Royal Australasian College of Surgeons

FATES 1 Impact of the quantity of supervisor oversight on specialty training outcomes

A scoping review

18 July 2025

CONTRIBUTORS

Dr Alun Cameron, ASERNIP-S, RACS

Dr Bridget Clancy FRACS, RSS, RHESC, RACS

CONTENTS

Abbreviations	2
Introduction	3
Methods	4
Results	5
Discussion	6
Conclusions	10
References	11
Appendix A: Search strategy	18
Appendix B: PRISMA chart	20
Appendix C: Summary of included studies	21

LIST OF TABLES

Table 1	SPICE (setting, perspective, intervention, comparison, evaluation) table	4
Table 2	Ovid MEDLINE search results (inception to May 28, 2025)	18
Table 3	Ovid Embase search results (1974 to 2025 May 28)	18
Table 4	Identified studies: study design	21
Table 5	Identified studies: study reporting and relevance to the research question	23

ABBREVIATIONS

FATES	Flexible Approach to Training in Expanded Settings
FRACS	Fellow of the Royal Australasian College of Surgeons
FTE	full-time equivalent
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RACS	Royal Australasian College of Surgeons
RHESC	Rural Health Equity Steering Committee
RRR	regional, rural and remote
RSS	Rural Surgery Section
SPICE	setting, perspective, intervention, comparison, evaluation

ACKNOWLEDGEMENTS

This report was undertaken as part of the Australian Government Department of Health, Disability and Ageing Flexible Approach to Training in Expanded Settings (FATES) 1 funded project titled Rural accreditation—addressing barriers to rural specialist training.

The project was conducted by the Royal Australasian College of Surgeons (RACS), alongside consortium partner RACMA (Royal Australasian College of Medical Administrators).

INTRODUCTION

Quality supervision is required for trainees to achieve the learning and graduate outcomes of specialty medical training. Supervision can be direct (direct observation), indirect (within the same facility or able to attend within a specified time) or remote (not at the same facility, available by phone or video, not available to attend).[1]

The quality of supervision can be measured in several ways. A common proxy is the quantity of supervision. Quantitative measures include: the number of supervisors, the number of hours of supervision, the number of patient encounters or procedures directly supervised, or a combination of one or more of these. The evidence supporting the ideal quantity of supervision (including number of supervisors and number of hours of supervision) for the best training outcomes is uncertain.

The Royal Australasian College of Surgeons (RACS) training committee/board standards for accreditation of hospital training posts include a requirement for a minimum number of FRACS (Fellow of the Royal Australasian College of Surgeons) surgeons per trainee; 7 out of 9 Specialty Training Boards have a quantified supervisor requirement (2 or 3 FRACS to supervise each trainee) in their training-post accreditation standards. Supervisors must be available to attend on-site when needed; that is, they must provide direct and indirect rather than remote supervision. The standards do not include a minimum requirement for number of hours of direct and indirect supervision per trainee per week.

In the context of rural surgical specialty training, regional, rural and remote (RRR) hospital surgical units tend to have fewer FRACS surgeons with higher per-surgeon FTE (full-time equivalent) compared to urban surgical units, with more FRACS surgeons and lower per-surgeon FTE. As a result, many RRR hospitals are unable to reach and maintain the requirements of training-post accreditation of specialist medical colleges and specialty societies.[2, 3]

Current Australian Medical Council accreditation standards are flexible and allow for remote and alternative models of supervision.[4, 5] Remote supervision cannot replace on-site supervision for surgical training but presents a solution to expanding the number of supervisors available to a trainee. The RACS Rural Health Equity Strategic Action Plan recommends that supervision be based on total FTE hours of direct consultant supervision rather than on the total number of FRACS employed, to preserve quality and remove the barrier to accreditation of the rural post.[6, 7] An example is, rather than having 3 surgeons providing a total of 4 sessions of on-site supervision per week, there could be 2 surgeons providing a total of 8 sessions.[8]

Project aim and research question

This literature review is one part of the Australian Government Department of Health, Disability and Ageing's Flexible Approach to Training in Expanded Settings (FATES)-funded project: *Rural accreditation—addressing barriers to rural specialist training*.

The objective of this scoping review is to identify and describe evidence that supports quantifiable supervisor requirements for specialty training (e.g. number of individual supervisors, number of hours or number of sessions of supervision per week) and how this relates to training effectiveness.

Research question:

What is the evidence supporting quantity of supervisor oversight on specialty training outcomes?

METHODS

A scoping review was conducted to map available evidence on quantity of supervision on training outcomes and to identify extant research gaps. A scoping review is particularly useful for collating heterogeneous literature and was considered the most appropriate approach given that the quantified impact of supervision on trainee outcome was anticipated to have a limited and varied evidence base. Where possible, the scoping review was performed according to standard methods and adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Extension for Scoping Reviews reporting guideline.[9-11] The project proposal was approved on 26 March 2025.

The project was formulated using the SPICE (setting, perspective, intervention, comparison, evaluation) framework, relevant to evaluating the outcomes of a service, project or intervention (Table 1**Error! Reference source not found.**).[12]

Table 1 SPICE (setting, perspective, intervention, comparison, evaluation) table

SPICE element	Description
Setting	Hospital-based training programs
Perspective	Medical specialist • Surgery • Procedural specialties • Non-procedural specialties General practice will be excluded
Intervention/exposure/interest	Supervision • Quantified by number of supervisors, FTE of supervisors, volume of supervision sessions, duration of supervision, number of patient encounters or procedures etc
Comparison	• Alternative supervision models • Non-comparative studies will also be included
Evaluation	Trainee outcome (e.g. successful training, progression to next training position, underperformance, failure to progress, patient-relevant outcomes)
Limits	English language

FTE = full-time equivalent; SPICE = setting, perspective, intervention, comparison, evaluation

Search strategy

Systematic literature searches were undertaken across Ovid MEDLINE and Embase (Appendix A Table 2, Table 3).

Study selection and data extraction

Titles and abstracts were screened for relevance according to the predefined inclusion criteria (Table 1) by one reviewer. Full-text reports of potentially relevant studies were selected for further inspection. A standardised data extraction template in Excel was used to extract data from the individual studies.

At full-text analysis, relevant studies were sorted based on 2 criteria:

- A quantitative or measurable aspect of the training was reported (with a focus on quality or quantity of supervision).
- The quality of the training outcome was determined (with a focus on objective measures).

Evidence analysis

A narrative synthesis was undertaken. Quality assessment was not planned for this scoping review.

RESULTS

Across all identified studies, no study reported across both metrics of quantitated aspect of training delivery and quantitated impact of training. However, 14 studies were included that provide evidence across one and partially the second metric, or partially across both metrics, and are briefly summarised below.[13-26]

A PRISMA chart of study selection is shown in Appendix B. Studies were a mix of reviews of individual training sites, or broader reviews of multiple or national programs. A summary of study design is shown in Table 4, and of study reporting in Table 5 (see Appendix C). The studies do not quantify how the training is provided (e.g. the number of supervisors or quantity of supervision), therefore it is not possible to link the quality of the training with any single aspect of training provision.

Studies were from the US (7) and UK (5), with additional studies from Pakistan (1) and Brazil (1). Reported specialties were a mix of specialties (6), surgery (2), paediatrics (3), neurology, anaesthesia and accident and emergency (1 each). In most studies, the population was trainees (trainees, residents or senior house officers (12), with 2 studies reporting information from trainees and trainers.[14, 19] The majority of studies were single arm (10), with 4 studies reporting comparative information of historical data or across groups.[14, 25, 26]

Osborne and colleagues compared the exam results of all paediatric residents at one university training site in the US across 3 different placement sites: traditional, university-based clinics; private practice offices; and publicly funded community-based clinics.[26] The experiences of the residents varied across the settings, with some settings offering greater patient numbers, different presentations and varied supervision time. There was no change in Paediatric In-Training Examination results, but residents from private practice offices had a more varied patient mix and scored higher on the Behavioural Paediatrics Examination (P value not provided).

