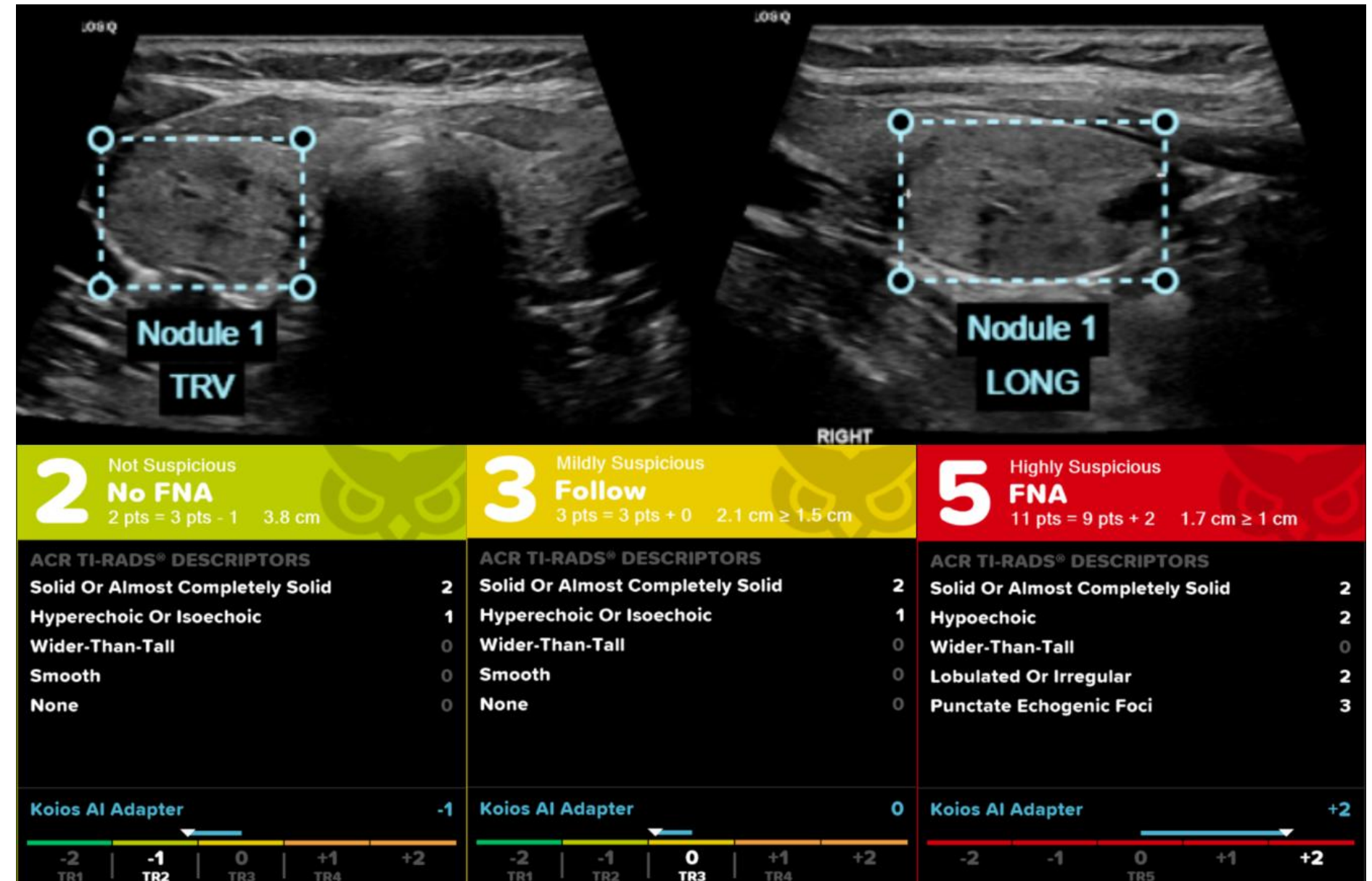


BMUS)))

