UniSR
Università Vita-Salute San Raffaele

UNIVERSITY OF CAMBRIDGE

Collaborative Diagnosis: Bridging Clinicians and AI for Respiratory Health Monitoring

Marco Montagna[1] & Jing Han[2]

[1] Vita-Salute San Raffaele University - San Raffaele Hospital (Milan)
[2] University of Cambridge

European Research Council
Established by the European Commission

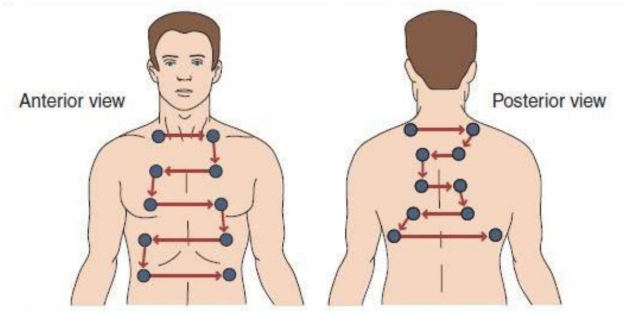
This talk will cover

Part 1: A Real-World Case Study

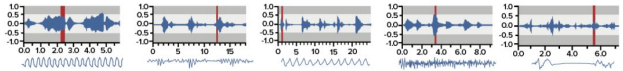
🔍 Key findings from our paper “Evaluating Listening Performance for COVID-19 Detection by Clinicians and Machine Learning: Comparative Study”, J Med Internet Res 2023;25:e44804

Part 2: Our Vision of the Future of AI-Clinician Collaboration

🚀 Exciting opportunities & critical challenges

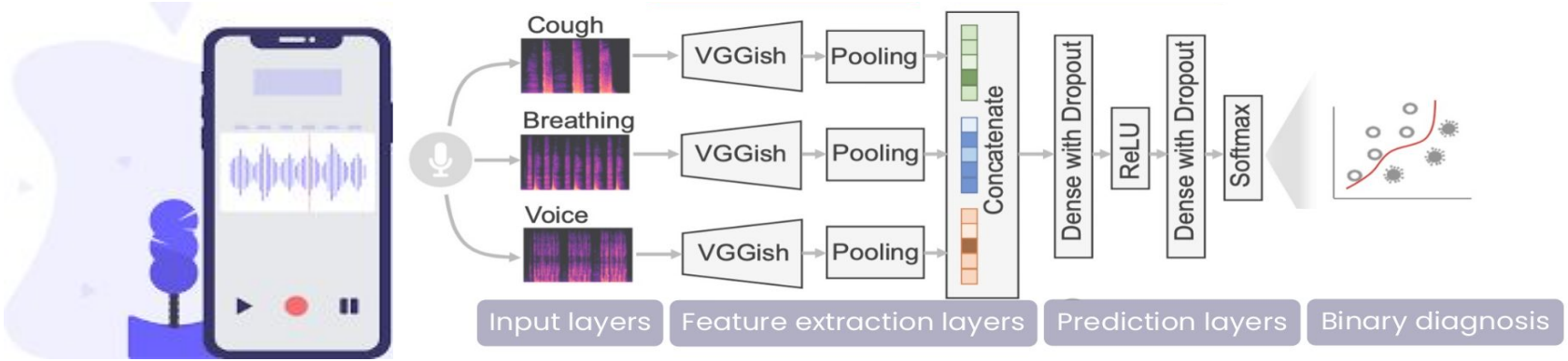


ABNORMAL LUNG SOUNDS



Lung sound	Wheeze	Crackles (rales)	Rhonchi	Stridor	Pleural friction rub
Sounds like	• High-pitched whistling sound • Commonly heard with expiration • May be heard	• Bubbling, popping, or clicking noises • Best heard during inhalation	Low-pitched rattling sounds that may sound like snoring, gurgling, or	High-pitched variable respiratory sound	Creaking, grating, rubbing