Three US studies reported mixed data, including quantitative evidence, investigating individual procedures, case numbers, level of supervision, and changes over time.[21, 23, 24] Together, these studies provide an insight into trainee autonomy and need for supervision. Trainees were observed to require less supervision over time and as their experience grew,[24] and as their volume of cases and supervised procedures increased.[21] In an analysis of the VASQIP (Veterans Affairs Surgical Quality Improvement Program) dataset, outcomes improved over time in cases where trainees were afforded autonomy, even though there was a decrease in overall trainee autonomy during training across this time period.[23]

Ten studies reported qualitative data and report evidence from surveys including perception and satisfaction, representing examples across various individual sites or national training programs.[13-20, 22, 25] While these studies provide an insight into aspects of specialty training and supervision, they were not directly related to an objective measure of training quality. These examples identify that there is variability in supervision and in training between units,[13, 16, 19, 22] and that the training requirements vary dependent on specific procedures and so needs to be tailored to individual requirements.[15] Supervision is valued by trainees, with dedicated time needed to allow for this to happen.[18] Supervision quality and frequency was linked to improved learning[17] and conversely, the lack of supervision (i.e. direct oversight) was linked to medical errors.[13] Also perceived to be related to mistakes were excessive work hours, lack of teamwork, inadequate supervision and problems with patient hand-over.[20, 25] Improvements were reported in trainee satisfaction when the format of the training program, including supervision, was enhanced.[14]

DISCUSSION

The scoping review identified a small number of studies, commonly examples of local specialty training programs with limited objective outcomes reporting. A small number of comparative studies were available, which provide insights into training programs over time and across different groups.[14, 21, 25, 26] None of the studies provide direct evidence on the impact of any single aspect of specialty training or supervision on the training program. Due to the variability in study design and reporting, no synthesis was possible. However, the studies provide an understanding of perceptions and outcomes of various training programs.

Postgraduate specialty training is a complex activity. Many factors can affect the performance of a surgical trainee in Australia, including personal issues and personality of the trainee and the broader training and working environment.[27] Multiple themes impact the experiences and wellbeing of the Australian surgical trainee in the workplace, including relationships with colleagues, work hours and workload, the learning environment and supervision, training program requirements, hospital and team administration, and hospital facilities and programs.[28]

There is an ongoing move towards competency-based (rather than time-based) assessment with a need for structured supervision, assessment and feedback set around tangible goals.[29] This includes the adoption of clinically relevant entrustable professional activities into curricula, with a focus on individualised supervision and high quality feedback.[30] Competency-based training is becoming more common for training in surgery and other specialties, but Australian programs can be variable and can be onerous for trainees and trainers for the assessment of professional skills.[31, 32]

Specialty training programs

Various aspects of training have an impact on the effectiveness of training programs. These include supervision, allowance of time for procedures and formal training, appropriate clinical workloads, the assessment of trainee competence as well as professionalism and the use of structured curricula.[33-37] However, studies show that training programs are rarely standardised in their approach to training. Examples from a number of hospitals and training institutions show variability of training across different sites, with training quality a result of a complex interplay between the trainee, trainer and training environment.[15, 38-40] While studies show that trainees are commonly satisfied with their training, there are often areas for improvement or differences in various indicators of training quality.[41-44] For instance, studies from the UK, the US and the Netherlands show that a lack of time for teaching has been found to be associated with poor quality supervision.[45-47] However, the size of the training site is not necessarily an indication of the quality of the training. In an evaluation of anaesthetic training sessions in UK hospitals, trainees in smaller hospitals received the same number of supervised sessions as in larger teaching hospitals.[48]

Internationally, there have been numerous changes to specialty training over recent years, including the requirement of reduced working hours in the US and in Europe.[49-51] These have led to a range of developments in the design of specialty training programs including a focus on optimising training time and on relevant competencies, and using available training tools such as simulation.[52-55] In these examples, a system-wide approach is used to address and improve the training experience.

While there are no formal working-hour restrictions in Australia, trainees in New Zealand have a maximum of 72 hours of rostered work per week, and RACS suggests an optimum of 50 to 60 hours per week.[56-58] However, there is a continual need to review the adequacy of medical education and make sure that the necessary time and resources are available to support the training experience.[59, 60]

In the delivery of specialty training based on competency-based assessments, a range of methods are available to assess trainees' performance in practical surgical skills; for example, for general surgery these include Objective Structured Assessment of Technical Skill (OSATS), Procedure Based Assessment (PBA) and Entrustability Scales.[61] These tools are valuable. However, a recent scoping review has showed that further research is needed to ensure adequate reliability, as well as clarification

from training programs to ensure that they are used in the most effective way: specifically, whether these assessments are an assessment *for* learning or an assessment *of* learning.[62]

Digital tools and other methods continue to be developed to enhance teaching and learning, and to provide improved and increasingly objective feedback for competency-based education.[63-65]

Supervision

Many attributes are associated with a good supervisor, including 'character' (approachability, patience, enthusiasm, encouraging/supportiveness), 'procedural' (willingness to let trainee operate, balance between supervision and independence), 'teamwork and communication' (sets educational aims and objectives, ability to use appropriate feedback, communication skills, and time availability to train) and 'clinical' (capable, good relationships with patients and the healthcare team).[66, 67] Supervisor training is essential.[68]

Beyond the supervisor, the quality of the supervisor–trainee dyad relationship is critical; this includes asking for advice and having the ability to discuss the training and the delivery of quality care.[69] The trainee and supervisor need to develop a relationship and develop trust.[70] A tool (the 7-point Connection Index) has been developed to quantify this interaction across various domains including empathy, psychological safety, educational alliance and feedback, with increased connection found to be correlated with positive trainee experience.[71] Interestingly, in one example from the US, perceptions of overall departmental supervision were less than overall averages of perceptions of individual supervisors, a concept that should be considered when interpreting unit or national data.[72]

Feedback is an important aspect of supervision. A multisite Dutch analysis showed that supervisor training in feedback with experienced medical educators led to a significant increase in the quality of feedback after the training.[73] The R2C2 feedback model has been established to improve feedback and plans for improvement for trainees, based on relationship, reaction, content and coaching. This model was found to be effective in guiding a productive, reflective feedback conversation in various specialties including surgery.[74, 75] Progress evaluations, based on the CanMEDS (Canadian Medical Education Directives for Specialists) framework, were found to be a constructive method for directing feedback and adding structure to the learning process.[76]

Tools have been developed to measure the quality of anaesthetists' clinical supervision of trainees; information can be collected on a daily basis to measure the overall quality of supervision of a trainee across a period of time.[77, 78] Studies showed that a reduction in supervision scores was associated with increased errors.[79] However, the quantitative nature of the supervision (i.e. how it was provided) was not described.

Observation and direct oversight

Many studies have investigated the presence or absence of supervision or direct oversight on various aspects of training outcome, including adverse effects. While these studies do not provide an indication of how the supervision is provided and how this relates to any specified training placement, they provide an insight into this aspect of trainee supervision.

A previous systematic review has found that increased oversight and enhanced supervision has been shown to be associated with improved patient- and education-related outcomes, although studies are limited by small sample sizes and a lack of objective measures of clinical supervision.[80] Patient-related outcomes were focused on defined events (e.g. patient interactions, procedures or diagnoses); education-related outcomes were limited to perceptions, satisfaction or performance on standardised testing. One study by Osborne and colleagues investigated differences in supervision from 3 separate locations, and has been included separately (see above).[26] However, recent randomised controlled trial evidence has shown that increased supervision did not significantly reduce the rate of medical errors but did result in trainees reporting a decreased level of autonomy.[81] This suggests that other factors are involved with procedural outcomes.

