#5659: Implementing an innovative safety-checkpoint process for clinical deployment of AI in breast cancer screening: First results and experiences at 14 sites

A. Ng[1], G. Fox[1], C. Oberije1, J. Nash[1], S. Kerruish[1], C. Brackstone[1], A. Arkle-Howard[1], A. Leaver[2], L. Bisset[3], R. Bradley[4], A. Christou[5], A. Ford[6], N. Forester[7], A. Helal[8], A. Kumar[9], N. Layt[10],
S. Puri[11], A. Redman[2], F. Razzaq[12], B. Rengabashyam[4], A. Talwalkar[13], O. Strukowska[5], W. Teh[14], B. Glocker[1, 15], P. Kecskemethy[1]

[1]Kheiron Medical Technologies, London, United Kingdom, [2]Gateshead Health NHS Foundation Trust, Gateshead, United Kingdom, [3]University Hospitals Dorset NHS Foundation Trust, Dorset, United Kingdom, [4]York and Scarborough Teaching Hospitals NHS Foundation Trust, York, United Kingdom, [5]University Hospitals Sussex NHS Foundation Trust, United Kingdom, [6]Chesterfield Royal Hospital NHS Foundation Trust, Chesterfield, United Kingdom, [7]Newcastle Upon Tyne Hospitals NHS Foundation Trust, Newcastle, United Kingdom, [8]University Hospitals of North Midlands NHS Trust, North Midlands, United Kingdom, [9]Frimley Health NHS Foundation Trust, Frimley, United Kingdom, [10]Northampton General Hospital NHS Trust, Northampton, United Kingdom, [11]University Hospitals of Derby and Burton NHS Foundation Trust, Derby, United Kingdom, [12]Warrington and Halton Teaching Hospitals NHS Foundation Trust, Warrington, United Kingdom, [13]Wrightington, Wigan and Leigh NHS Foundation Trust, Wigan, United Kingdom, [14]Royal Free London NHS Foundation Trust, London, United Kingdom, [15]Department of Computing, Imperial College London, United Kingdom.

deephealth

Disclosures

Dr. Peter Kecskemethy was the CEO and co-founder of Kheiron Medical Technologies and is the Head of Business Development for AI at DeepHealth.

Ethics Committee Approval

Ethics committee approval was not required as it was local service evaluation of a CE certified medical device

Funding for this study

Funding was received from the AI in Health and Care Award.

deephealth

Purpose

To demonstrate the implementation of initial stages of a safe deployment process for AI.

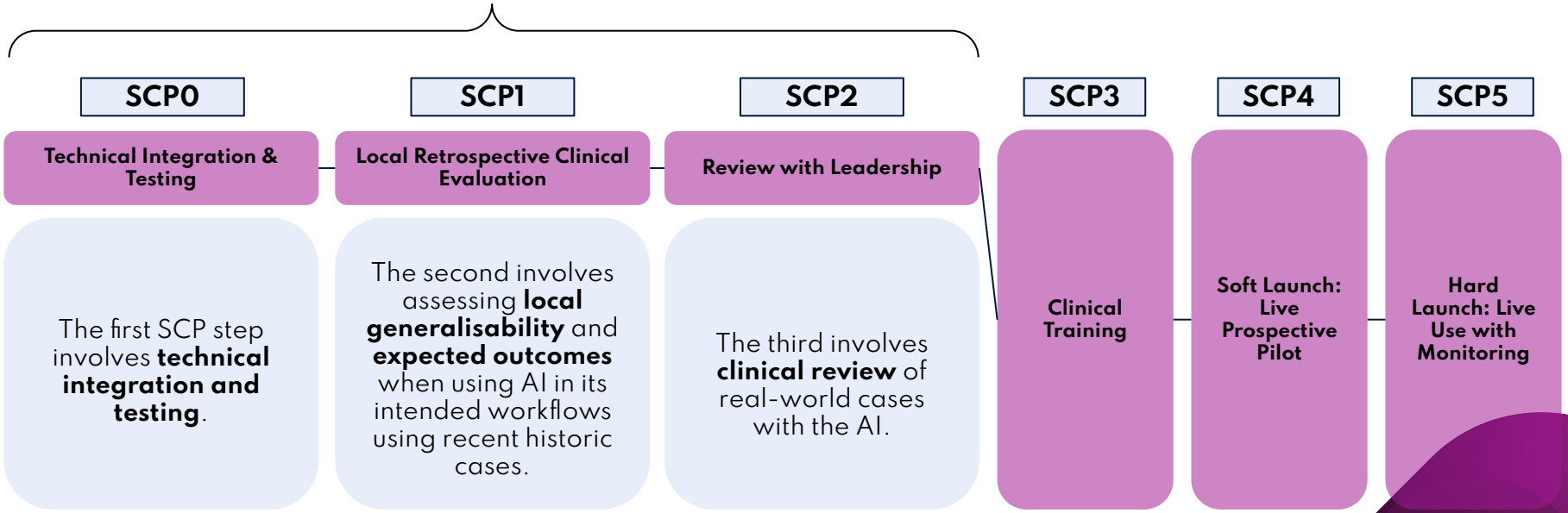
This process evolved through extensive experience in developing, validating, deploying, monitoring and working with AI in clinical practice.

