

## From national policy to Trust-level practice: designing and evaluating a structured mentorship pilot for International Medical Graduate and Locally Employed doctors

Vosinakis C<sup>1,2</sup>, Chaudhary P<sup>3</sup>, Booth F<sup>4</sup>, Alia R<sup>4,5</sup>, Kastelik J<sup>4,6</sup>, Sayeed S<sup>7</sup>

1. Research & Development Department, Hull University Teaching Hospitals NHS Trust, Humber Health Partnership, Castle Hill Hospital, Castle Road, Cottingham, East Yorkshire HU16 5JQ, UK
2. Department of Plastic and Reconstructive Surgery, Hull University Teaching Hospitals NHS Trust, Humber Health Partnership, Castle Hill Hospital, Castle Road, Cottingham, East Yorkshire HU16 5JQ, UK
3. Department of Medicine for the Elderly, Hull University Teaching Hospitals NHS Trust, Humber Health Partnership, Hull Royal Infirmary, Anlaby Road, Hull HU2 3JZ, UK
4. Medical Education Centre, Hull University Teaching Hospitals NHS Trust, Hull Royal Infirmary, Anlaby Road, Hull, HU3 2JZ
5. Department of Paediatric Medicine, Hull University Teaching Hospitals NHS Trust, Humber Health Partnership, Hull Royal Infirmary, Anlaby Road, Hull HU2 3JZ, UK
6. Department of Respiratory Medicine, Hull University Teaching Hospitals NHS Trust, Humber Health Partnership, Castle Hill Hospital, Castle Road, Cottingham, East Yorkshire HU16 5JQ, UK
7. Department of Vascular Radiology, Hull University Teaching Hospitals NHS Trust, Hull Royal Infirmary, Humber Health Partnership, Anlaby Rd, Hull HU2 3JZ, UK

### Corresponding Author:

Mr Christos Vosinakis [christos.vosinakis@nhs.net](mailto:christos.vosinakis@nhs.net)

---

### Abstract

International medical graduates now constitute 42% of UK licensed doctors, yet documented differential attainment in postgraduate assessment, sustained departures from NHS employment and the absence of structured, evaluated, dyadic mentorship aspiring to Kirkpatrick Level 3 represent an unresolved implementation gap in national policy. This paper describes the design of a structured six-month mentorship pilot at Hull University Teaching Hospitals NHS Trust, comprising ten mentor-mentee pairs matched through an eight-factor scored matrix weighted towards clinical and career alignment. Three mandatory structured meetings are held at months one, three and six, with a mid-pilot review and mentee-completed check-in forms capturing progress and wellbeing after each session. The evaluation framework triangulates Kirkpatrick Levels one to three, Moore's continuing

professional development model and the RE-AIM implementation framework, with outcome measures captured at baseline and end of pilot. This QIP-classified pilot aims to close the gap between national policy and local implementation, providing a replicable, low-cost framework for Trusts to support a workforce where connection to the NHS has halved in only three years.

## Introduction

International medical graduates (IMGs) now account for 42% of all licensed doctors in the UK, some 138,405 individuals<sup>1</sup>. The Locally Employed Doctor (LED) workforce in England and Wales has more than tripled since 2014 to over 38,000, and the Specialty and Associate Specialist (SAS) grade has grown by 48% over the same period to approximately 12,500<sup>2</sup>. In 2024, 20,060 IMGs joined the register and 4,880 left, (3,869 in 2023), against a stable pre-pandemic (2017-2019) departure rate of approximately 3.5-4% of the non-UK PMQ licensed workforce. Over the same period, the proportion of PLAB joiners still connected to a designated body six months after arrival fell from one in four to one in eight<sup>1</sup>. Recruitment has plateaued and retention has become the defining workforce question.