For all procedures, there is a learning curve for trainees to reach a minimal threshold of competency.[82-85] However, analyses of logbooks show that there remains a variability in case numbers and in teaching

styles, or a lack of access to certain experiences.[86-88] In an analysis of 21 procedures for trauma in a Spanish training program, the expectations of trainees and supervisors did not always align in terms of number of procedures, and an assessment of the logbooks showed that about half of all clinical procedures were carried out fewer times than expected.[89]

Reviews of specific of training programs through survey and feedback, for example for endoscopy in the UK and a pan-European study of endoscopic retrograde cholangiopancreatography, aid in highlighting improvement opportunities,[90, 91] and have shown limitations and variability in training programs.[92-95]

Autonomy and entrustment

An important part of specialty training is the preparation of the trainee for surgical competence and progressive independence in clinical practice. Examples show that increasing directly supervised practice with inexperienced clinicians, leading to less direct supervision for more experienced trainees, will reduce the rate of adverse reactions as in the case of obstetrics.[96] Over time, and with experience, the opportunities for the surgical trainee to act independently increases.[24, 97] In balancing the requirements of entrustment and oversight, the supervisor often needs to weigh the needs of stepping in and stepping away for specific patient care tasks.[98] However, as identified in a recent US survey of program directors in adolescent medicine, there can be a variability in expectations for the levels of supervision deemed necessary for safe and effective practice.[99] Suggested improvements to allow more clinical independence includes allowing more clinical time, the use of surgical simulation to allow for more experience to develop technical skills within a controlled environment, and the use of competency-based education to allow for improved and more frequent feedback.[54]

In a thematic analysis of published evidence, a recent systematic review identified that for surgical trainees, the decision to allow entrusted autonomy was primarily linked to patient safety, itself driven by the interactions between the learner, the learning environment and the preferences and confidence of the supervising surgeon.[100] A Dutch study showed that the entrustment process develops over 3 progressive phases: prior knowledge about the resident, observed general competencies (as well as the supervisor's intuition), and re-evaluated decisions and the trainee's overall growth in competencies.[101] A tool developed to assess intraoperative independence measures interactions across 5 domains: types of questions asked, operative plan, instruction, problem-solving, and leadership by the surgical resident.[102]

Examples from US training programs have shown that despite opportunities for trainee autonomy decreasing over the years, outcomes in cases where they are afforded autonomy are improving over time.[23, 54, 103] For plastic surgery, the majority of trainees felt that their training program prepared them for independent practice, although there were perceived gaps in certain procedures.[104]

Limitations of this research

As a scoping review, this research was by its nature speculative, intended to provide an overview of this complex topic and an understanding of the type of evidence available. As a consequence of the broad approach, the literature searches may not have been entirely comprehensive. However, the evidence base identified is likely to be representative of the type of evidence available on this theme.

The quality of training was not linked to one specific element of the training placement. Studies were varied in terms of their design, and most were reviews of local or national training programs. There were few comparative studies that reported on retrospective analysis of data sets, or comparisons of different placements or changes over time. As a result, there are likely to be many confounding factors associated with these analyses. In addition, most outcomes were from surveys and represent perspectives on aspects of training quality rather than a subjective analysis of the training outcome. There were no Australian studies, therefore applicability of the evidence to Australian specialty training is uncertain.

Many additional themes were identified in terms of aspects of the supervision, direct observation and other aspects of training programs. Further research would be valuable to investigate these additional themes in detail, if needed.

Future research in Australia could include using existing tools to measure the quality of training and supervision locally and, if possible, link that with the overall outcome of the placement. Additionally, a formal comparison of the exam outcomes of trainees in rural compared to metropolitan training posts in Australia would provide further insight into these training activities and outcomes across these settings.

CONCLUSIONS

There was no published evidence on any quantitative aspect of supervision or other single component of postgraduate specialty training. Specifically, there was no evidence for the current practice of requiring defined supervisor-to-trainee ratios. Studies suggest that specialty training programs, training terms and training sites are inherently variable, providing different experiences; where reported, the training outcomes are perceived to be highly satisfactory but commonly with areas for improvement. Tailoring the training to individual trainee requirements and allowing an appropriate level of autonomy is needed.

Specialty training is a complex activity involving a number of different variables. In any defined training program, the impact of any single element on the quality of training is not well understood. Allowing time for teaching, having an appropriate workload and good quality supervision are commonly reported. Particularly with less experienced trainees, more direct oversight can lead to improved outcomes, but the manner in which this is best provided per placement is not clear.

Future research to quantify the quality of training and supervision in Australia would further inform aspects of local training and help identify any points of improvement.

REFERENCES

1. Ahpra. Supervised practice framework, https://www.ahpra.gov.au/Resources/Supervised-practice/Supervised-practice-framework.aspx (2022, accessed 19 Feb 2025).

2. Department of Health and Aged Care. How Accreditation Practices Impact Building a Non-General Practice Rural Specialist Medical Workforce, https://www.health.gov.au/sites/default/files/2023-03/how-accreditation-practices-impact-building-a-non-general-practice-rural-specialist-medical-workforce.pdf (2022, accessed 29 July 2023).

3. RACS. *Rural Accreditation Workshop*. 2022. Royal Australasian College of Surgeons.

4. RACS. Training post accreditation, https://www.surgeons.org/en/Trainees/training-post-accreditation (2023, accessed 3 August 2023).

5. Australian Medical Council. Accreditation standards and procedures, https://www.amc.org.au/accredited-organisations/accreditation-standards-and-procedures/ (2025, accessed 25 Feb 2025).

6. RACS. Rural Health Equity Strategic Action Plan, https://www.surgeons.org/News/News/Rural-Health-Equity-Strategic-Action-Plan (2020, accessed 15 August 2023).

7. RACS. Train for Rural; Rural Health Equity Strategy; April 2021, https://www.surgeons.org/-/media/Project/RACS/surgeons-org/files/interest-groups-sections/Rural-Surgery/2-Train-for-Rural.pdf?rev=2e53ad7ecb3e4864be9dcc69e0556bb1&hash=6EF8D4FD8F5C13332C8C625D920E42C6 (2021, accessed July 31 2023).

8. Clancy B. Royal Australasian College of Surgeons Rural Health Equity Strategic Action Plan: excellence through equity. *ANZ journal of surgery* 2022; 92: 1990-1994. DOI: https://dx.doi.org/10.1111/ans.17954.

9. Peters MD, Godfrey CM, Khalil H, et al. Guidance for conducting systematic scoping reviews. *Int J Evid Based Healthc* 2015; 13: 141-146. DOI: 10.1097/XEB.0000000000000050 .

10. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. *JBI Evid Synth* 2020; 18: 2119-2126. DOI: 10.11124/JBIES-20-00167.

11. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med* 2018; 169: 467-473. 20180904. DOI: 10.7326/M18-0850.

12. Booth A, Noyes J, Flemming K, et al. Formulating questions to explore complex interventions within qualitative evidence synthesis. *BMJ Glob Health* 2019; 4: e001107. 20190125. DOI: 10.1136/bmjgh-2018-001107.

13. Baldwin DC, Jr., Daugherty SR, Ryan PM, et al. Residents' Ratings of Their Clinical Supervision and Their Self-Reported Medical Errors: Analysis of Data From 2009. *J Grad Med Educ* 2018; 10: 235-241. DOI: https://dx.doi.org/10.4300/JGME-D-18-00200.1.

14. Butt AS, Shamim MS, Ali MA, et al. Applying a Mixed-Method Approach to Improve On-the-Job Learning and Job Satisfaction in a Cohort of Interns at a University Hospital. *Cureus* 2021; 13: e15905. DOI: https://dx.doi.org/10.7759/cureus.15905.

15. Dale J, Williams S, Wellesley A, et al. Training and supervision needs and experience: a longitudinal, cross-sectional survey of accident and emergency department senior house officers. *Postgrad Med J* 1999; 75: 86-89. Research Support, Non-U.S. Gov't. DOI: https://dx.doi.org/10.1136/pgmj.75.880.86.

16. Davies JH, Tan K and Jenkins HR. The current status of senior house officer postgraduate education in a single region. *Medical Education* 2000; 34: 367-370. DOI: https://dx.doi.org/10.1046/j.1365-2923.2000.00518.x.

17. de Melo Silva Junior ML, Valenca MM and Sampaio Rocha-Filho PA. Characteristics of residency programs and residents' learning experience in Brazil: a multispecialty, nation-wide study. *Psychol Health Med* 2023; 28: 2329-2340. DOI: https://dx.doi.org/10.1080/13548506.2022.2093928.

