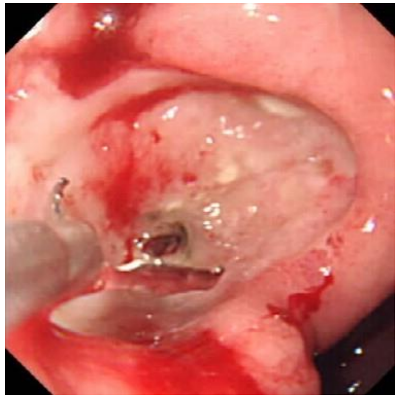


Outcome:

1. Transfusion
2. Intervention
3. Die within 30 days







Machine Learning Model vs Clinical Models

Category	Parameters
Demographic	Age Sex
Co-morbidity	ASA score Ischemic heart disease Cardiac failure, Renal failure, Liver disease Any malignancy
Medication	Aspirin, thienopyridines Anti-coagulants NSAID
Clinical feature at presentation	Pulse, systolic BP Syncope, altered mental state Melena, hematochezia
Initial Lab value	Hb, Urea, Cr, Albumin, INR

A Blatchford Score

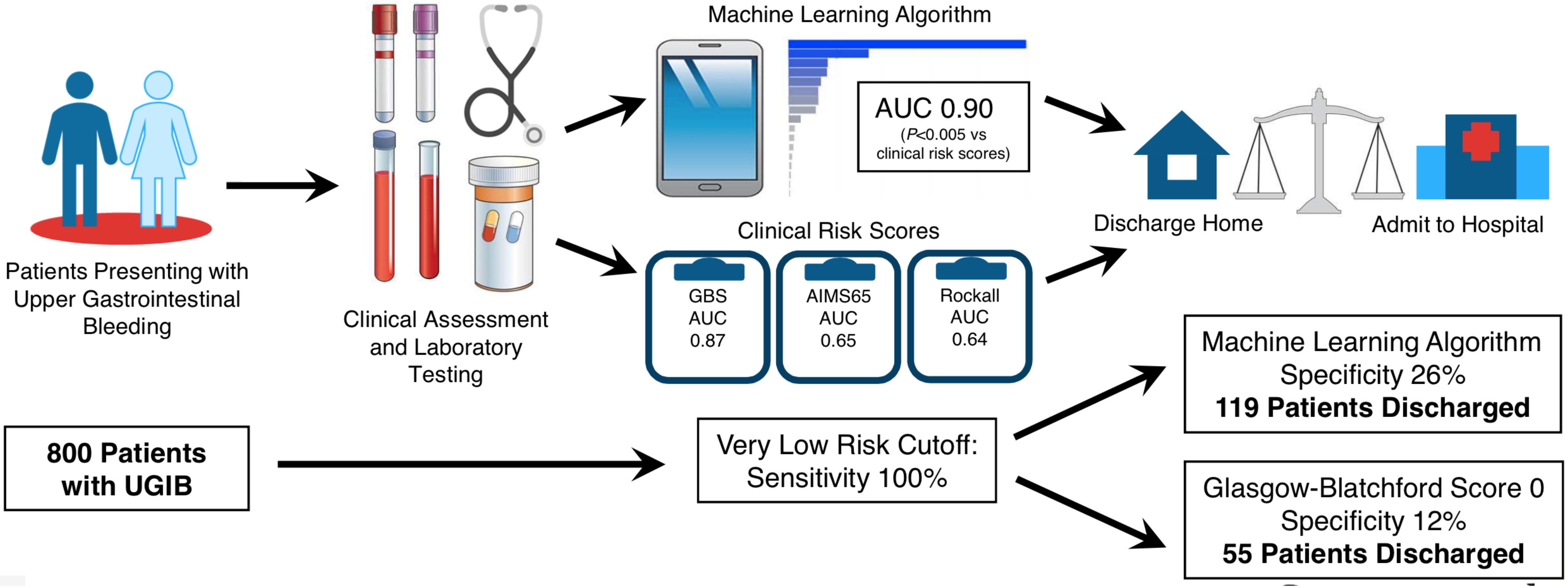
At Presentation	Points
Systolic blood pressure	
100–109 mm Hg	1
90–99 mm Hg	2
<90 mm Hg	3
Blood urea nitrogen	
6.5–7.9 mmol/liter	2

B Rockall Score

Complete Rockall Score (whole table); Clinical Rockall Score (Age, Shock, Coexisting illness)

Variable	Points
Age	
<60 yr	0
60–79 yr	1
≥80 yr	2
Shock	
Heart rate >100 beats/min	1
Systolic blood pressure <100 mm Hg	2
Coexisting illness	
Ischemic heart disease, congestive heart failure, other major illness	2
Renal failure, hepatic failure, metastatic cancer	3
Endoscopic diagnosis	
No lesion observed, Mallory–Weiss tear	0
Peptic ulcer, erosive disease, esophagitis	1
Cancer of upper GI tract	2
Endoscopic stigmata of recent hemorrhage	
Clean base ulcer, flat pigmented spot	0
Blood in upper GI tract, active bleeding, visible vessel, clot	2

Validation of a machine learning model that outperforms clinical risk scoring system for UGIB



Shung et al. Gastroenterol 2019 (in press)

Prediction of outcome impact on treatment and allocation of resources

Prediction	N	AUC	Publication
In-hospital mortality, unplanned readmission, prolonged LOS	216,221	0.93, 0.75, 0.85	Rajkomar et al
All cause 3-12 month mortality	221,284	0.93	Avati et al
Developing diseases	704,587	range	Miotto et al
Alzheimer's Disease	273	0.91	Cleret de Langavant et al
Mortality after cancer chemotherapy	26,946	0.94	Elfiky et al
Disease onset for 133 conditions	298,000	range	Razavian et al
Suicide	5,543	0.84	Walsh et al