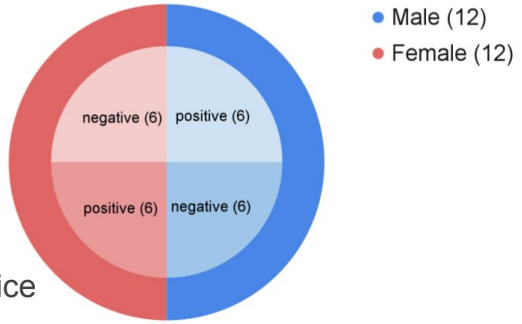
we want to explore the synergy between the two … …



A Real-World Case Study

24 audio clips selected

- covering
  - COVID-19 Positive/Negative
  - (a)symptomatic
  - w/ or w/o asthma
  - non- or heavy smoking
- each audio format:
  - breathing + 1s silence + cough + 1s silence + voice
- total duration ***17*** min



Virtual:

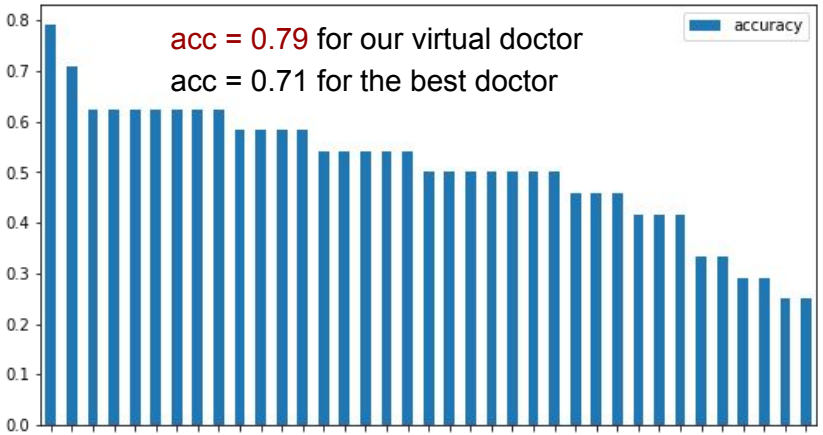
- Machine learning model [Han, J. (2022)]
  - Train and test sets (1162:329)
  - SE = 0.65, SP = 0.69
- Outputs
  - binary decision
  - confidence score

Real:

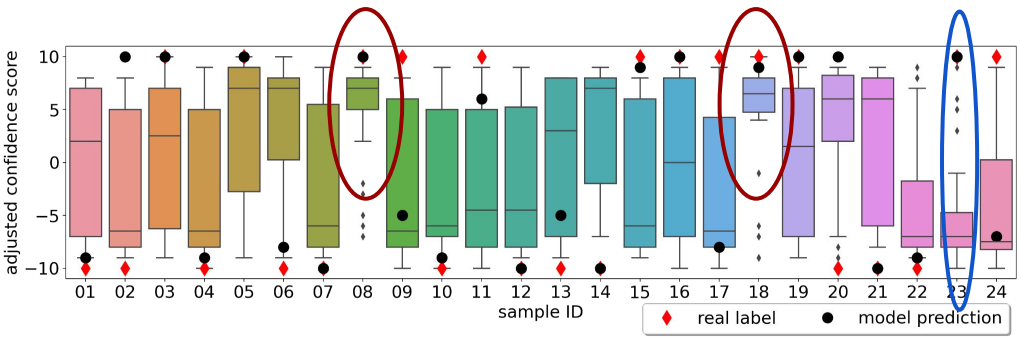
- 36 respiratory clinicians
  - w/ varied experiences
  - listen clips and Answer questions in spreadsheet
- Outputs
  - binary decision
  - confidence score [1-10]