18. Fuller GN. Clinical training of specialist registrars in neurology: the registrar's perspective. *J R Coll Physicians Lond* 1999; 33: 56-59.

19. Grant J, Kilminster S, Jolly B, et al. Clinical supervision of SpRs: where does it happen, when does it happen and is it effective? Specialist registrars. *Medical Education* 2003; 37: 140-148. DOI: https://dx.doi.org/10.1046/j.1365-2923.2003.01415.x.

20. Jagsi R, Kitch BT, Weinstein DF, et al. Residents report on adverse events and their causes. *Arch Intern Med* 2005; 165: 2607-2613. Research Support, Non-U.S. Gov't. DOI: https://dx.doi.org/10.1001/archinte.165.22.2607.

21. Jensen RM, Anand A, Kearse LE, et al. Looking Beyond the Numbers: A Comparison of Operative Self-Efficacy, Supervision, and Case Volume in General Surgery Residency. *J Grad Med Educ* 2024; 16: 280-285. Comparative Study. DOI: https://dx.doi.org/10.4300/JGME-D-23-00461.1.

22. Kerr J and Skinner A. Supervision of trainee anaesthetists: a survey of opinions and practices. *Anaesthesia* 2002; 57: 76-81. DOI: https://dx.doi.org/10.1046/j.1365-2044.2002.02359.x.

23. Oliver JB, McFarlane JL, Kunac A, et al. Declining Resident Surgical Autonomy and Improving Surgical Outcomes: Correlation Does Not Equal Causality. *J Surg Educ* 2023; 80: 434-441. DOI: https://dx.doi.org/10.1016/j.jsurg.2022.10.009.

24. Schwartz A, Borman-Shoap E, Carraccio C, et al. Learner Levels of Supervision Across the Continuum of Pediatrics Training. *Acad Med* 2021; 96: S42-S49. Research Support, Non-U.S. Gov't. DOI: https://dx.doi.org/10.1097/ACM.0000000000004095.

25. Sorokin R, Riggio JM, Moleski S, et al. Physicians-in-training attitudes on patient safety: 2003 to 2008. *J Patient Saf* 2011; 7: 133-138. DOI: https://dx.doi.org/10.1097/PTS.0b013e31822a9c5e.

26. Osborn LM, Sargent JR and Williams SD. Effects of time-in-clinic, clinic setting, and faculty supervision on the continuity clinic experience. *Pediatrics* 1993; 91: 1089-1093.

27. Paice E. Identification and management of the underperforming surgical trainee. *ANZ J Surg* 2009; 79: 180-184; discussion 185. DOI: https://dx.doi.org/10.1111/j.1445-2197.2008.04837.x.

28. Balhatchet B, Schutze H and Williams N. 'Hey, can I go home?': a qualitative case study of wellbeing and the work environment in surgical training. *Korean J Med Educ* 2024; 36: 401-414. DOI: https://dx.doi.org/10.3946/kjme.2024.313.

29. Javanmard-Emamghissi H and Tierney GM. Modernizing assessment of surgical trainees in the workplace. *Surgery (United Kingdom)* 2021; 39(12): 785-789. Review. DOI: https://dx.doi.org/10.1016/j.mpsur.2021.10.004.

30. Pinilla S, Bauer W, Breckwoldt J, et al. Introducing entrustable professional activities for postgraduate medical training in Switzerland. *GMS J Med Educ* 2024; 41: Doc60. DOI: https://dx.doi.org/10.3205/zma001715.

31. Ashman B, Cosenza A, Frank J, et al. Competency-Based Training for Orthopedic Surgery: An Evaluation of the Australian Orthopedic Association Training Program 2017-2021. *J Surg Educ* 2025; 82: 103550. 20250524. DOI: 10.1016/j.jsurg.2025.103550.

32. Royal Australasian College of Surgeons. Surgical Competence and Performance, https://www.surgeons.org/-/media/Project/RACS/surgeons-org/files/reports-guidelines-publications/manuals-guidelines/surgical-competence-and-performance-framework_final.pdf (2020, accessed 1 July 2025).

33. Ikkos G. Responding to trainee doctors in difficulty. *Hosp Med* 2000; 61: 348-351. DOI: https://dx.doi.org/10.12968/hosp.2000.61.5.1335.

34. Ortwein H, Blaum WE and Spies CD. Anesthesiology residents' perspective about good teaching--a qualitative needs assessment. *Ger* 2014; 12: Doc05. DOI: https://dx.doi.org/10.3205/000190.

35. Schumacher DJ, Slovin SR, Riebschleger MP, et al. Perspective: beyond counting hours: the importance of supervision, professionalism, transitions of care, and workload in residency training. *Acad Med* 2012; 87: 883-888. DOI: 10.1097/ACM.0b013e318257d57d.

36. van der Leeuw RM, Lombarts KM, Arah OA, et al. A systematic review of the effects of residency training on patient outcomes. *BMC Med* 2012; 10: 65. DOI: https://dx.doi.org/10.1186/1741-7015-10-65.

37. Strand P, Edgren G, Borna P, et al. Conceptions of how a learning or teaching curriculum, workplace culture and agency of individuals shape medical student learning and supervisory practices in the clinical workplace. *Adv Health Sci Educ Theory Pract* 2015; 20: 531-557. DOI: https://dx.doi.org/10.1007/s10459-014-9546-0.

38. Iyizoba-Ebozue Z, Fatimilehin A, Kayani M, et al. Unveiling Disparities: Exploring Differential Attainment in Postgraduate Training Within Clinical Oncology. *Clin Oncol (R Coll Radiol)* 2024; 36: e119-e127. DOI: https://dx.doi.org/10.1016/j.clon.2024.03.021.

39. Tessmann T, Marty AP, Stricker D, et al. There is no 'too small' for frequent workplace-based assessment: Differences between large and small residency programs in anesthesia when using a mobile application to assess EPAs. *GMS J Med Educ* 2024; 41: Doc54. DOI: https://dx.doi.org/10.3205/zma001709.

40. Yamamoto S, Tanaka P, Madsen MV, et al. Comparing Anesthesiology Residency Training Structure and Requirements in Seven Different Countries on Three Continents. *Cureus* 2017; 9: e1060. DOI: https://dx.doi.org/10.7759/cureus.1060.

41. Marwan Y, Luo L, Toobaie A, et al. Operating Room Educational Environment in Canada: Perceptions of Surgical Residents. *J Surg Educ* 2021; 78: 60-68. DOI: https://dx.doi.org/10.1016/j.jsurg.2020.07.010.

42. Nestel D, Harlim J, Bryant M, et al. Surgical education and training in an outer metropolitan hospital: a qualitative study of surgical trainers and trainees. *Adv Health Sci Educ Theory Pract* 2017; 22: 639-651. DOI: https://dx.doi.org/10.1007/s10459-016-9697-2.

43. Chan J, Oo S and Yeatman M. Is there a national variation in cardiothoracic trainees' satisfaction? Evaluation of the United Kingdom General Medical Council national training surveys. *J Card Surg* 2019; 34: 1069-1074. Multicenter Study. DOI: https://dx.doi.org/10.1111/jocs.14211.

44. Oker N, Escabasse V, Pensky H, et al. Training satisfaction and work environment in Otorhinolaryngology, Head and Neck surgery: a comparison between France and Germany. *Eur Arch Otorhinolaryngol* 2014; 271: 2565-2573. DOI: https://dx.doi.org/10.1007/s00405-014-3046-y.

45. Busari JO and Koot BG. Quality of clinical supervision as perceived by attending doctors in university and district teaching hospitals. *Medical Education* 2007; 41: 957-964. Comparative Study. DOI: https://dx.doi.org/10.1111/j.1365-2923.2007.02837.x.

46. Hexom B, Trueger NS, Levene R, et al. The educational value of emergency department teaching: it is about time. *Intern* 2017; 12: 207-212. DOI: https://dx.doi.org/10.1007/s11739-016-1447-1.