IMGs are also more likely than UK graduates to struggle in postgraduate assessment. In the membership examination of the Royal College of General Practitioners, where IMGs now make up 58% of all candidates, the first-time pass rate in 2024 to 2025 was 87% for UK graduates against 65% for IMGs in the Applied Knowledge Test, and 94% against 54% in the Simulated Consultation Assessment<sup>3</sup>. Training in a country where the primary medical qualification is not taught in English is independently associated with lower odds of passing

and lower scores across every component of the examination<sup>4</sup>. In surgery, black and Asian candidates pass the MRCS Part B at adjusted odds ratios of 0.41 (95% CI 0.18-0.92) and 0.49 (0.35-0.69) relative to white candidates respectively<sup>5</sup>. Country of primary medical qualification is the strongest single predictor of an unsuccessful application to intensive care medicine training<sup>6</sup>.

These gaps are not explained by ability alone, and the evidence does not point to overt examiner bias as the principal cause. A national qualitative study of postgraduate trainees identified the relationship with senior colleagues as the most powerful channel of learning and progression, and the one most vulnerable to the effects of perceived bias, cultural distance and a lack of trust, with IMGs frequently turning to peers rather than seniors for support<sup>7</sup>. The policy response has been consistent for over half a decade. The 2019 review *Fair to Refer* recommended that enhanced induction should include allocating a mentor<sup>8</sup>. NHS England set out minimum standards for the induction of IMGs in 2022<sup>9</sup>. The General Medical Council committed in 2021 to eliminate disadvantage in medical education and training by 2031<sup>10</sup>. The NHS Equality, Diversity and Inclusion Improvement Plan made comprehensive induction for internationally recruited staff a high-impact action<sup>11</sup>. The 2026 survey of LE and SAS doctors will, for the first time since 2019, give individual trusts named data on the experience of this workforce<sup>2</sup>. Existing Trust-level evaluations report on induction rather than structured dyadic mentorship and typically evaluate at Kirkpatrick Levels 1 and 2 only<sup>12-14</sup>. Guidance is abundant; structured implementation at trust level is not. This pilot asks whether a structured, time-limited and evaluated mentorship programme, rather than an informal

buddy arrangement, can close the gap between national recommendation and local delivery, and generate the evidence a trust needs to decide whether to embed it.

## Methods

The pilot enrolls up to ten mentor-mentee pairs at Hull University Teaching Hospitals NHS Trust (HUTH) for a period of six months. Three structured meetings are mandatory (at months one, three and six) with a built-in mid-pilot review at the month-three meeting (Figure 1). Cohort size is calibrated to allow meaningful qualitative evaluation while remaining administratively feasible for a single Associate Medical Education Fellow as programme lead. Six months is the minimum duration at which Kirkpatrick Level 3 outcomes become evaluable within a single academic year<sup>15</sup>. The dyadic design is grounded in Kram's foundational framework of career and psychosocial mentor functions, supported by a systematic review of 39 studies establishing dyadic mentorship's influence on career choice, productivity and personal development<sup>16,17</sup>. A structured but flexible design is consistent with the reframing of mentorship as a complex adaptive system: the absence of a fixed syllabus enables the pair to respond adaptively to mentee needs that vary and cannot be prescribed in advance<sup>18</sup>.

Mentors are eligible from specialty training year three or above with at least twelve months of NHS experience, or from the SAS and LED grades, or as consultants. Mentees are IMGs and LED doctors at any grade. Pairs are matched using a scored eight-factor matrix before the programme begins, with factors weighted toward clinical and career alignment. The three highest-weighted factors (specialty alignment, career and pathway match, and needs match)

each carry twice the weight of the remaining five, which address site compatibility, meeting format preferences, availability and whether the mentee has expressed a preference for an IMG mentor. The matrix addresses the evidence that cultural distance and lack of trust from seniors drive IMGs toward peer rather than senior support in the absence of structured [pairing](#)<sup>7</sup>. Systematic review evidence identifies deliberate matching as a consistently active mechanism in successful acculturation programmes<sup>19</sup>. Each match is confirmed jointly by the Programme Lead and IMG Tutor, with rationale recorded.