Background

- AI shows promise for improving the quality of breast screening.
- Study results are encouraging about AI, but the translation of study performance into everyday practice is not guaranteed.
- A step-wise, safety-focused deployment process with comprehensive monitoring is required to confirm whether experimentally measured benefits translate into real-world deployments.

Methods – Safety-checkpoint (SCP) process

The initial 3 stages of a safety-checkpoint (SCP) process was implemented with a commercially available AI system at 14 sites:



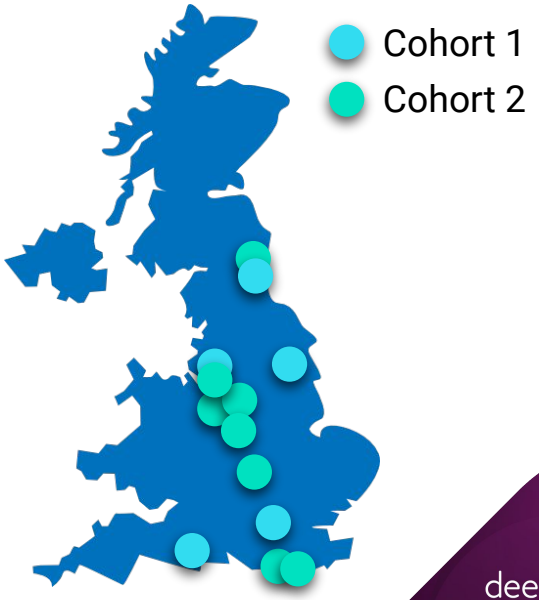
Methods – Site cohorts

The 14 deployment sites covered different regions of the UK and were never previously used in the AI's development or validation.

The safety-checkpoint process was implemented at the sites in 2 cohorts:

Cohort 1: 6 sites - underwent SCP 0 and 2

Cohort 2: 8 sites - underwent SCP 0, 1, and 2



Methods – Safety-checkpoint (SCP) process

SCP 0:

- Technical integration with the site's PACs and RIS was conducted.

SCP 1:

- A checklist was developed to ensure standardised reporting and facilitate performance comparisons.
- The workflow presented for SCP 2 uses AI as Reader-2 when the AI and Reader-1 agreed to not recall, otherwise the human Reader-2 opinion was used (Figure 1).

SCP 2:

- The AI was deployed as a 'silent' reader of prospective cases.
- Clinical teams reviewed AI-flagged cases not recalled by standard double reading (DR) (positive discordant cases).
- Cases categorised as a type-3 interval cancer (i.e. missed cancer with strong visible signs on the base screen) were recalled.