47. Leatherby RJ, Chandrasekar B, Kausar A, et al. Is there enough time to train? An audit of core surgical training rotas. *Surg* 2022; 20: 268-274. DOI: https://dx.doi.org/10.1016/j.surge.2021.04.009.

48. Green A, Tatham KC, Yentis SM, et al. An analysis of the delivery of anaesthetic training sessions in the United Kingdom. *Anaesthesia* 2017; 72: 1327-1333. DOI: https://dx.doi.org/10.1111/anae.13950.

49. Kirton OC. The Accreditation Council for Graduate Medical Education duty hour regulations: how do we make the best of an unpopular situation in training the surgeons of tomorrow? *JAMA Surg* 2013; 148: 433-434. DOI: 10.1001/jamasurg.2013.372.

50. Giles JA. Surgical training and the European Working Time Directive: The role of informal workplace learning. *Int J Surg* 2010; 8: 179-180. 20100206. DOI: 10.1016/j.ijsu.2010.01.011.

51. Datta ST and Davies SJ. Training for the future NHS: training junior doctors in the United Kingdom within the 48-hour European working time directive. *BMC Med Educ* 2014; 14 Suppl 1: S12. Review. DOI: https://dx.doi.org/10.1186/1472-6920-14-S1-S12.

52. Allum W and King M. Improving surgical training. *Surgery (United Kingdom)* 2023; 41(8): 464-467. Review. DOI: https://dx.doi.org/10.1016/j.mpsur.2023.05.012.

53. DiSegna ST, Kelley TD, DeMarco DM, et al. Effects of Accreditation Council for Graduate Medical Education Duty Hour Regulations on Clinical Preparedness of First-Year Orthopaedic Attendings: A Survey of Senior Orthopaedic Surgeons. *J Surg Orthop Adv* 2018; 27: 42-46.

54. Dougherty PJ, Cannada LK, Murray P, et al. Progressive Autonomy in the Era of Increased Supervision: AOA Critical Issues. *J Bone Joint Surg Am* 2018; 100: e122. DOI: https://dx.doi.org/10.2106/JBJS.17.01515.

55. Sofka CM. Developments and innovations in resident and fellowship education: review article. *Hss J* 2014; 10: 225-229. Review. DOI: https://dx.doi.org/10.1007/s11420-014-9396-6.

56. Gough IR. The impact of reduced working hours on surgical training in Australia and New Zealand. *Surgeon* 2011; 9 Suppl 1: S8-9. 20110212. DOI: 10.1016/j.surge.2010.11.010.

57. O'Grady G, Harper S, Loveday B, et al. Appropriate working hours for surgical training according to Australasian trainees. *ANZ J Surg* 2012; 82: 225-229. 20120117. DOI: 10.1111/j.1445-2197.2011.05992.x.

58. Royal Australasian College of Surgeons. Standards for Safe Working Hours and Conditions Guide, https://www.surgeons.org/-/media/Project/RACS/surgeons-org/files/position-papers/2019-04-16_pos_fes-fel-085_standards_for_safe_working_hours_and_conditions_guide.pdf?rev=f8c38ea775e94391960c9f1e49ea13e5&hash=E4312B07857FF748535CEF0EA6619793 (2018, accessed 1 July 2025).

59. Chong A, Weiland TJ, Mackinlay C, et al. The capacity of Australian ED to absorb the projected increase in intern numbers. *Emerg Med Australas* 2010; 22: 100-107. DOI: https://dx.doi.org/10.1111/j.1742-6723.2010.01268.x.

60. Jelinek GA, Weiland T and Mackinlay C. The Emergency Medicine Capacity Assessment Study: perceived resource requirements to support a major increase in intern numbers in Australian emergency departments. *Emerg Med Australas* 2011; 23: 76-83. Research Support, Non-U.S. Gov't. DOI: https://dx.doi.org/10.1111/j.1742-6723.2010.01377.x.

61. Hackney L, O'Neill S, O'Donnell M, et al. A scoping review of assessment methods of competence of general surgical trainees. *Surgeon* 2023; 21: 60-69. 20220314. DOI: 10.1016/j.surge.2022.01.009.

62. Mayne A, Wilson L and Kennedy N. The Usefulness of Procedure-Based Assessments in Postgraduate Surgical Training Within the Intercollegiate Surgical Curriculum Programme; A Scoping Review. *J Surg Educ* 2020; 77: 1227-1235. Scoping Review. DOI: https://dx.doi.org/10.1016/j.jsurg.2020.03.005.

63. Baboolal SO and Singaram VS. Implementation and Impact of an Adapted Digital Perioperative Competency-Building Tool to Enhance Teaching, Learning And Feedback in Postgraduate Competency-Based Medical Education. *J Surg Educ* 2024; 81: 722-740. DOI: https://dx.doi.org/10.1016/j.jsurg.2024.01.015.

64. Badreldin AMA, Muehle A, Misic J, et al. Objective method to evaluate the competency of residents in cardiac surgery. *Eur J Cardiothorac Surg* 2021; 59: 1059-1068. DOI: https://dx.doi.org/10.1093/ejcts/ezaa467.

65. Diwersi N, Gass JM, Fischer H, et al. Surgery goes EPA (Entrustable Professional Activity) - how a strikingly easy to use app revolutionizes assessments of clinical skills in surgical training. *BMC Med Educ* 2022; 22: 559. DOI: https://dx.doi.org/10.1186/s12909-022-03622-1.

66. Dean B, Jones L, Garfjeld Roberts P, et al. What is Known About the Attributes of a Successful Surgical Trainer? A Systematic Review. *J Surg Educ* 2017; 74: 843-850. DOI: https://dx.doi.org/10.1016/j.jsurg.2017.01.010.

67. Nisar PJ and Scott HJ. Key attributes of a modern surgical trainer: perspectives from consultants and trainees in the United kingdom. *J Surg Educ* 2011; 68: 202-208. DOI: https://dx.doi.org/10.1016/j.jsurg.2010.12.010.

68. Cottrell D, Kilminster S, Jolly B, et al. What is effective supervision and how does it happen? A critical incident study. *Medical Education* 2002; 36: 1042-1049. Research Support, Non-U.S. Gov't. DOI: https://dx.doi.org/10.1046/j.1365-2923.2002.01327.x.

69. Ebneter AS, Vonlanthen R and Eychmueller S. Quality of care as an individual concept: Proposition of a three-level concept for clinical practice. *Z* 2022; 173: 17-21. DOI: https://dx.doi.org/10.1016/j.zefq.2022.05.003.

70. Castanelli DJ, Weller JM, Molloy E, et al. How Trainees Come to Trust Supervisors in Workplace-Based Assessment: A Grounded Theory Study. *Acad Med* 2022; 97: 704-710. Research Support, Non-U.S. Gov't. DOI: https://dx.doi.org/10.1097/ACM.0000000000004501.

71. Puder D, Dominguez C, Borecky A, et al. Assessing Interpersonal Relationships in Medical Education: the Connection Index. *Acad Psychiatry* 2022; 46: 683-691. DOI: https://dx.doi.org/10.1007/s40596-021-01574-0.

72. Hindman BJ, Dexter F and Smith TC. Anesthesia residents' global (departmental) evaluation of faculty anesthesiologists' supervision can be less than their average evaluations of individual anesthesiologists. *Anesth Analg* 2015; 120: 204-208. Observational Study. DOI: https://dx.doi.org/10.1213/ANE.0000000000000444.

73. Renting N, Jaarsma D, Borleffs JC, et al. Effectiveness of a supervisor training on quality of feedback to internal medicine residents: a controlled longitudinal multicentre study. *BMJ Open* 2023; 13: e076946. Multicenter Study. DOI: https://dx.doi.org/10.1136/bmjopen-2023-076946.

74. Sargeant J, Lockyer JM, Mann K, et al. The R2C2 Model in Residency Education: How Does It Foster Coaching and Promote Feedback Use? *Acad Med* 2018; 93: 1055-1063. Research Support, Non-U.S. Gov't. DOI: https://dx.doi.org/10.1097/ACM.0000000000002131.