A Real-World Case Study

Performance per single doctor



Decision confidence per audio clip



A Real-World Case Study

Committee of Multiple Members

# Members (Decision weighted)	Acc_max (1000 runs)
4	0.792
5	0.792
10	0.583

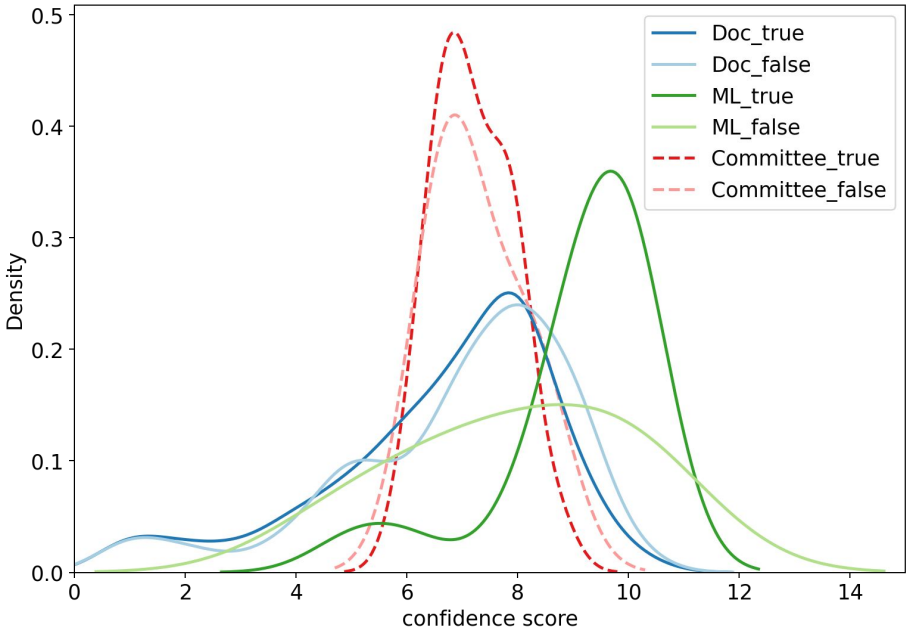
Committee with a Virtual Member

# Members (Decision weighted)	Acc_max (1000 runs)
4 **+1**	0.830
5 **+1**	**0.875**
10 **+1**	0.792

ML model help the clinicians committee reach a more accurate decision.

A Real-World Case Study

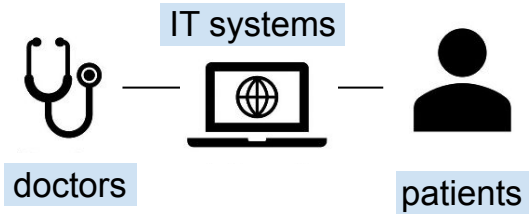
Confidence Score Kernel Density Estimate



Clinicians help alleviate the overconfidence issue of the ML model.

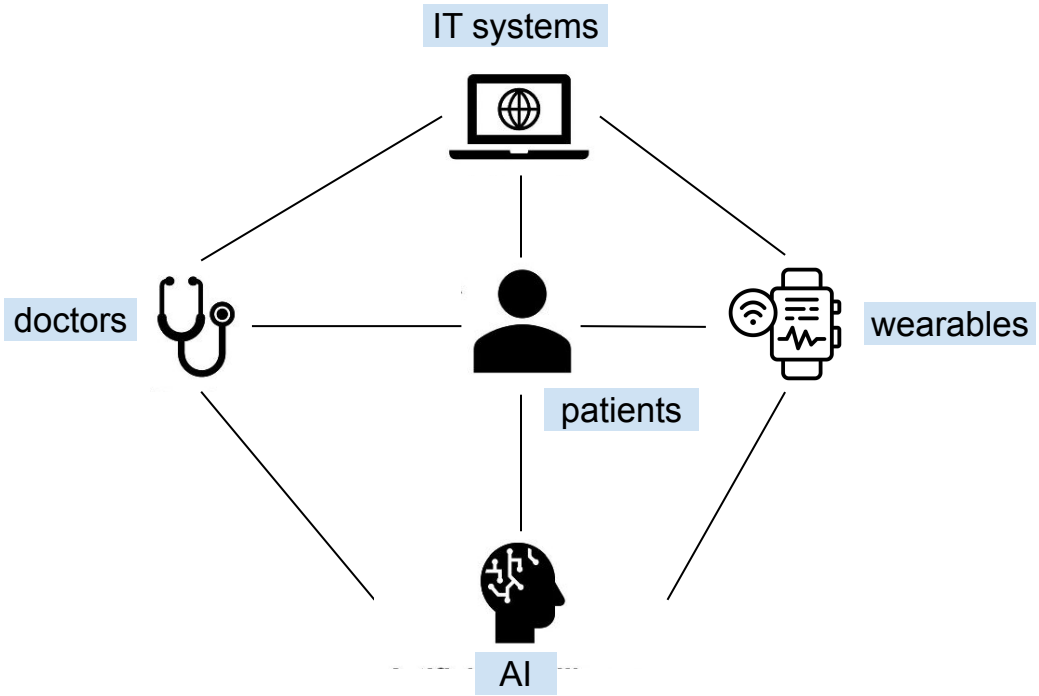
Part 2: Our Vision of the Future of AI-Clinician Collaboration

