ASCILITE 2026

The Future is Connection:

Collaboration, Creativity and Compassion in a Digital Society

Enabling AI-ready medical graduates: A working group approach to curriculum and assessment change in a graduate medical program

Joanne Hart, Phoebe Adams, Sally Ayesa, Adrian Boscolo, Christine Bowles, Kellie Charles, Katryna Dening, Lisa Hampshire, Kevin Hou, Priya Joshi, Paul Roth, Matthew Shu, Dominic Joel Symon, Ian Todd, Alan Wong, Caryl Barnes

University of Sydney

Artificial intelligence (AI) is transforming clinical practice, and digital-health and assessment frameworks increasingly identify capabilities that involve using AI. Translating these expectations into a crowded, vertically integrated medical curriculum is a substantial implementation challenge. This case study describes how one Australian graduate medical program has approached integrating AI literacy education, teaching with AI and assessment reform. A cross-functional working group, convened by the Dean of Medicine and chaired by the first author, comprising academic, clinical, professional, library and student members and reporting to the School Executive, has coordinated the program's early response through monthly meetings, program planning day sessions and engagement with grassroots pilots, including a trial of an AI-powered communication skills simulation platform. We describe the challenges encountered: defining AI literacy content without a settled national curriculum, negotiating curriculum space across all year levels, sustaining communication among diverse stakeholders, developing staff capability across attitudes ranging from enthusiasm to sustained concern, responding to student concerns about AI-enabled learning tools, and redesigning assessment to comply with institution-wide reform while preserving a programmatic assessment system. We share practical lessons for other programs, arguing that an enabling, relationship-centred approach may help coordinate change.

Keywords: artificial intelligence, medical education, curriculum change, assessment reform, faculty development, AI literacy

Introduction

Artificial intelligence (AI) is already influencing clinical practice and medical education. The advent of large language models and generative AI has sharply accelerated the integration of AI into clinical work, from documentation and decision support to patient communication, creating practical questions for education providers (Artsi et al., 2025; Thirunavukarasu et al., 2023). Medical graduates entering the workforce will be expected to use, appraise and supervise these tools safely, making AI literacy a relevant graduate capability (Ng et al., 2021; Tolentino et al., 2024). In Australia, this expectation is reinforced at several levels. The Australian Medical Council and the Australian Digital Health Agency have articulated digital health capabilities expected of medical practitioners, positioning digital and data literacy within professional practice (Australian Medical Council & Australian Digital Health Agency, 2021). The national higher education regulator has called on providers to reform assessment for an era in which generative AI is ubiquitous, emphasising assessment validity, academic integrity and the responsible use of AI by students (Lodge et al., 2023). Universities have responded with institution-wide policies committing to AI-ready graduates and with structural assessment reforms, including two-lane models that distinguish secured, supervised assessment of learning from open contexts in which students may work with AI (Bridgeman & Liu, 2025).

For medical programs, the practical questions are what to teach, where to place it in an already crowded curriculum, and how to bring staff and students along. This paper presents a case study of how one Australian graduate medical program is tackling these questions. The work is ongoing and deliberately adaptive: in a field moving this quickly, we regard flexibility as a design requirement rather than a weakness and expect our approach to keep evolving with the technology, the evidence and the regulatory environment. We describe the drivers for change, the working group model adopted to lead it, the challenges being encountered in

defining content, negotiating curriculum space, developing staff capability, managing resistance and redesigning assessment, and the lessons we believe are transferable to other programs. Consistent with the conference theme, our central argument is that AI integration is as much a human and organisational challenge as a technical one, requiring sustained connection and communication across academics, clinicians, professional staff and students.

Context

The setting is a four-year graduate-entry Doctor of Medicine (MD) program at a large Australian university. The program is vertically integrated, delivered across multiple metropolitan and rural clinical schools, and underpinned by a programmatic assessment system in which learner progress is judged on accumulated evidence across many low-stakes assessment points rather than single high-stakes examinations (van der Vleuten et al., 2012). The curriculum is tightly scheduled: foundational biomedical sciences, clinical skills, professional development and research training compete for limited timetable space, and new content requires explicit trade-offs. Curriculum governance sits within the MD Program itself, operating under Faculty-level curriculum governance and University-level policy and quality processes, with change typically progressing through annual planning cycles. These features, common to many medical programs, mean that responding to a fast-moving external driver such as generative AI stretches the usual machinery of curriculum change.

This paper is a descriptive case study of a single program, addressing the question: how can a graduate medical program coordinate curriculum and assessment change in response to generative AI across a large and distributed teaching community? It draws on the working group's terms of reference, agendas and meeting records from June 2025 onwards; materials and participant responses from MD Program planning and quality days; curriculum and assessment mapping artefacts; and feedback relayed to the group through its staff and student members. The authors are participants in the work described, so the account is reflective as well as documentary, and we distinguish throughout between documented outcomes, our own observations as participants, and lessons we regard as still emerging. This work did not require human research ethics review. Evaluation studies arising from it, including the trial described below, run under separate approvals.