Dig Dis Sci (2016) 61:913–919
DOI 10.1007/s10620-015-3952-x

CrossMark

ORIGINAL ARTICLE

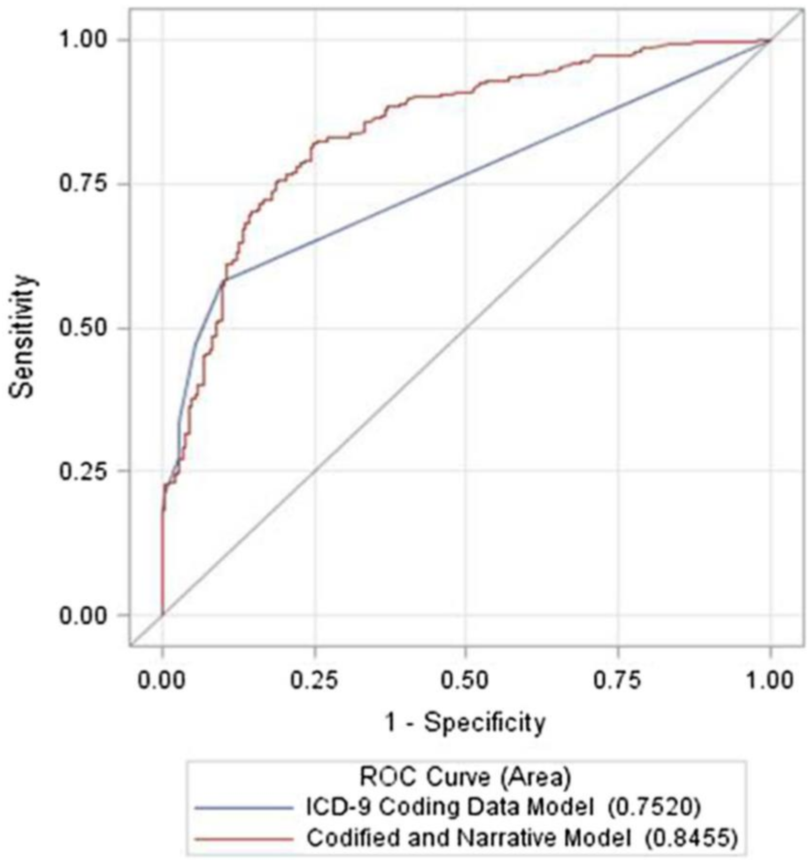
Development and Validation of an Algorithm to Identify Nonalcoholic Fatty Liver Disease in the Electronic Medical Record

Kathleen E. Corey[1,2] · Uri Kartoun[2,3] · Hui Zheng[4] · Stanley Y. Shaw[2,3]

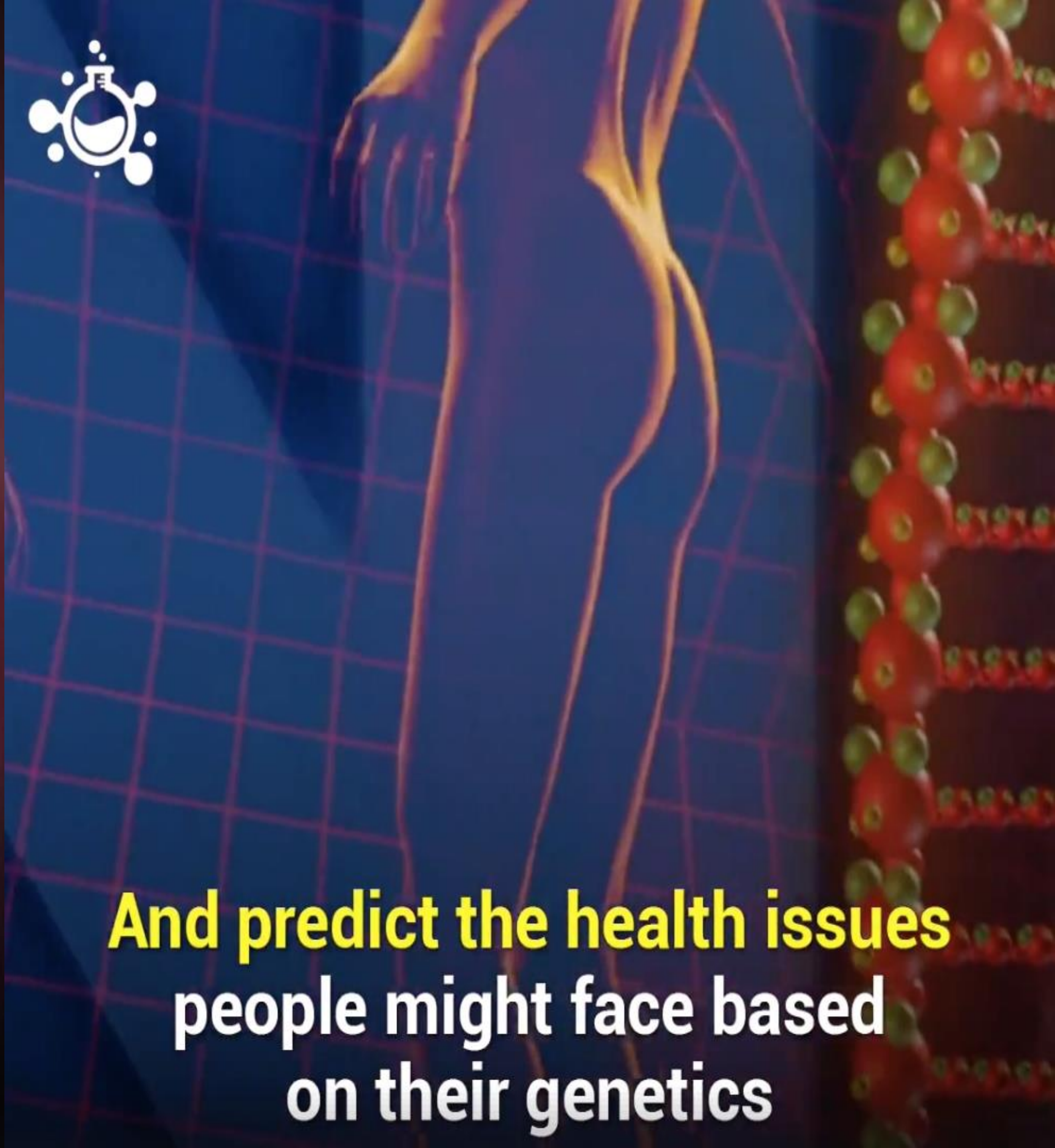
$$\begin{aligned}\text{Linear Score} = {} & -1.0742 + 0.449 * \text{fatty\_liver\_codes\_life} \\ & + 0.0792 * \text{Number\_all\_NAFLD} \\ & + 0.00765 * \text{Triglycerides}\end{aligned}$$

The probability for NAFLD was calculated using the inverse logit function:

$$\text{Probability(NAFLD)} = \exp(\text{linear score}) / (1 + \exp(\text{linear score})).$$