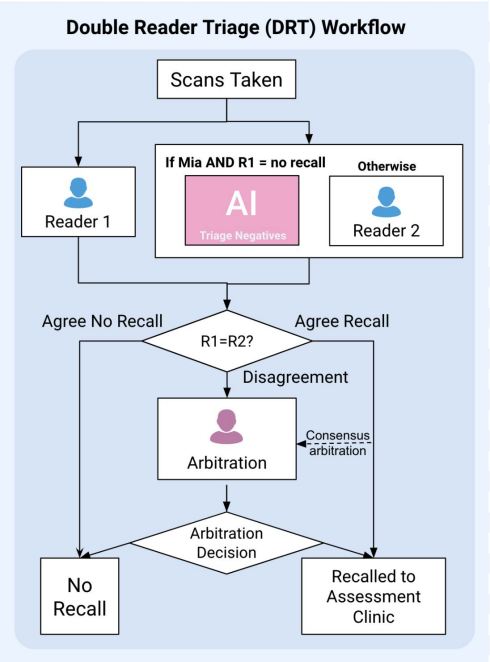
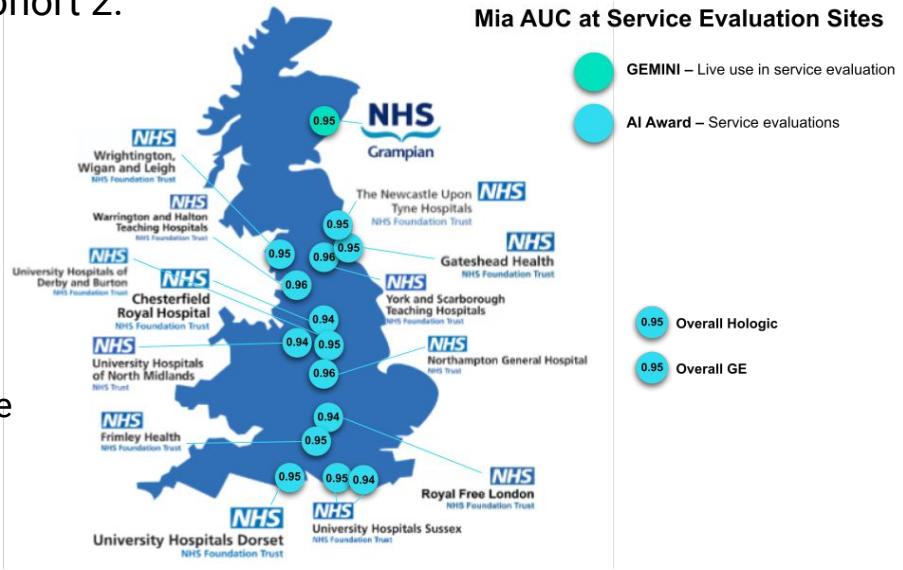


Figure 1

Results – SCP 1: Generalisability across sites

- 118,584 cases were included in SCP-1 across Cohort 2.
- AI's AUC ranged from 0.94-0.96 across sites.
- Modelling the double reading triage workflow with AI across sites resulted in an expected:
  - relative -3.0% recall rate
    (reduction range: 1.6-4.4%),
  - relative -0.3% 'screen-detected' cancer detection rate
    (reduction range: 0.0-1.1%)
  - relative +2.8% positive predictive value
    (increase range: 1.6-4.6%)
  - 20-28% workload savings.



deephealth

Results – SCP 2: Clinical reviews

111,2604 eligible cases were included in the SCP 2 clinical review period across the 14 deployment sites

→ 10,028 cases (9.0%) were positive discordant cases – all were reviewed

→ 48 women (0.4%) were recalled from these reviews

→ 28 additional type-3 interval cancers were found across 27 women, increasing cancer detection rate by 0.25/1000

Based on similar evaluations such as the GEMINI evaluation, additionally recalling type-2 cancers (with only minimal signs of cancer on the base screen), a cancer detection rate increase of +1/1000 is expected.

Conclusion

- The use of a specially-designed safety-checkpoint process for clinical deployment of AI has been demonstrated.
- Conducting it at 14 sites with varying levels of clinical experience with AI enabled a standardised way of confirming the generalisability of the AI system and it's expected benefits translating locally and showed that the approach is scalable.
- Trialling the initial stages of the safety-checkpoint process across multiple sites from different regions has demonstrated its suitability for moving safely towards the next safety-checkpoints: training and live use with monitoring.

Limitations

- Only a single AI was evaluated
- Only the first 3 steps of the safety checkpoint process was assessed