The Future

Artificial Intelligence

Presented by : Amy Barnes



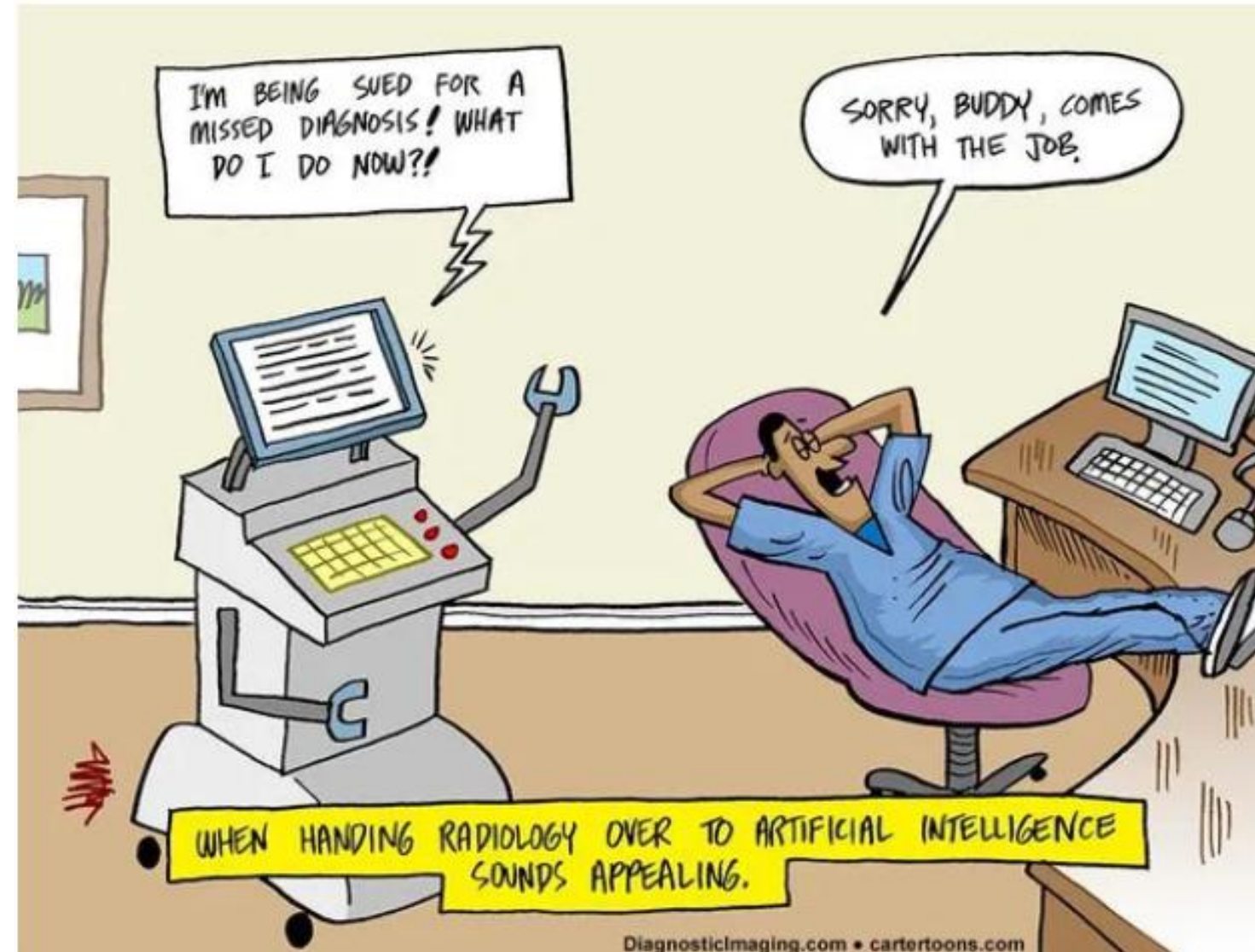
Insert NHS Trust Logo if relevant/preferred

Why are we talking about AI?