With the help of AI, body scans can spot cancer and vascular diseases early


And predict the health issues people might face based on their genetics

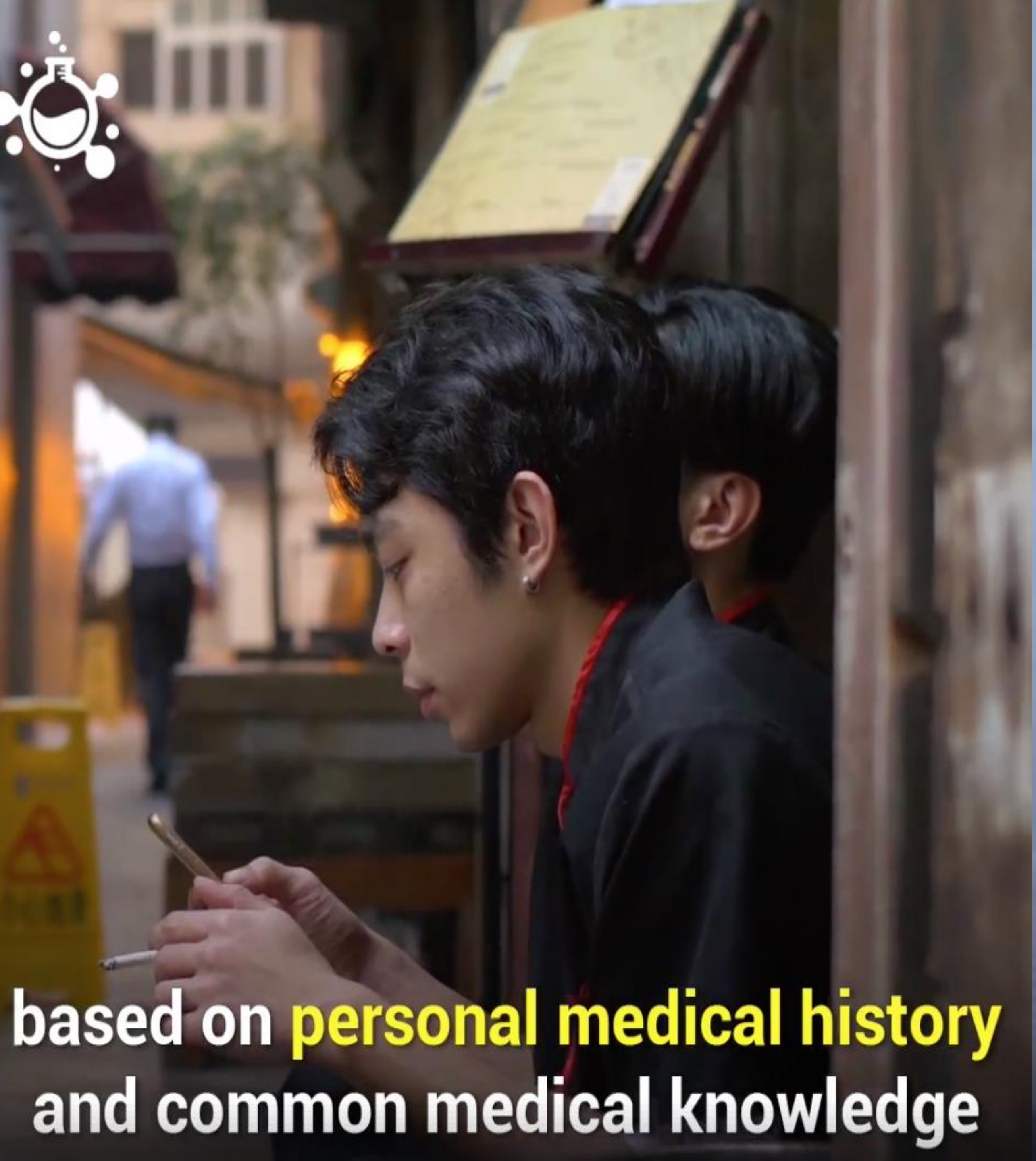
AI will have impact at three different levels

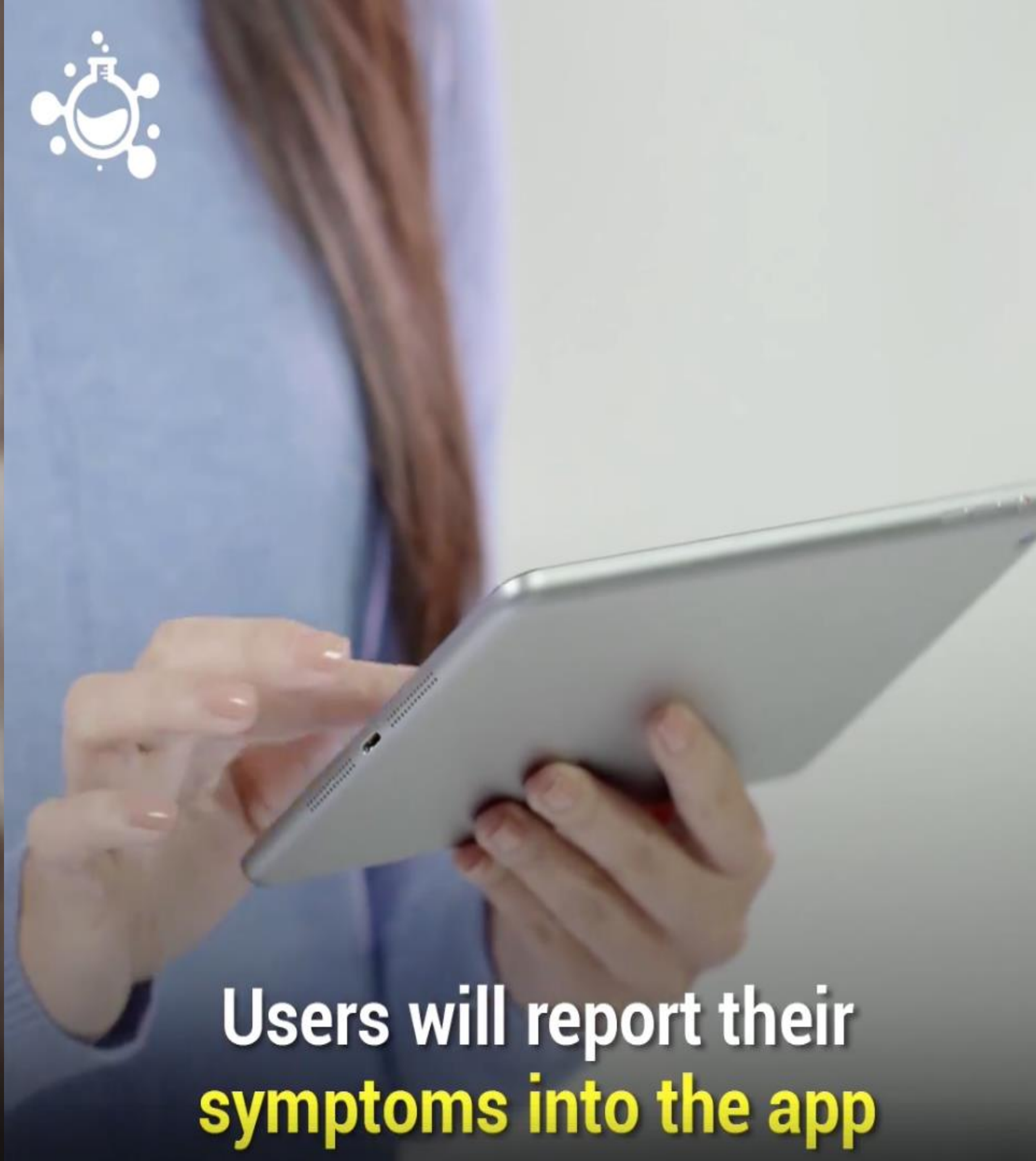
Clinician: Rapid and accurate image interpretation