The mentee completes a structured check-in form after each meeting, capturing progress against stated personal goals, the degree of psychological safety experienced in the relationship and any concerns requiring escalation. Within the educator-learner dyad, psychological safety, the shared belief that the relationship is safe for interpersonal risk-taking, is the mechanism by which structured mentoring translates into earlier disclosure of difficulty and earlier access to support<sup>20</sup>. The month-three check-in includes a brief wellbeing pulse, the responses to which are not shared with the mentor and are held as programme-evaluation data. End-of-pilot evaluation captures Kirkpatrick Levels one to three through paired measures across eleven domains (Figure 2), using one-to-ten Likert scales<sup>15,21</sup>. Reach, adoption and longer-term sustainability prospects are assessed against the RE-AIM implementation framework<sup>22</sup>.

The evaluation framework triangulates Kirkpatrick, Moore's continuing professional development model and RE-AIM<sup>15,18,21,22</sup>. Within RE-AIM, reach is operationalised through

recruitment numbers relative to the feasibility thresholds; adoption through mentor willingness to continue; and sustainability through the MEC Committee scaling decision. The explicit ambition to evaluate at Kirkpatrick Level three distinguishes this pilot from the published comparator UK Trust-level induction programmes identified in the literature, which report at Levels one and two only<sup>12-14</sup>. Pre-specified minimum thresholds for completion, impact and feasibility, and the four-scenario decision logic that follows from them, are shown in [Tables 1](#) and [2](#). The pilot is formally classified as a Quality Improvement Project (QIP). The UK Policy Framework for Health and Social Care Research defines research as the attempt to derive generalisable or transferable new knowledge and explicitly excludes QIPs, service evaluations and audits<sup>23</sup>. The Health Research Authority online decision tool, completed in May 2026, confirmed that the pilot would not be considered research by the NHS; the output is available from the corresponding author on request.

Figure 1. Programme structure: HUTH IMG/LED Mentorship Programme, Pilot Cohort 1

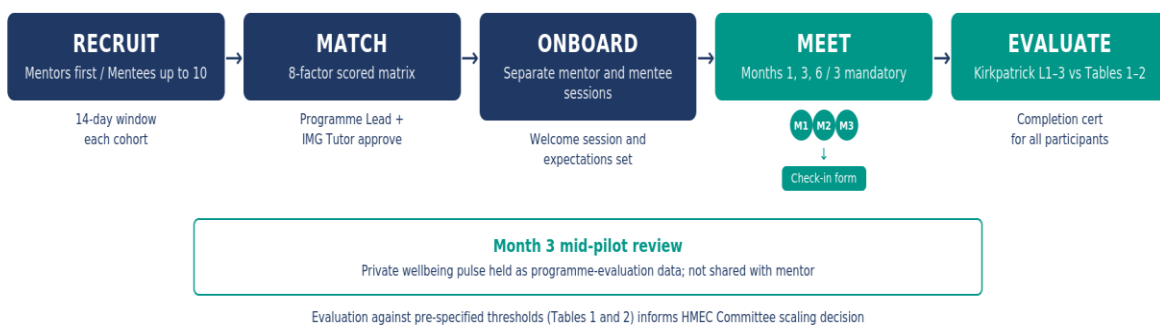


Figure 2. Evaluation domains: eleven paired baseline to end-of-pilot measures. All domains rated on a 1 to 10 Likert scale at baseline and end of pilot.

Trust Connection	Clinical and Psychological Safety	Career Agency
<p>How supported do you feel in the Trust?</p> <p>How much do you feel you belong in your team?</p> <p>Confidence with NHS processes (study leave, Datix, exception reporting, escalation)</p> <p>Understanding NHS workplace culture and expectations</p>	<p>Confidence escalating concerns or calling seniors</p> <p>Safety to ask questions or admit uncertainty at work</p> <p>Current wellbeing at work</p> <p>Knowing where to seek help for wellbeing or workplace concerns</p>	<p>Clarity about UK career pathways (training vs Portfolio Pathway)</p> <p>Confidence about portfolio evidence and applications</p> <p>Manageable work-life balance</p>

All 11 domains rated 1-10 at baseline (application) and end of pilot (feedback form). Success threshold: mean improvement  $\geq +1.0$  point across all domains.