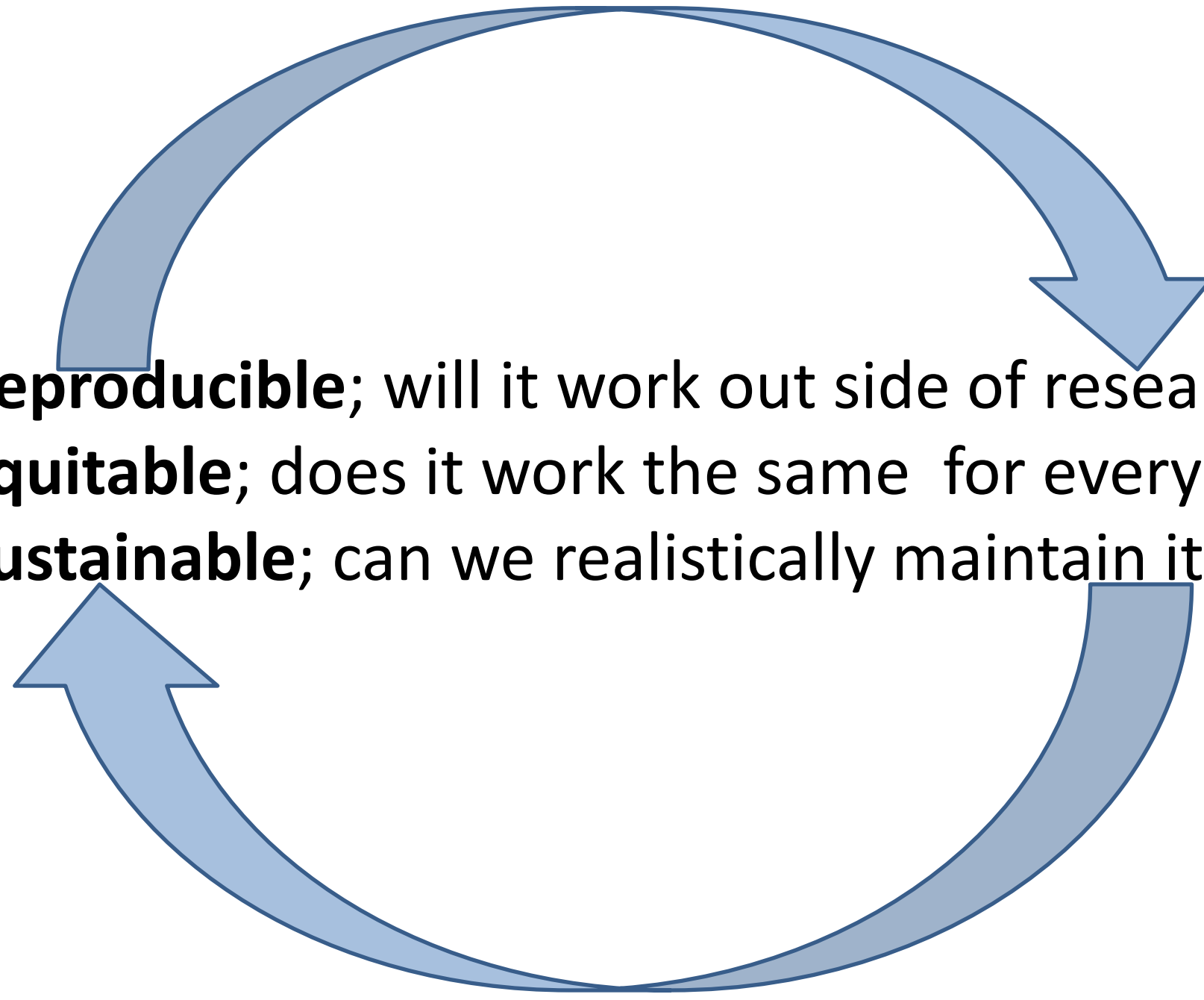
“ people should stop training radiologists now – its just completely obvious within 5 years deep learning is going to do better than radiologists”

Geoffrey Hinton, 2016

What should AI deliver?