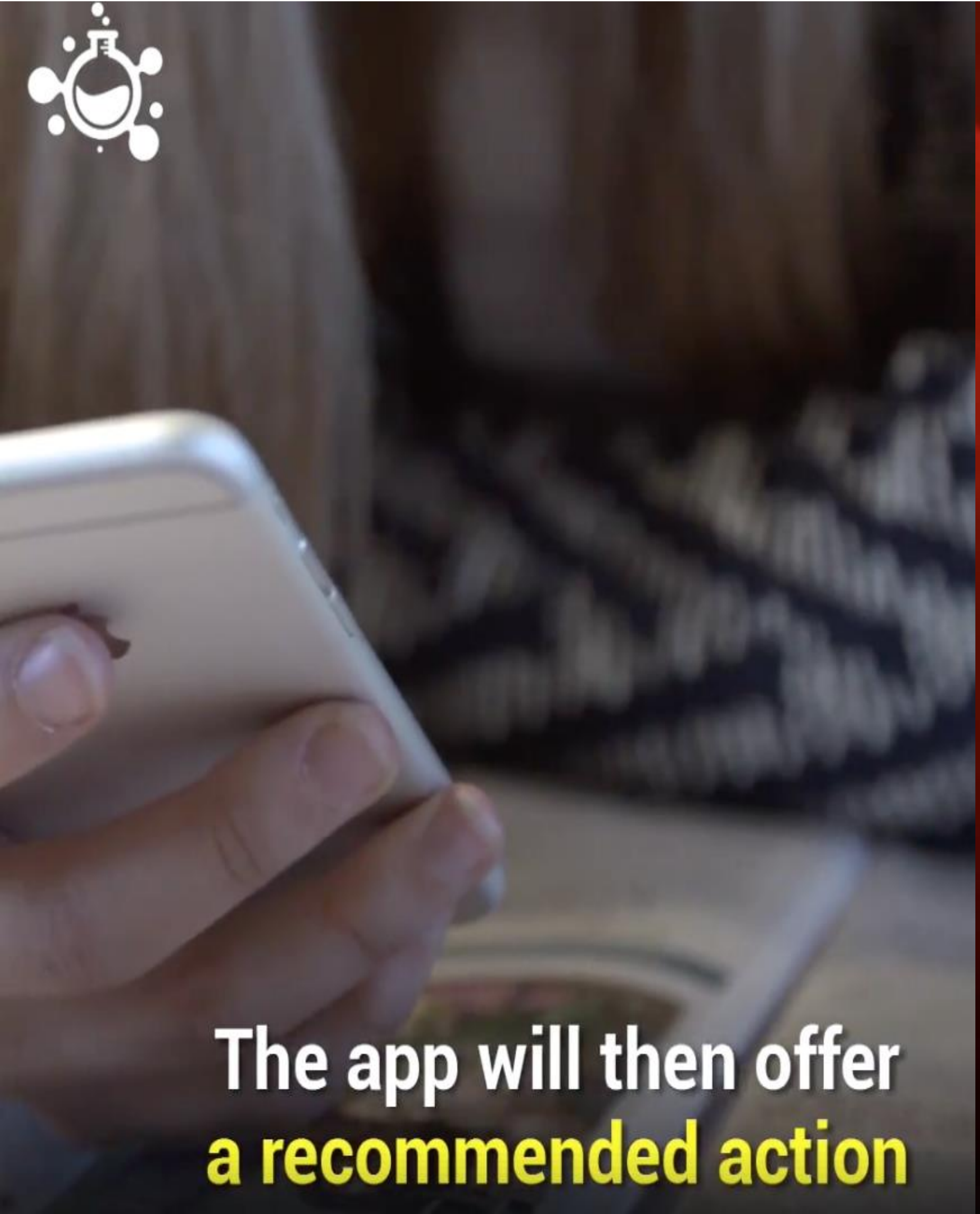
Health system: improve efficiency and allocation of resources