IT systems appeared as intruders



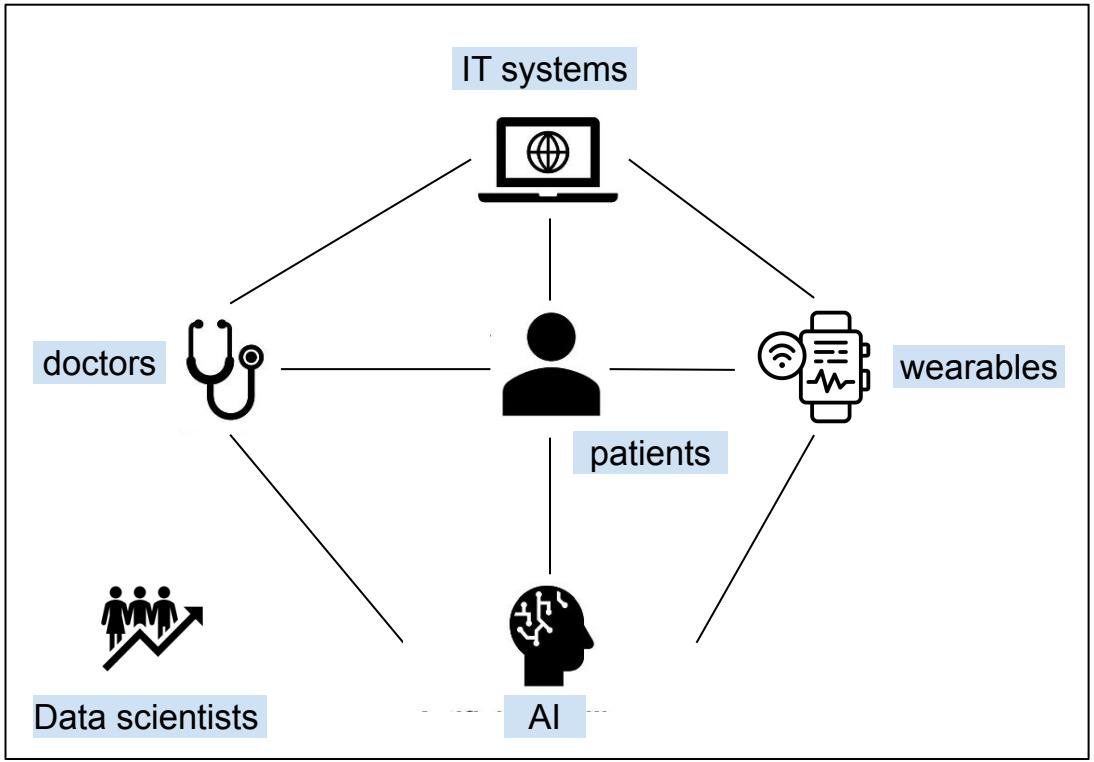
Adapted from Allen, M.R., Webb, S., Mandvi, A. *et al. BMC Prim. Care* (2024)

Patients’ ecosystem now is complex



Adapted from Allen, M.R., Webb, S., Mandvi, A. *et al. BMC Prim. Care* (2024)