75. Sargeant J, Mann K, Manos S, et al. R2C2 in Action: Testing an Evidence-Based Model to Facilitate Feedback and Coaching in Residency. *J Grad Med Educ* 2017; 9: 165-170. DOI: https://dx.doi.org/10.4300/JGME-D-16-00398.1.

76. Westein MPD, Koster AS, Daelmans HEM, et al. How progress evaluations are used in postgraduate education with longitudinal supervisor-trainee relationships: a mixed method study. *Adv Health Sci Educ Theory Pract* 2023; 28: 205-222. DOI: https://dx.doi.org/10.1007/s10459-022-10153-3.

77. de Oliveira Filho GR and Dexter F. Interpretation of the association between frequency of self-reported medical errors and faculty supervision of anesthesiology residents. *Anesth Analg* 2013; 116: 752-753. DOI: https://dx.doi.org/10.1213/ANE.0b013e3182831346.

78. Dexter F, Hindman BJ and Epstein RH. Overall anesthesia department quality of clinical supervision of trainees over a year evaluated using mixed effects models. *J Clin Anesth* 2023; 87: 111114. DOI: https://dx.doi.org/10.1016/j.jclinane.2023.111114.

79. De Oliveira GS, Jr., Rahmani R, Fitzgerald PC, et al. The association between frequency of self-reported medical errors and anesthesia trainee supervision: a survey of United States anesthesiology residents-in-training. *Anesth Analg* 2013; 116: 892-897. DOI: https://dx.doi.org/10.1213/ANE.0b013e318277dd65.

80. Farnan JM, Petty LA, Georgitis E, et al. A systematic review: the effect of clinical supervision on patient and residency education outcomes. *Acad Med* 2012; 87: 428-442. DOI: https://dx.doi.org/10.1097/ACM.0b013e31824822cc.

81. Finn KM, Metlay JP, Chang Y, et al. Effect of Increased Inpatient Attending Physician Supervision on Medical Errors, Patient Safety, and Resident Education: A Randomized Clinical Trial. *JAMA Intern Med* 2018; 178: 952-959. DOI: https://dx.doi.org/10.1001/jamainternmed.2018.1244.

82. Arbelot C, Dexheimer Neto FL, Gao Y, et al. Lung Ultrasound in Emergency and Critically Ill Patients: Number of Supervised Exams to Reach Basic Competence. *Anesthesiology* 2020; 132: 899-907. DOI: https://dx.doi.org/10.1097/ALN.0000000000003096.

83. Auerbach MA, White ML, Bhargava S, et al. Are Graduating Pediatric Residents Prepared to Perform Infant Lumbar Punctures?: A Multi-Institutional Descriptive Study. *Pediatr Emerg Care* 2018; 34: 116-120. Multicenter Study. DOI: https://dx.doi.org/10.1097/PEC.0000000000000914.

84. Garcia-Cano J. 200 supervised procedures: the minimum threshold number for competency in performing endoscopic retrograde cholangiopancreatography. *Surg Endosc* 2007; 21: 1254-1255. DOI: https://dx.doi.org/10.1007/s00464-006-9013-y.

85. Steiger HJ. Training of Cerebrovascular Specialists: The Surgeon's View. *Acta Neurochir Suppl* 2018; 129: 129-134. DOI: https://dx.doi.org/10.1007/978-3-319-73739-3_19.

86. Baugh TP and Franzese CB. Extremes in Otolaryngology Resident Surgical Case Numbers: An Update. *Otolaryngol Head Neck Surg* 2017; 156: 1060-1066. DOI: https://dx.doi.org/10.1177/0194599817696305.

87. Brady AK, Town JA, Robins L, et al. Bronchoscopy Teaching Without a Gold Standard: Attending Pulmonologists' Assessment of Learners, Supervisory Styles, and Variation in Practice. *Chest* 2021; 160: 1799-1807. DOI: https://dx.doi.org/10.1016/j.chest.2021.06.006.

88. Pantelides NM, Highton L, Lamb A, et al. An analysis of the cosmetic surgery experience acquired through UK plastic surgery training. *J Plast Reconstr Aesthet Surg* 2018; 71: 1532-1538. DOI: https://dx.doi.org/10.1016/j.bjps.2018.07.012.

89. Santonja-Medina F, Garcia-Sanz MP, Santonja-Renedo S, et al. Mismatch between student and tutor evaluation of training needs: a study of traumatology rotations. *BMC Res Notes* 2018; 11: 826. DOI: https://dx.doi.org/10.1186/s13104-018-3925-1.

90. Brown F, Weidner A, Wells C, et al. UK endoscopy trainer survey: perspectives on current endoscopy training delivery, experience, barriers and opportunities. *Frontline Gastroenterology* 2024; 15(3): 214-221. DOI: https://dx.doi.org/10.1136/flgastro-2023-102557.

91. Hamesch K, Cahyadi O, Dimitriadis S, et al. Endoscopic retrograde cholangiopancreatography training conditions, results from a pan-European survey: Between vision and reality. *United European Gastroenterol* 2025; 13: 474-487. DOI: https://dx.doi.org/10.1002/ueg2.12684.

92. Mousset RA, De Vos Tot Nederveen Cappel WH, Pierie JPEN, et al. Endoscopy training in the Netherlands: A national survey among gastroenterology residents. *Endoscopy International Open* 2022; 10(12): E1548-E1554. DOI: https://dx.doi.org/10.1055/a-1953-9288.

93. Mousset RA, Diemers A, de Vos Tot Nederveen Cappel WH, et al. Advances and challenges in endoscopy training: A mixed methods study among endoscopy trainers in the Netherlands. *Endoscopy International Open* 2024; 12: E1006-E1014. DOI: https://dx.doi.org/10.1055/a-2370-5812.

94. Ricotta DN, Smith CC, McSparron JI, et al. When Old Habits Train a New Generation: Findings From a National Survey of Internal Medicine Program Directors on Procedural Training. *Am J Med Qual* 2018; 33: 383-390. DOI: https://dx.doi.org/10.1177/1062860617743786.

95. Reddavid R, Allum W, Polom K, et al. Upper GI training of young surgeons: a reality full of hurdles. An international survey. *Updates Surg* 2021; 73: 627-637. DOI: https://dx.doi.org/10.1007/s13304-020-00955-8.

96. Aiken CE, Aiken AR, Park H, et al. Factors associated with adverse clinical outcomes among obstetrics trainees. *Medical Education* 2015; 49: 674-683. DOI: https://dx.doi.org/10.1111/medu.12741.

97. McCaughran PW, Burton H, Mohammadi M, et al. The Benefits of Promoting Junior Trainees in Vascular Surgery. *Cureus* 2023; 15: e50517. DOI: https://dx.doi.org/10.7759/cureus.50517.

98. Hatala R, Ginsburg S, Gauthier S, et al. Supervising the senior medical resident: Entrusting the role, supporting the tasks. *Medical Education* 2022; 56: 1194-1202. DOI: https://dx.doi.org/10.1111/medu.14883.

99. Cifra N, Pitts S, Mink R, et al. Analysis of fellowship program director opinions of entrustable professional activities in adolescent medicine fellowship. *Int J Adolesc Med Health* 2024; 36: 237-242. DOI: https://dx.doi.org/10.1515/ijamh-2023-0154.

100. Keuning MC, Lambert B, Nieboer P, et al. Perceptions and Guiding Strategies to Regulate Entrusted Autonomy of Residents in the Operating Room: A Systematic Literature Review. *J Surg Educ* 2024; 81: 93-105. Systematic Review. DOI: https://dx.doi.org/10.1016/j.jsurg.2023.09.008.

101. Sagasser MH, Fluit C, van Weel C, et al. How Entrustment Is Informed by Holistic Judgments Across Time in a Family Medicine Residency Program: An Ethnographic Nonparticipant Observational Study. *Acad Med* 2017; 92: 792-799. Observational Study. DOI: https://dx.doi.org/10.1097/ACM.0000000000001464.