Table 1. Pilot Success Threshold Framework

Domain	Minimum threshold	Notes
Completion	At least 7 of 10 pairs complete at least 2 of 3 structured meetings	Pairs withdrawing due to circumstances beyond the programme (e.g. maternity leave, sickness) noted separately and not counted as failures

Domain	Minimum threshold	Notes
	At least 5 of 10 pairs complete all 3 structured meetings	Aspirational target; used alongside the 2-meeting threshold
	End-of-pilot feedback form completion rate at least 70% of eligible participants	Mentor and mentee forms counted separately
	Zero unresolved safeguarding or escalation events at pilot closure	Any events must have been managed through Trust channels
Impact	Mean improvement of at least +1.0 point across all 11 wellbeing and confidence domains (paired baseline to end-of-pilot, 1 to 10 scales)	End score minus baseline per domain, averaged across responding mentees
	Overall programme usefulness mean at least 7.0 out of 10	Mentee end-of-pilot feedback form, Section 2

Domain	Minimum threshold	Notes
	Mentor-mentee match satisfaction mean at least 7.0 out of 10	Both mentor and mentee end-of-pilot feedback forms
	At least 70% of mentees rate likelihood to recommend at least 7 out of 10	Mentee end-of-pilot feedback form, Section 2
Feasibility	At least 8 mentors recruited for 10 pairs	Indicates programme attractiveness to potential mentors
	At least 10 eligible mentee applications received	Indicates demand
	No escalation of administrative burden to senior team during monitoring period	Assessed by Programme Lead at mid-pilot review

Domain	Minimum threshold	Notes
	Mentor administrative ease mean at least 6.0 out of 10	Mentor end-of-pilot feedback form, Section 2

Table 2. Post-Pilot Scaling Decision Matrix

Scenario	Recommended decision
All completion, impact and feasibility thresholds met	Proceed to Cohort 2 with minor refinements based on qualitative feedback
Completion and feasibility thresholds met; impact thresholds partially met	Proceed to Cohort 2 with targeted improvements to matching or programme structure; present pilot findings as early-stage evidence
Completion or feasibility thresholds not met; impact data promising	Run revised Cohort 1b with structural changes before scale-up; identify and address specific completion or feasibility barriers

Scenario	Recommended decision
Multiple thresholds not met across domains	Do not scale; full review; identify root causes; consider whether the programme model requires fundamental redesign

## Results

The pilot is scheduled to launch in Q3 of 2026. No outcome data currently exist. Tables 1 and 2 pre-specify the minimum thresholds and decision logic against which Cohort 1 findings will be judged when evaluation is complete.

## Discussion

The Medical Education Centre (MEC) Committee will use the thresholds in Table 1 and the decision logic in Table 2 to determine whether to proceed to Cohort 2, introduce targeted refinements, run a revised pilot, or stand down the programme. This positions the HUTH pilot as the operational implementation layer the national policy architecture has not yet specified: national guidance has operated at policy and standards level; Royal College programmes at examination and pastoral support; Trust-level structured delivery has remained the gap. The pre-specified impact thresholds address the equity argument. The feasibility thresholds address the workforce retention argument. Neither displaces the other.

The principal limitations are single-site design, ten pairs and no comparator group. Findings will represent observations from Cohort 1 at one teaching trust and are not designed to be generalisable. The programme lead also leads the evaluation, a position that carries risk of optimism bias toward positive findings and reduced sensitivity to programme weaknesses.

Structural mitigations include independent IMG Tutor oversight (S.S.), MEC Committee scrutiny and the pre-specification of thresholds and decision criteria prior to launch, which limits post-hoc rationalisation of outcomes. End-of-pilot measures are primarily self-reported. Objective Kirkpatrick Level 3 evidence will be limited at six months. A methods-plus-outcomes manuscript on this pilot specifically, drawing on a second cohort with longer follow-up, is planned for submission to a PubMed-indexed journal on completion of Cohort 1 evaluation.