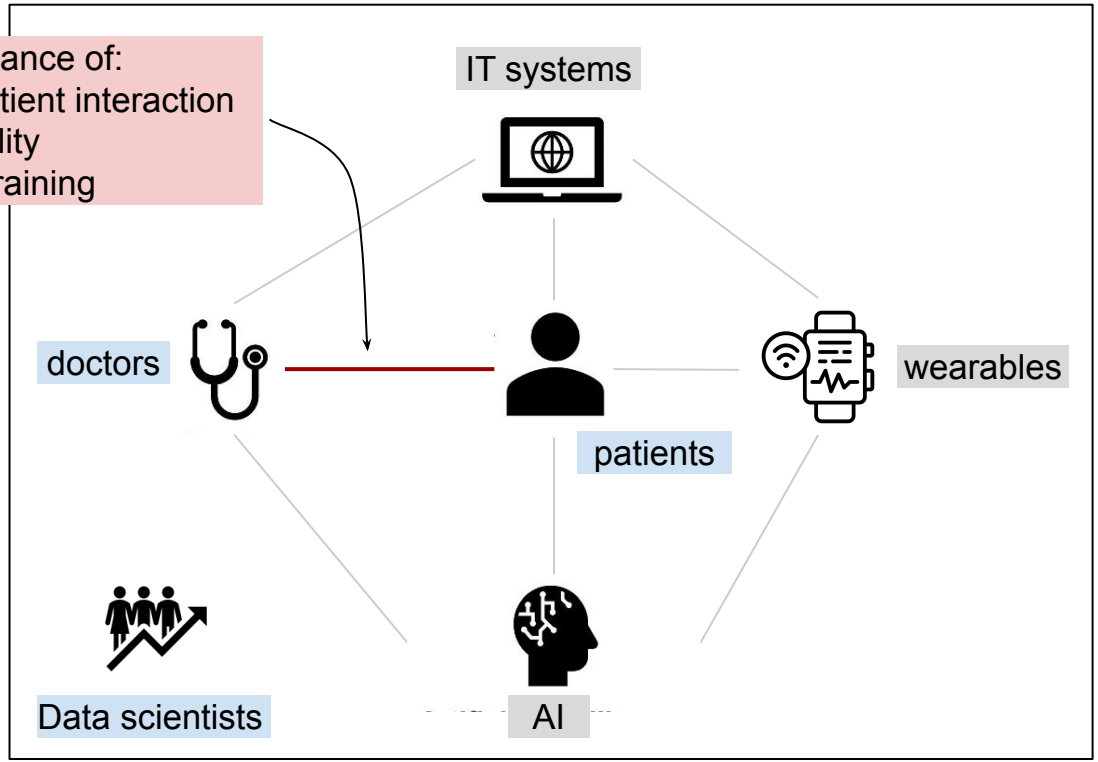
Multidisciplinarity is the new paradigm: the “datascope”



Adapted from Allen, M.R., Webb, S., Mandvi, A. *et al. BMC Prim. Care* (2024)

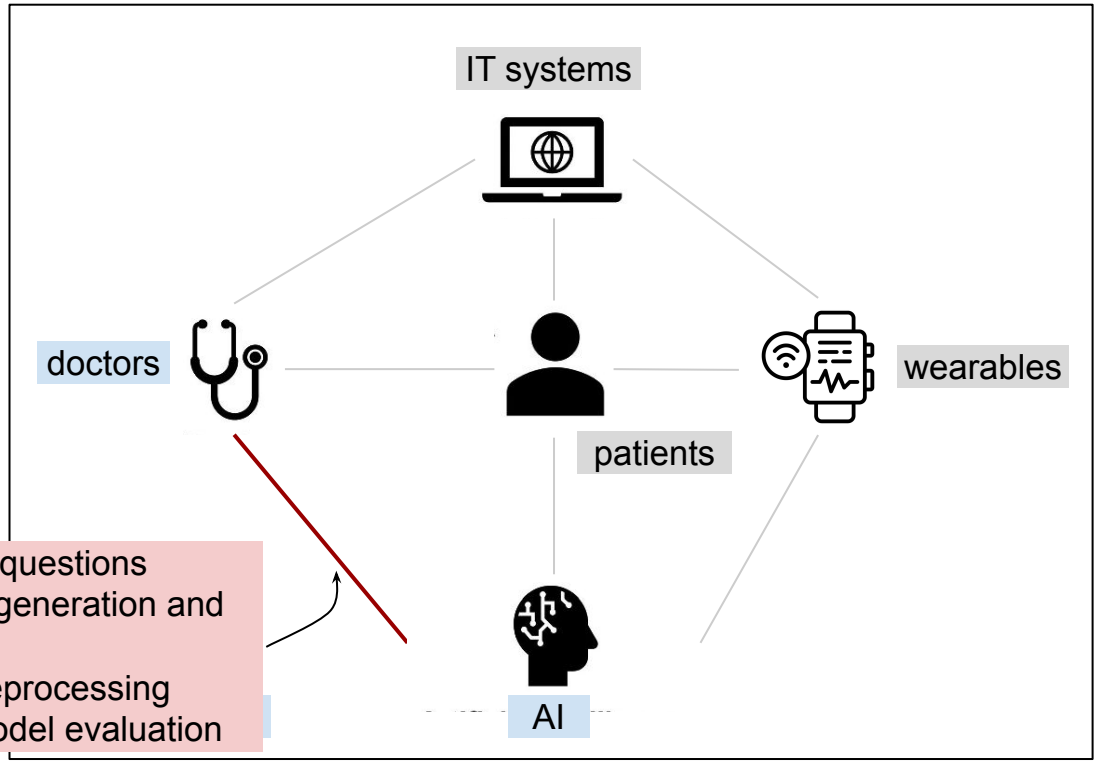
We call for a health-data-informed clinical workforce