What makes “good AI”?

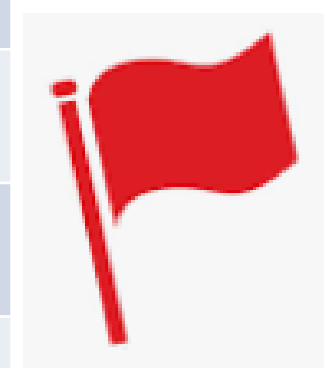
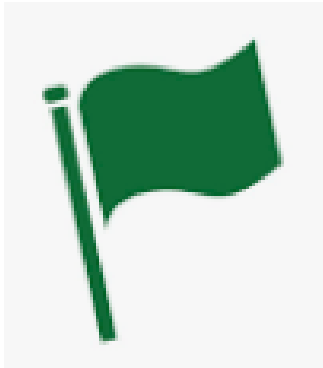


Reproducible; will it work out side of research?
Equitable; does it work the same for everyone?
Sustainable; can we realistically maintain it?

What good AI research should look like?

Not all AI evidence is equal

Green flags	Red flags
Prospective / pragmatic design	Tiny dataset
Clearly designed reference standard	Single centre
Representative population	No external validation
Transparent methodology	Retrospective convenience samples only
External validation	No comparison to clinician performance
Real world implementation thinking	Accuracy claims without workflow/ patient outcomes
Reporting bias acknowledgement	Black box thinking



How to critically review AI?

STARD-AI

Clear patient selection (*diversity*)

Index test clearly described (*AI*)

Human comparator (*operator*)

Reference standard defined (*BTA / Thy*)

Diagnostic pathway described (*SoC*)

Threshold pre specified (*p-hacking*)

Missing data explained (*why, mitigation*)

Performance reported transparently

STANDING Together

Who was included?

Who was excluded?

Is the dataset diverse?

Could bias affect performance?

Would this work in my population?

[The STARD-AI reporting guideline for diagnostic accuracy studies using artificial intelligence | EQUATOR Network](#)

Lancet Digit Health. 2025 Jan;7(1):e64-e88.

doi: 10.1016/S2589-7500(24)00224-3. Epub 2024 Dec 18.

AI must earn its place...

- Does it solve a real clinical problem?
- Is the evidence trustworthy?
- Will it fit into practice?
- Is it fair and sustainable?
- Is it safe?



Investigation of artificial intelligence–based clinical decision support system’s performance in reducing the fine needle aspiration rate of thyroid nodules: A pilot study

Amy Barnes¹ , Rebecca White², Heather Venables² , Vincent Lam³ and Ram Vaidhyanath³

Ultrasound
1–9
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1742271X241299220
journals.sagepub.com/home/ult



- n=29
- 31% reduction in FNA
- TI-RADS

TI- RADS -> BTA project

Amy Barnes
Dr Steve Colley
Dr Andrew McQueen
Dr Ram Vaidyhanath

FUNDED BY

NIHR | National Institute for
Health and Care Research

Diagnostic Accuracy Study Cost Consequence Analysis Focus group and Interviews Consensus



20
Years

Thyroid segmentation project for machine learning (AI)

- Looking for expert volunteers to perform annotation on an anonymised large clinical dataset of ultrasound examinations, from the Leeds Teaching Hospital Trust (LTHT).
- Full dataset is intended for release (via a restricted data repository) for open research into segmentation models and benchmarking.
- If you are interested, please contact James McLaughlan (j.r.mclaughlan@leeds.ac.uk) for more information.

Image details

Multi-class segmentation

Ultrasound settings

Radiology report

- Enlarged thyroid
- Single thyroid nodule
- U3 grading

Cytology report

- FNA
- Benign

An example of the Label Studio platform that would be used for this process (all website based no local install etc needed).

The future.....

BMUS AI CIG & ASM AI stream.....

Thank you