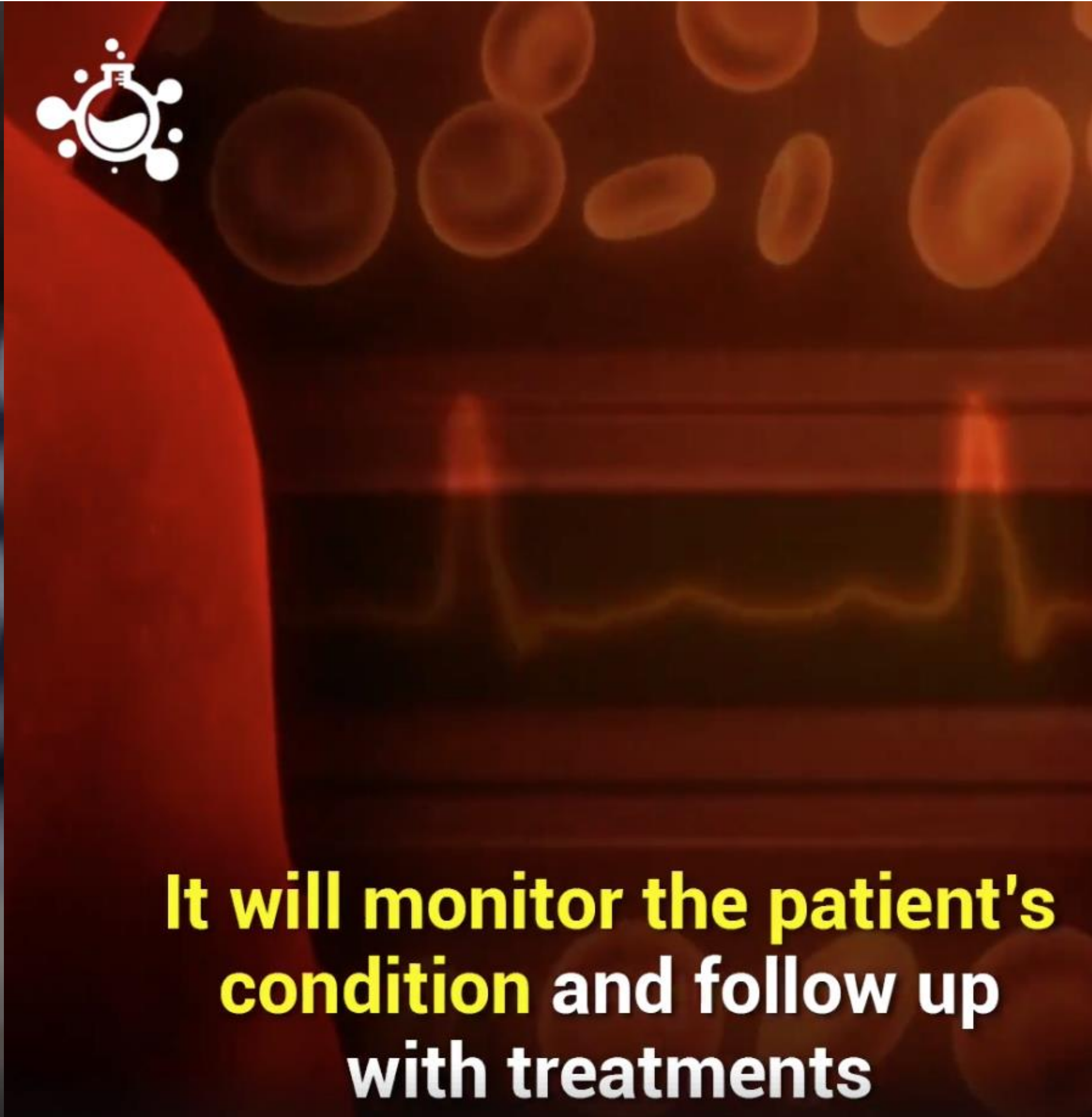
Patient: Promote personal health by processing own data

- Machine vision- clinician handwashing, ICU patient movement, Falling
- Wearable monitors- BP, HR, rhythm, PaO_2, T^oC, Respiration
- Nearest-neighbor analysis- providing the easiest access to health service


based on personal medical history
and common medical knowledge


Users will report their
symptoms into the app


The app will then offer
a recommended action


It will monitor the patient's
condition and follow up
with treatments


A smartphone's webcam is integrated with an advanced AI system


Can make sure that patients are taking their prescriptions on time


Medication Management is another area set to thrive especially among seniors

What the Apple Watch's FDA clearance actually means

The FDA-cleared features aren't supposed to be used by those under 22

By Angela Chen | @chengela | Sep 13, 2018, 12:23pm EDT

SHARE

Yesterday, Apple announced that the US Food and Drug Administration cleared two new features for the Apple Watch Series 4. One is an advanced method of monitoring the heart called an electrocardiogram (EKG), and the other is the Watch's ability to detect and notify

The Apple Watch is in Class II. For Class II and Class I, the FDA doesn't give "approval," it just gives clearance. Class I and Class II products are lower-risk products — as Speer puts it, a classic Class I example is something like a tongue depressor — and it's much easier to get clearance than approval.

Most of the time, products are cleared because they're sufficiently similar to an existing medical device that the FDA already regulated. Apple, however, has emphasized that it has received a "de novo" classification for the EKG feature. That means that, although it's still in Class II in terms of risk and hasn't gone through as much testing as an "approved" device, it's unlike anything else on the market. It is the first direct-to-consumer EKG wearable. (Last year, the FDA approved the AliveCor KardiaBand, a watch accessory that essentially does the same thing, but that wasn't direct-to-consumer.)

The Apple Watch's EKG won't be nearly as comprehensive as the one produced by a traditional electrocardiograph, which hooks up to multiple parts of the body, like the one the cardiologist used on me. The watch is a single-lead EKG device, meaning it will record one angle of the heart's electrical signals — enough to collect data about arrhythmia but not to diagnose a heart attack.

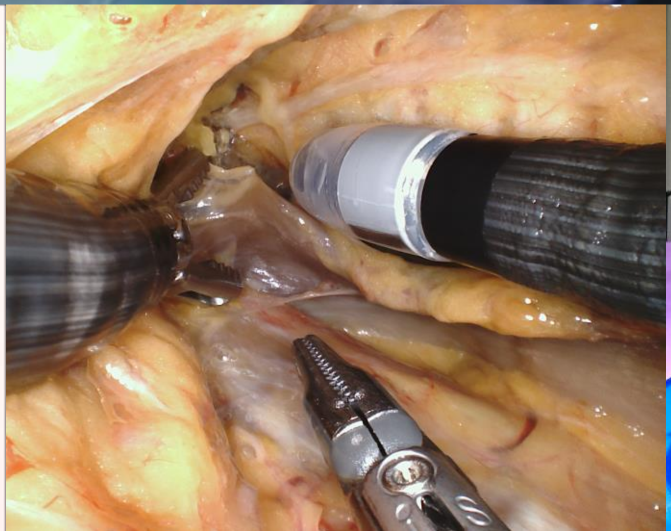

As well relay the data back
to their doctor
Electronics embedded in clothes
will be normal in the future
Smart wearable health trackers
will monitor heart rate
and activity levels

With tons of digital health data available, do we know how to interpret and make sense out of them?

Problems that may arise from wearable device?

- There is a lack of validation study to show that these digital parameters are useful
- False-positive and false-negative results
- Choosing the right person for the right device
- AI to monitor patient compliance: how reliable it is?
- Data privacy is a concern

Challenges of using AI and Robots in Medicine



Inherent biases in the data used to train AI system

HOUSE OF LORDS

Select Committee on Artificial Intelligence

Report of Session 2017–19

AI in the UK: ready, willing and able?

Select Committee on Artificial Intelligence
The Select Committee on Artificial Intelligence was appointed by the House of Lords on 29 June 2017 "to consider the economic, ethical and social implications of advances in artificial intelligence."

Membership
The Members of the Select Committee on Artificial Intelligence were:

Baroness Bakewell
Lord Clement-Jones (Chairman)
Lord Giddens
Baroness Grender
Lord Hollick
Lord Holmes of Richmond
Lord Levene of Portsoken
The Lord Bishop of Oxford
Lord Puttnam
Viscount Ridley
Baroness Rock
Lord St John of Bletso
Lord Swinfen

Declaration of interests
See Appendix 1.