## Conclusion

National guidance on IMG and LED support in the NHS has been unbroken since 2019. Programmes that strengthen the relational and professional infrastructure around IMGs and LEDs have the potential to reduce both the attainment gap and the departure rate that now stands at 4,880 leavers a year. The HUTH pilot is a measurable, evaluable response to the gap between that guidance and Trust-level delivery. The 2026 GMC LE/SAS survey will make Trust performance on this workforce visible for the first time in seven years. Trusts that have built structured support will be positioned to answer what it reveals. Mentorship is not expensive; the absence of it is.

## Practice Points

- Senior relationships are high stakes: Relationships with senior clinicians are the most powerful drivers of IMG progression but are the most vulnerable to bias and cultural distance.
- The retention window is closing: The proportion of new IMG joiners connected to a designated body within six months has halved since 2021; structured support is now a retention necessity, not an optional extra.
- Beyond "buddying": Effective Trust-level support requires structured matching weighted by career goals and structured mentor selection to ensure quality and methodological rigour.
- Evaluate at Kirkpatrick Level 3: Implementation should move beyond "satisfaction" surveys to measure changes in behaviour, such as increased confidence in escalating concerns or admitting uncertainty.
- Low-cost, high-governance: A structured 5-step model (Recruit, Match, Onboard, Meet, Evaluate) is replicable at low cost but requires explicit success thresholds to guide scaling decisions.

## References

1. General Medical Council. *The state of medical education and practice in the UK*. 2025 [Accessed 14th April 2026]. <https://www.gmc-uk.org/about/what-we-do-and-why/data-and-research/the-state-of-medical-education-and-practice-in-the-uk> [Accessed 14th April 2026].
2. General Medical Council. *GMC to launch new survey of tens of thousands of UK doctors*. <https://www.gmc-uk.org/news/news-archive/gmc-to-launch-new-survey-of-tens-of-thousands-of-uk-doctors> [Accessed 28th May 2026].
3. Withnall R. *MRCGP annual report covering 2024/25*. 2025 [Accessed 28th May 2026]. <https://www.rcgp.org.uk/mrcgp-exams> [Accessed 28th May 2026].
4. Tzortziou Brown V, Haviland J, Priyadarshini G, Turner M, George RE, Siriwardena AN, et al. Language of primary medical qualification and differential MRCGP exam attainment: an observational study. *British Journal of General Practice*. 2025;75(753): e285–e291. <https://doi.org/10.3399/BJGP.2024.0296>.
5. Ellis R, Brennan PA, Lee AJ, Scrimgeour DS, Cleland J. Differential attainment at MRCS according to gender, ethnicity, age and socioeconomic factors: a retrospective cohort study. *Journal of the Royal Society of Medicine*. 2022;115(7): 257–272. <https://doi.org/10.1177/01410768221079018>.
6. Tridente A, Parry-Jones J, Chandrashekaraiyah S, Bryden D. Differential attainment and recruitment to Intensive Care Medicine Training in the UK, 2018–2020. *BMC Medical Education*. 2022;22(1): 672. <https://doi.org/10.1186/s12909-022-03732-w>.