Reconsider the importance of:
- Data-diligent patient interaction
- Scientific reliability
- Statistics core training



Adapted from Allen, M.R., Webb, S., Mandvi, A. *et al. BMC Prim. Care* (2024)

Doctors should stay in the loop



- Relevant clinical questions
- Insights on data generation and context
- Assistance in preprocessing decisions and model evaluation

Adapted from Allen, M.R., Webb, S., Mandvi, A. *et al. BMC Prim. Care* (2024)

Human intelligence + Artificial intelligence

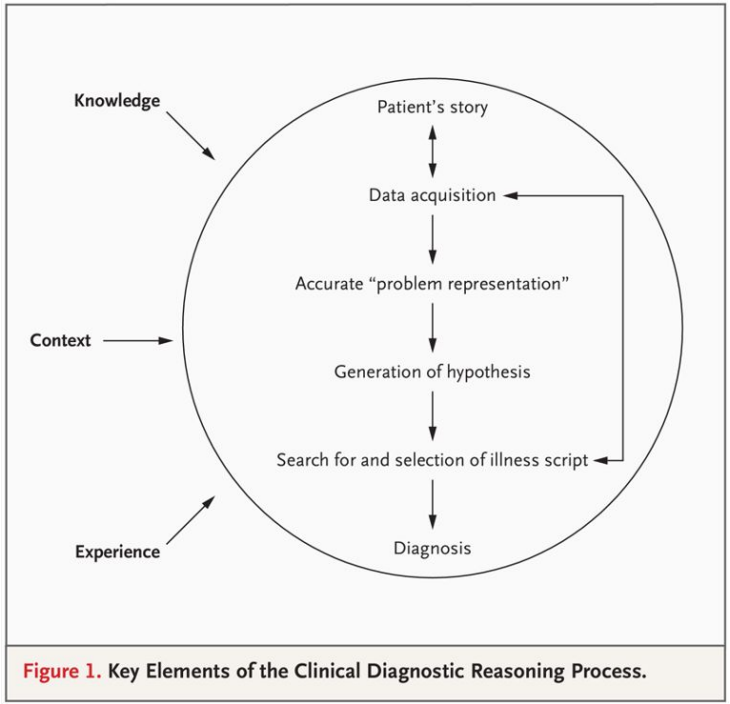


Figure 1. Key Elements of the Clinical Diagnostic Reasoning Process.

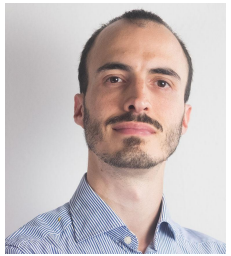
Educational Strategies to Promote Clinical Diagnostic Reasoning. Judith L. Bowen, *NEJM* (2006)

We would like to express **our deepest gratitude** to

- 🩺 **doctors** who contributed to the labeling with their expertise.
- 🤳 **participants** who donated their voices and other information.
- 👥 **researchers** who actively evolved and supported the project.
- 🎤 **audience** today for your valuable questions and feedback.

Jing Han

jh2298@cam.ac.uk

Marco Montagna

montagna.marco@hsr.it