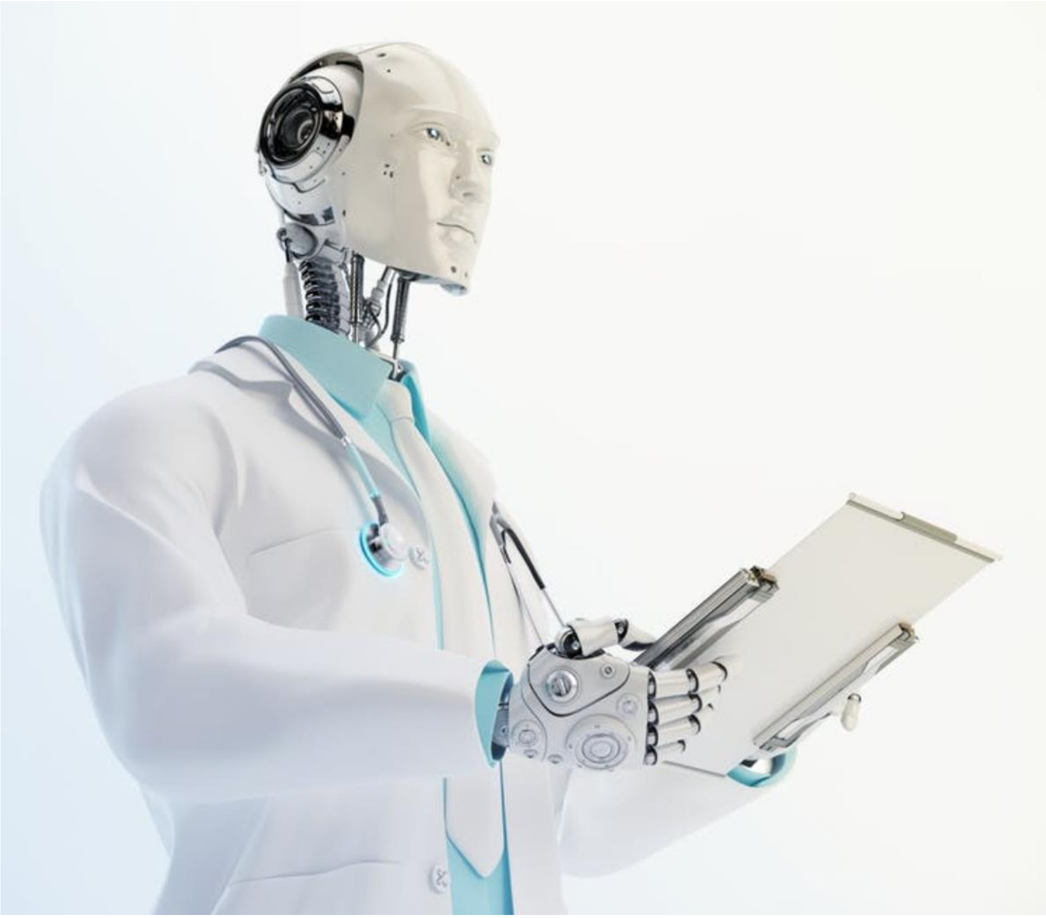
A full list of Members' interests can be found in the Register of Lords' Interests:
http://www.parliament.uk/mps-lords-and-offices/standards-and-interests/register-of-lords-interests

Publications
All publications of the Committee are available at:
http://www.parliament.uk/ai-committee

Ensuring data security for privacy

We are generating more data about people's health and biology than ever before. Combined with advances in IT and data science, this offers significant opportunities to generate new knowledge and improve medical practice; but it also raises concerns about individuals' privacy.

Securing TRUST in the use of AI technology



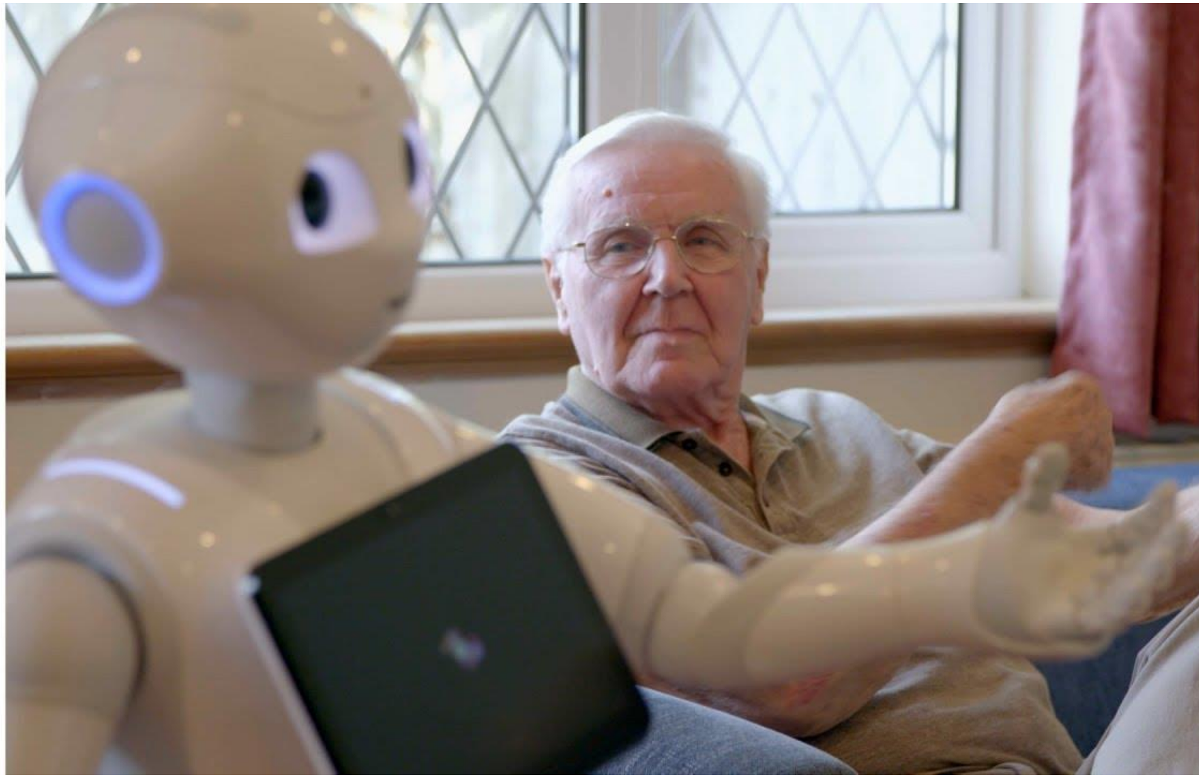
The problem with Watson for Oncology was that doctors simply didn't trust it. Human trust is often based on our understanding of how other people think and having experience of their reliability. This helps create a psychological feeling of safety.