102. Sandhu G, Nikolian VC, Magas CP, et al. OpTrust: Validity of a Tool Assessing Intraoperative Entrustment Behaviors. *Ann Surg* 2018; 267: 670-676. DOI: https://dx.doi.org/10.1097/SLA.0000000000002235.

103. Sehat AJ, Oliver JB, Yu Y, et al. Increasing volume but declining resident autonomy in laparoscopic inguinal hernia repair: an inverse relationship. *Surg Endosc* 2023; 37: 3119-3126. DOI: https://dx.doi.org/10.1007/s00464-022-09476-4.

104. Braza ME, Adams NS and Ford RD. Perceptions of Preparedness in Plastic Surgery Residency Training. *Plast* 2020; 8: e3163. DOI: https://dx.doi.org/10.1097/GOX.0000000000003163.

APPENDIX A: SEARCH STRATEGY

Table 2 Ovid MEDLINE search results (inception to 28 May 2025)

#	Query	Results
1	residen*.mp.	408,753
2	(intern or interns or internship or internships).mp.	74,467
3	registrar*.mp.	5,454
4	fellowship*.mp.	22,479
5	(training or education).mp.	1,639,527
6	"Internship and Residency"/	65,303
7	residency/	65,303
8	education, medical, graduate/	35,326
9	(post$graduate adj2 (education or placement* or train*)).mp.	10,896
10	(graduate adj2 (education or placement* or train*)).mp.	61,458
11	(international medical graduate$ or img).mp.	2,116
12	1 or 2 or 3 or 4 or 6 or 7 or 8 or 9 or 10 or 11 [training]	480,590
13	surg*.mp. or surgeons/ or exp Specialties, Surgical/	4,055,786
14	ophthalm*.mp.	194,161
15	urolog$.mp.	105,926
16	(obstetrics and gyna$).mp.	8,051
17	(anaesthe* or anesthe*).mp.	528,110
18	physician$.mp.	693,283
19	cardiolog*.mp.	90,663
20	(paediatri* or pediatri*).mp.	587,576
21	dermatolog*.mp.	123,841
22	gastroenterolog*.mp.	39,834
23	oncolog*.mp.	267,777
24	rheumatolog*.mp.	46,185
25	specialist*.mp.	139,941
26	specialty.mp.	64,015
27	emergency medicine.mp.	34,948
28	generalist.mp.	10,136
29	or/13-28 [specialty]	6,029,528
30	((numbe$ or individual$ or FTE or full time equivalen$ or full-time equivalent$ or minut$ or hour$ or session$ or frequency or criteria or standard or requirment$ or duration or time) adj8 (supervi$ or trainer$)).mp.	12,207
31	12 and 29 and 30 [training and specialty and supervisor metric]	936

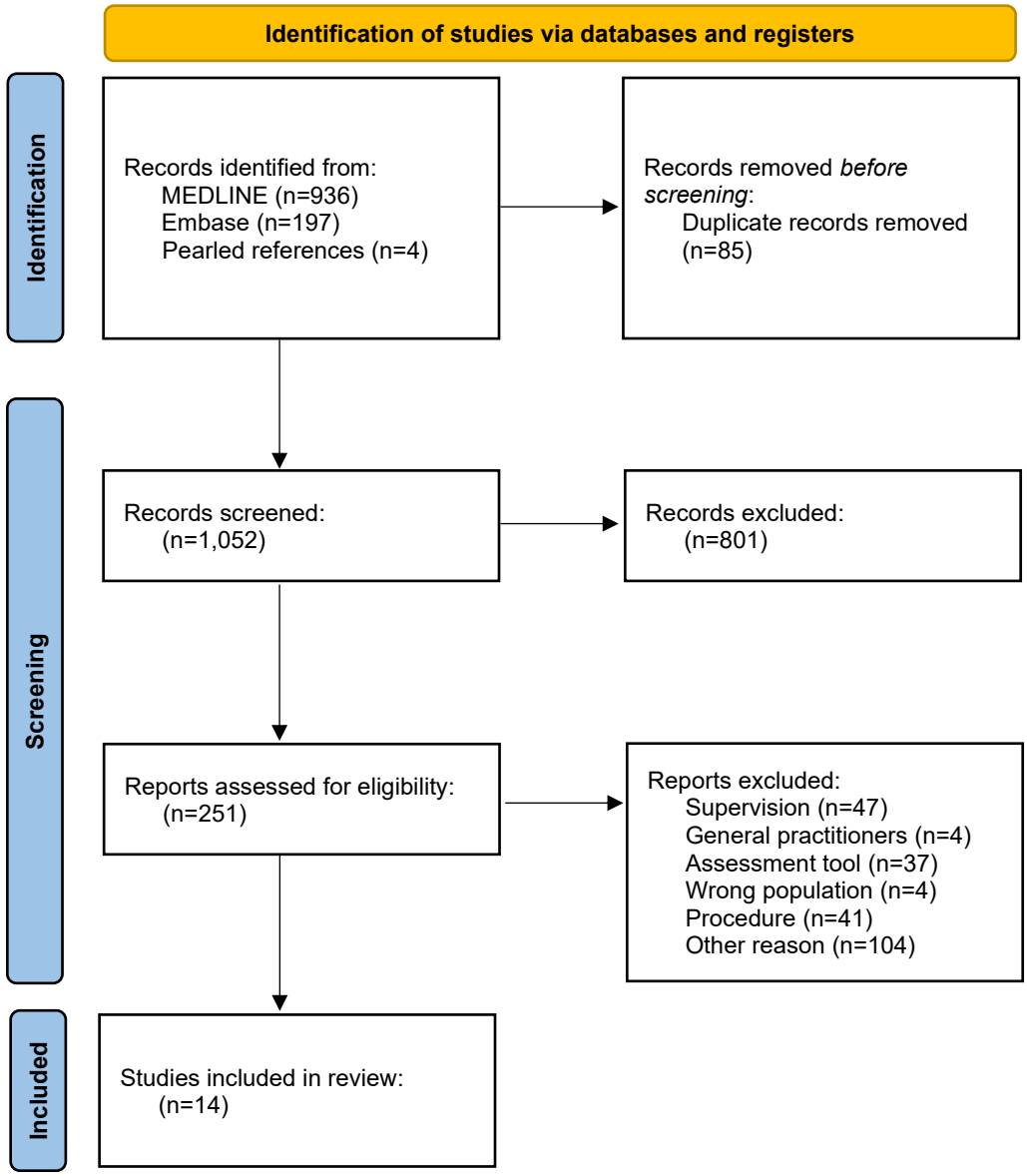
Table 3 Ovid Embase search results (1974 to 28 May 2025)

#	Query	Results
1	residen*.mp.	500,788
2	(intern or interns or internship or internships).mp.	19,122
3	registrar*.mp.	9,461
4	fellowship*.mp.	28,300
5	(training or education).mp.	2,255,297
6	"Internship and Residency"/	272,230
7	residency/	15
8	education, medical, graduate/	272,230
9	(post$graduate adj2 (education or placement* or train*)).mp.	25,932
10	(graduate adj2 (education or placement* or train*)).mp.	18,454

11	(international medical graduate$ or img).mp.	2,988
12	1 or 2 or 3 or 4 or 6 or 7 or 8 or 9 or 10 or 11 [training]	782,199
13	surg*.mp. or surgeons/ or exp Specialties, Surgical/	7,766,391
14	ophthalm*.mp.	264,576
15	urolog$.mp.	159,669
16	(obstetrics and gyna$).mp.	14,338
17	(anaesthe* or anesthe*).mp.	768,581
18	physician$.mp.	926,388
19	cardiolog*.mp.	239,985
20	(paediatri* or pediatri*).mp.	933,609
21	dermatolog*.mp.	189,051
22	gastroenterolog*.mp.	118,488
23	oncolog*.mp.	570,944
24	rheumatolog*.mp.	143,526
25	specialist*.mp.	271,192
26	specialty.mp.	82,261
27	emergency medicine.mp.	68,844
28	generalist.mp.	10,491
29	or/13-28 [specialty]	10,502,363
30	((numbe$ or individual$ or FTE or full time equivalen$ or full-time equivalent$ or minut$ or hour$ or session$ or frequency or criteria or standard or requirment$ or duration or time) adj8 (supervi$ or trainer$)).mp.	19,189
31	12 and 29 and 30 [training and specialty and supervisor metric]	1,735
32	limit 31 to conference abstract status	635
33	31 not 32	1,100
34	limit 33 to "remove medline records"	197