7. Woolf K, Rich A, Viney R, Needleman S, Griffin A. Perceived causes of differential attainment in UK postgraduate medical training: a national qualitative study. *BMJ Open*. 2016;6(11): e013429. <https://doi.org/10.1136/bmjopen-2016-013429>.
8. Atewologun D, Kline R. *Fair to refer? Reducing disproportionality in fitness to practise concerns reported to the GMC*. 2019 [Accessed 28th May 2026]. <https://www.gmc-uk.org/about/what-we-do-and-why/data-and-research/research-and-insight-archive/fair-to-refer> [Accessed 28th May 2026].
9. NHS England. *Welcoming and valuing international medical graduates: a guide to induction for IMGs recruited to the NHS*. 2022 [Accessed 28th May 2026]. <https://www.england.nhs.uk/long-read/welcoming-and-valuing-international-medical-graduates/> [Accessed 28th May 2026].
10. General Medical Council. *Tackling disadvantage in medical education Analysis of postgraduate outcomes by ethnicity and the interplay with other personal characteristics*. 2023 [Accessed 28th May 2026]. [https://www.gmc-uk.org/-/media/documents/96887270\\_tackling-disadvantage-in-medical-education-020323.pdf](https://www.gmc-uk.org/-/media/documents/96887270_tackling-disadvantage-in-medical-education-020323.pdf) [Accessed 28th May 2026].
11. NHS England. NHS equality, diversity, and inclusion improvement plan. 2023 [Accessed 28th May 2026]. <https://www.england.nhs.uk/long-read/nhs-equality-diversity-and-inclusion-improvement-plan/> [Accessed 28th May 2026].
12. Baryeh K, Tarannum ST, Cotzias C. Welcome to the NHS—Evaluating an International Medical Graduate Induction Programme. *British Journal of Hospital Medicine*. 2025;86(4): 1–14. <https://doi.org/10.12968/hmed.2024.0621>.

13. Sheokand A, Khanna R, Estrada D, Dudha S, Rathee R. Introduction of a Dedicated Induction Programme for International Medical Graduates in Psychiatry Training at South London and Maudsley National Health Service (NHS) Foundation Trust and Oxleas NHS Foundation Trust. *BJPsych Open*. 2025;11(S1): S105–S105. <https://doi.org/10.1192/bjo.2025.10308>.
14. Moghrabi S, Roy S, Matras B. Evaluation of the Extended Induction Programme for International Medical Graduate Core Psychiatry Trainees in CNWL. *BJPsych Open*. 2024;10(S1): S113–S113. <https://doi.org/10.1192/bjo.2024.315>.
15. Kirkpatrick DL, Kirkpatrick JD. *Evaluating training programs: the four levels*. 3rd ed. San Francisco: Berrett-Koehler Publishers; 2006.
16. Sambunjak D, Straus SE, Marušić A. Mentoring in Academic Medicine. *JAMA*. 2006;296(9): 1103. <https://doi.org/10.1001/jama.296.9.1103>.
17. Kram KE. *Mentoring at Work: Developmental Relationships in Organizational Life*. Glenview, IL: Scott, Foresman and Company; 1985.
18. Teo MYK, Ibrahim H, Lin CKR, Hamid NABA, Govindasamy R, Somasundaram N, et al. Mentoring as a complex adaptive system – a systematic scoping review of prevailing mentoring theories in medical education. *BMC Medical Education*. 2024;24(1): 726. <https://doi.org/10.1186/s12909-024-05707-5>.
19. Saxena SG, Tisdell E, Farace E, Godfrey T, Aumiller B, Dell E, et al. Achieving equity for International Medical Graduates: a systematic review. *Frontiers in Medicine*. 2025;12. <https://doi.org/10.3389/fmed.2025.1601492>.

20. McClintock AH, Kim S, Chung EK. Bridging the Gap Between Educator and Learner: The Role of Psychological Safety in Medical Education. *Pediatrics*. 2022;149(1).  
<https://doi.org/10.1542/peds.2021-055028>.
21. Moore DE, Green JS, Gallis HA. Achieving desired results and improved outcomes: Integrating planning and assessment throughout learning activities. *Journal of Continuing Education in the Health Professions*. 2009;29(1): 1–15. <https://doi.org/10.1002/chp.20001>.
22. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *American Journal of Public Health*. 1999;89(9): 1322–1327. <https://doi.org/10.2105/AJPH.89.9.1322>.
23. Health Research Authority. *UK Policy Framework for Health and Social Care Research*. 2022 [Accessed 8th June 2026]. <https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/uk-policy-framework-health-social-care-research/> [Accessed 8th June 2026].