AI, on the other hand, is still fairly new and unfamiliar to most people. It makes decisions using a complex system of analysis to identify potentially hidden patterns and weak signals from large amounts of data.

AI's decision-making process is usually too difficult for people to understand. And interacting with something we don't understand can cause anxiety and make us feel like we're losing control.

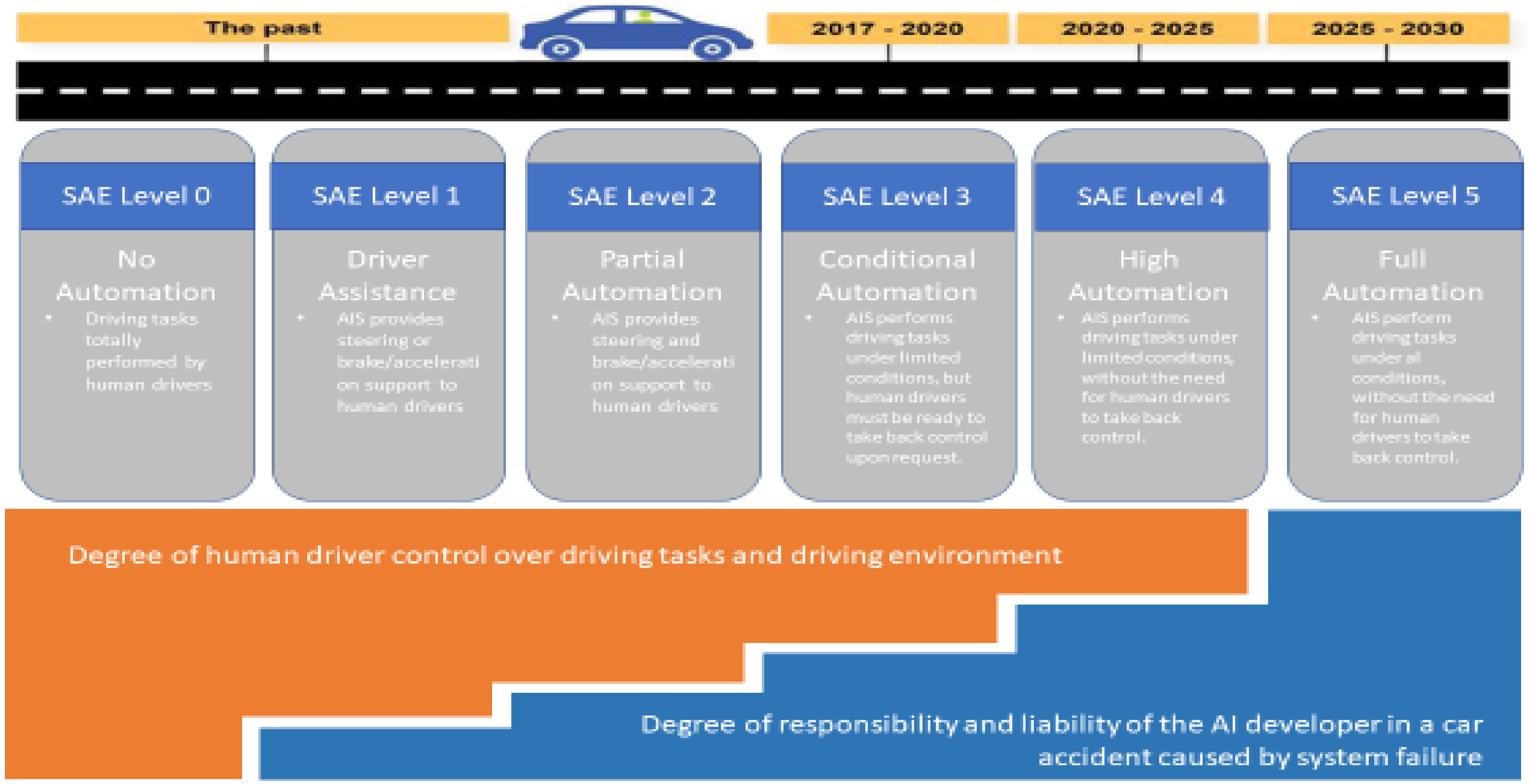
Guardian Nov 3 2017. Why we can't leave AI in the hands of big tech