APPENDIX B: PRISMA CHART

Figure 1 **PRISMA diagram showing study selection**



APPENDIX C: SUMMARY OF INCLUDED STUDIES

Table 4 Identified studies: study design

Study ID Country	Specialty	Study design	Population	Population detail	Population number	Study type	Study themes	Study themes detail
Baldwin 2018[13] US	Many Internal medicine Obstetrics and gynaecology Paediatrics Surgery	Qualitative	Trainees	PGY 2	759	Single arm	Supervision	Errors
Butt 2021[14] Pakistan	Many Mixed (various disciplines)	Qualitative	Interns Residents Supervisors	PGY 1	71	Comparative (before-after)	Mixed	Orientation Supervisors Supervision Mentorship
Dale 1999[15] UK	Accident and emergency Pre-specialty	Qualitative	Senior house officers	PGY 1, 2 (foundation, not specialist)	171	Single arm	Mixed	Training needs Supervision Wellbeing
Davies 2000[16] UK	Paediatrics Pre-specialty	Qualitative	Senior house officers	PGY 1, 2 (foundation, not specialist)	88	Single arm	Mixed	Teaching Study time Supervision
de Melo Silva Junior 2023[17] Brazil	Many Clinical Surgical Diagnostic	Qualitative	Residents	Varied Median=PGY 2	1,421	Single arm	Mixed	Supervision Hours Activities
Fuller 1999[18] UK	Neurology Neurology	Qualitative	Trainees	Specialist registrars (PGY 1–5)	61	Single arm	Supervision	Supervision Effectiveness
Grant 2003[19] UK	Mixed Anaesthesia General practice Laboratory science Medicine	Qualitative	Mixed	Medical directors Educational supervisors Specialist registrars	310	Single arm	Supervision	Supervision effectiveness Variation in training

Study ID Country	Specialty	Study design	Population	Population detail	Population number	Study type	Study themes	Study themes detail
	Paediatrics Psychiatry Surgery							
Jagsi 2005[20] US	Mixed All medical specialties including surgery	Qualitative	Residents	PGY 1–6+	689	Single arm	Mixed	Supervision Work hours Fatigue
Jensen 2024[21] US	Surgery General surgery	Mixed	Residents	PGY 5	1,145	Comparative (group)	Mixed	Supervision Self-efficacy Case volume
Kerr 2002[22] UK	Anaesthesia Anaesthesia	Qualitative	Trainees	Senior house officers Specialist registrars Consultants	274 (86 senior house officer, 83 specialist registrar, 105 consultant)	Single arm	Supervision	Supervision (Duration and degree)
Oliver 2023[23] US	Surgery Surgery	Quantitative	Trainees	Residents	1,346,461 cases	Single arm	Mixed	Resident autonomy
Osborne 1993[26] US	Paediatrics Paediatrics	Qualitative	Residents	Residents	24 (private offices) 11 (community clinic) 27 (university clinic)	Comparative	Mixed	Clinic setting Time in clinic Supervision
Schwartz 2021[24] US	Paediatrics Paediatrics	Qualitative	Residents	PGY 1, 2, 3	248	Single arm	Supervision	Supervision Entrustable professional activities
Sorokin 2011[25] US	Mixed Mixed (various disciplines)	Qualitative	Trainees	House staff Medical students	784 (2003=321, 2008=463)	Comparative (historical 2003 compared to 2008)	Mixed	Attitudes on patient safety

PGY = postgraduate year

Table 5 Identified studies: study reporting and relevance to the research question

Study ID Country	Study duration	Outcomes	Summary	Quality of training quantified?	Quality of outcome quantified?
Baldwin 2018[13] US	Long-term 12 months	Adequacy of supervision	Example, local review. There was substantial variability in supervision, although residents were least likely to rate 'lack of supervision'' as a problem. A perception that the lack of supervision is related to medical errors.	Y	P
Butt 2021[14] Pakistan	Long-term 12 months	Assessment of a change in an internship program	Example, local review. A re-design of an internship program. Supervision and rotations were 'enhanced' leading to improved satisfaction (not a formal quality measure): detail not provided.	Y	P
Dale 1999[15] UK	Medium-term 6 months	Perceived need for further training	Example, local review. Satisfaction of training and supervision. Identifies the training needs (for certain procedures) over time for PGY 1, 2	Y	P
Davies 2000[16] UK	Not defined NR	Assessing postgraduate education	Example, local review. A review of aspects of their training and satisfaction. Trainees were satisfied with their training but there was considerable variation between units.	Y	P
de Melo Silva Junior 2023[17] Brazil	Not defined NR	Assessing training programs	Broad review across multiple training programs. Supervision quality and frequency, adequate structure for healthcare, formal assessments and reduced duty hours were linked to improved learning. Determined by trainee perception.	Y	P
Fuller 1999[18] UK	NR NR	Effectiveness of clinical training	Broad review across multiple training programs. Supervision is valued by trainees. The survey provided information on how to improve supervision, including in different clinical scenarios, and that time needs to be made available to allow for supervision and training.	Y	P
Grant 2003[19] UK	NR NR	Perception of supervision experiences and practice	Broad review across multiple training programs. Supervisory practice is highly variable and there are significant differences between educational supervisors and specialist registrar in perceptions of frequency and effectiveness of supervision. None of the supervisory activities, including ensuring patient safety, were rated as receiving significant or full coverage. There is a need for more explicit guidance on supervision requirements and expectations. Perceptions reported on supervision but not related directly to training quality.	Y	P
Jagsi 2005[20] US	NR NR	Adverse events Near misses	Example, local review (2 training sites). Across all adverse events, 24% were attributed to mistakes. The most common reasons for mistakes, as perceived by residents, were excessive work hours (19%), inadequate supervision (20%) and problems with handoffs (15%).	P	Y

Study ID Country	Study duration	Outcomes	Summary	Quality of training quantified?	Quality of outcome quantified?
Jensen 2024[21] US	Long-term 12 months	Self-reported operative self-efficacy data Self-reported operative independence Case numbers	Analysis of 3 national datasets. Data were categorised to high, middle (mid) and low (self-reported) self-efficacy tiers. Case supervision and case volume are highly related to self-efficacy.	P	P
Kerr 2002[22] UK	NR NR	Different degrees of supervision Duration of supervision	Broad review across multiple training programs. The total duration of supervision has increased over time (from 1970s to 1998). However, the degree or duration of supervision is still variable. Training outcome reported as overall satisfaction.	Y	P
Oliver 2023[23] US	Long-term 15 years	Level of supervision 30-day outcome (from the VASQIP database, in 3x 5-year periods)	Analysis of national database. Despite resident autonomy decreasing, outcomes in cases where they are afforded autonomy are improving over time.	P	Y
Osborne 1993[26] US	Long-term 6 years	Performance in exam	Three clinical settings in one training program. Increased time in clinic resulted in a broader exposure to patients. Residents placed in private offices had a more varied patient mix, were more closely supervised and seemed to gain primary care skills more rapidly than residents at other sites. Clinic setting, time in clinic and faculty supervision affect the quality of the clinic experience.	P	Y
Schwartz 2021[24] US	Long-term 3 years (analysis of 3 datasets)	Meeting entrustable professional activities	Broad review across multiple training programs. Within each longitudinal training period, learners were rated as needing less supervision over time in each activity.	P	P
Sorokin 2011[25] US	NR NR	Attitudes to patient safety and whether this had changed with time	Example, local review. There was an improvement with time (a comparison of survey results from 2003 versus 2008). Issues associated with patient safety and reduced adverse events included teamwork, supervision, work hours.	P	Y

NR = not reported; P = partial; PGY = postgraduate year; VASQIP = Veterans Affairs Surgical Quality Improvement Program; Y = yes