People's sense of dignity and social isolation in care environment



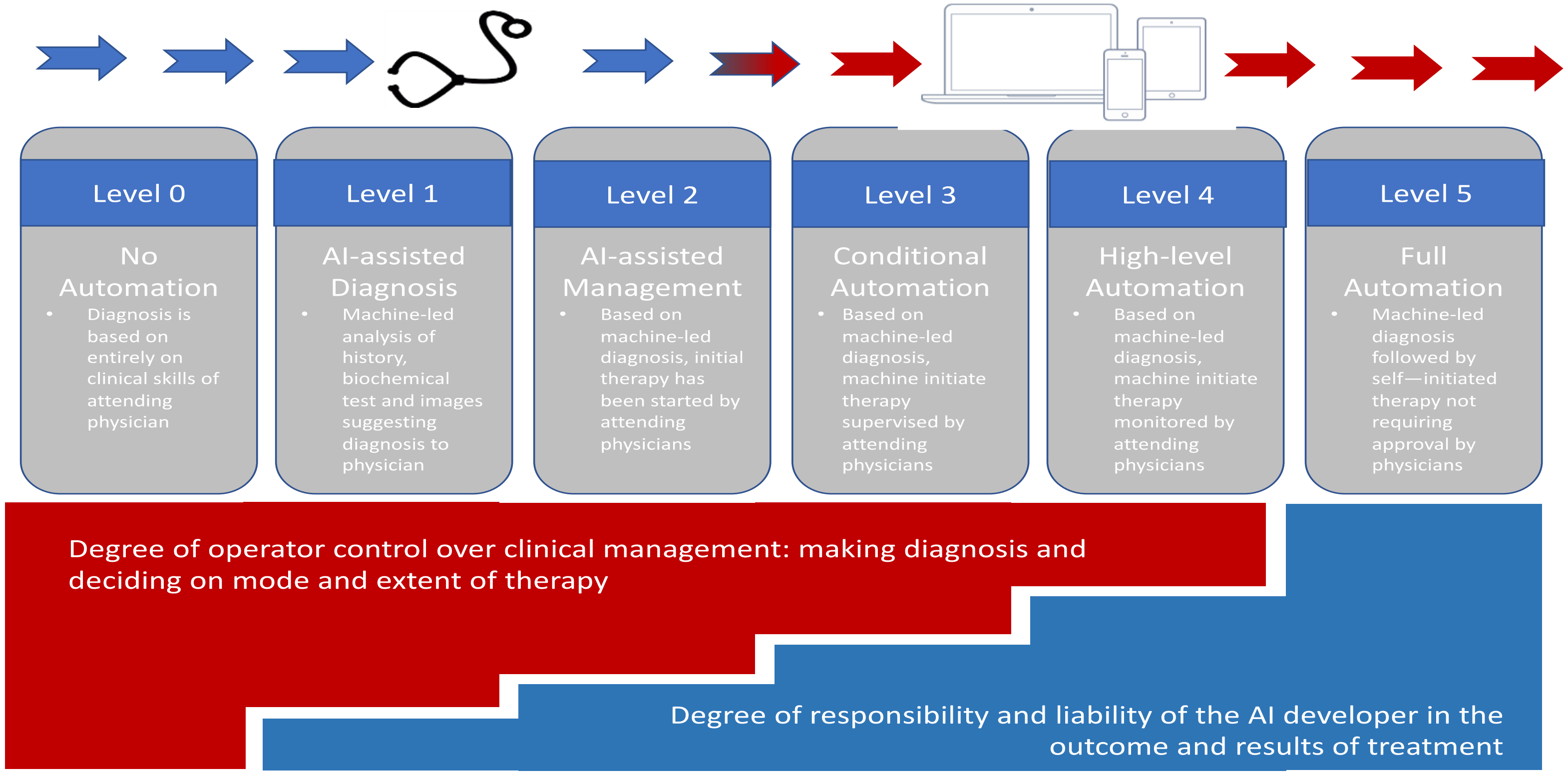
Step-wise shifting of responsibility

- Automobile industry: J3016™ ‘Levels of Driving Automation’ standard, SAE International



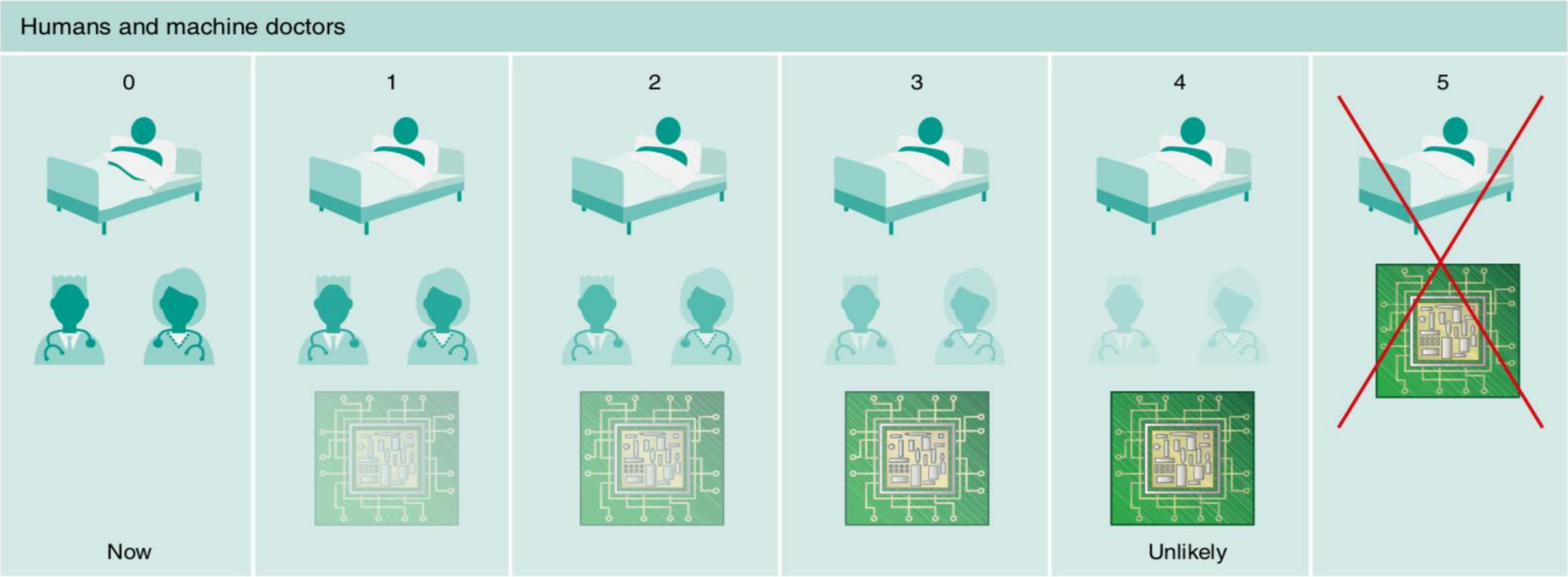
Step-wise shifting of responsibility

- Levels of AI-assisted decision in diagnosis and clinical management



Sung and Poon. JGH 2019

Human driver monitors environment			System monitors environment		
0 No automation	1 Driver assistance	2 Partial automation	3 Conditional automation	4 High automation	5 Full automation
The absence of any assistive features such as adaptive cruise control.	Systems that help drivers maintain speed or stay in lane but leave the driver in control.	The combination of automatic speed and steering control—for example, cruise control and lane keeping.	Automated systems that drive and monitor the environment but rely on a human driver for backup.	Automated systems that do everything—no human backup required—but only in limited circumstances.	The true electronic chauffeur: retains full vehicle control, needs no human backup, and drives in all conditions.



AI will have impact at three different levels

Clinician: Rapid and accurate image interpretation

- Radiology, Pathology, Neurology, Gastroenterology, Cardiology

Health system: improve efficiency and allocation of resources

- Prediction of key outcome- mortality and morbidity

Patient: Promote personal health by processing own data

- Machine vision- clinician handwashing, ICU patient movement, Falling
- Wearable monitors- BP, HR, rhythm, PaO_2, T°C, Respiration
- Nearest-neighbor analysis- providing the easiest access to health service

Will AI take over everything?

- A thousand algorithm producing a thousand management
- Garbage in, garbage out
- Trust… by patients and by doctors
- Predicted outcome: do I really want it?
- Legal responsibility
- Artificial intelligence vs Artificial humanity