Our approach: a cross-functional working group

Rather than assigning AI integration to an individual academic lead or an existing committee, the Dean of Medicine approved a dedicated AI in Medical Education working group which commenced in June 2025, reporting to the School of Medicine Executive. The working group includes academic staff with education and assessment leadership roles, practising clinicians who teach into the program, non-clinician academic educators with expertise in curriculum, pedagogy and assessment, professional staff including educational designers, library staff with expertise in information and AI literacy, and student representatives. This composition was a conscious design decision: each group contributes a different perspective. Clinicians ground decisions in the realities of AI use in practice, non-clinician academic educators contribute expertise in learning theory, curriculum design and evidence-informed educational practice, educational designers translate intent into deliverable learning activities, library colleagues bring established frameworks for literacy education, and students provide direct insight into how AI is being used and perceived by learners. Constituting the group in this way was intended to make it a boundary-crossing space in which perspectives that ordinarily sit in separate parts of the organisation are brought into contact and negotiated (Akkerman & Bakker, 2011). The reporting line to the School of Medicine Executive connects the group's recommendations to formal governance and resourcing decisions.

The group meets monthly to work through needs, content and strategy, providing a standing pattern that keeps momentum between formal curriculum governance cycles. Its remit spans three interconnected streams: defining AI literacy learning outcomes and content for students; supporting the responsible use of AI in teaching and learning; and contributing to assessment redesign in line with institutional reform that is in response to the availability of AI. Beyond meetings, the group has run AI sessions at MD Program planning days, reaching the wider teaching community. Notably, the group does not direct the many grassroots pilots of

AI in teaching and learning underway across the program, but it is increasingly a focal point for them. One such pilot, a trial of an AI-powered virtual patient communication platform offering students additional low-stakes conversational practice alongside existing simulated patient teaching, came to the group's attention through student feedback, as discussed below.

Challenges

Defining the content. To date there is no settled national AI curriculum for medical education, so the working group has had to define what AI literacy means for our graduates. We have anchored this work to the digital health capabilities expected of practitioners (Australian Medical Council & Australian Digital Health Agency, 2021) and to broader conceptions of AI literacy encompassing knowing and understanding AI, using and applying it, and evaluating it critically and ethically (Ng et al., 2021). Even with these anchors, decisions about depth and emphasis; for example, how much technical understanding of machine learning a graduate requires versus critical appraisal of AI outputs, remain contested and require ongoing negotiation.

Negotiating curriculum space. The MD curriculum is full. Inserting AI content at every year level so that literacy develops progressively from foundational concepts to application in clinical placements requires negotiation with year coordinators, theme leads and discipline heads, each of whom is protective of hard-won curriculum time. Progress has depended on identifying natural integration points within existing teaching, such as evidence-based medicine, clinical reasoning, ethics and professional practice, rather than demanding standalone hours, and on being explicit about trade-offs when new content genuinely requires displacement of old.

Communication across many parties. A change program touching every year level, multiple clinical schools and several professional teams generate a substantial communication burden. Keeping year leads, clinical supervisors, professional staff and students informed, consulted and aligned has been a challenge. The working group's diverse membership helps here, as each member acts as a conduit to their constituency, but we have learned that communication must be planned and resourced rather than assumed.

Faculty development and differing views. Staff views on AI range from enthusiasm to scepticism and sustained concern, reflecting the diversity of attitudes towards generative AI reported among medical educators (Cervantes et al., 2024). Building capability across this spectrum is slow work. Planning day sessions have raised general awareness, but developing the confidence to redesign teaching and assessment around AI requires sustained, hands-on support, consistent with emerging evidence of ongoing faculty development needs in this area (McCoy et al., 2025). More difficult still is scepticism that shades into sustained opposition: a small number of colleagues remain unconvinced of the direction of the work, and unaddressed concerns can slow engagement with new initiatives. We have found it more productive to treat objection as signal rather than noise, engaging sceptical colleagues directly, acknowledging the legitimate concerns that often underlie resistance including workload, academic integrity, and the place of human judgement, and adjusting implementation where those concerns are well founded.

Student perspectives. Students have also raised concerns. The AI communication platform trial came to the working group's attention through student objections, with concerns including the authenticity of practising human communication with a machine, data privacy, and the appropriateness of AI in a domain (i.e. empathic communication) that students regard as distinctly human. Responding has required transparency about the purpose and limits of the tool, emphasising its role as supplementary practice rather than replacement of simulated patient teaching, and providing genuine avenues for student feedback to shape the trial. The episode reinforced the value of student membership of the working group: early feedback creates an opportunity to address concerns before positions become entrenched. It also illustrated the group's emerging role as the place where concerns about AI in the program, from any quarter, can be raised and addressed.

AI in everyday teaching practice. A second set of issues has surfaced as staff begin using AI tools in their own teaching work. Colleagues have reported drafting assessment items with generative AI and finding the output inaccurate or pitched at the wrong level, such that checking it took as long as writing from scratch. The group has also begun scoping AI-assisted curriculum mapping, where the technical promise of automated alignment to learning outcomes is constrained by the quality and consistency of the source material and by the absence of dedicated resourcing. Both illustrate a role we did not anticipate: the group has been as useful for testing claims made about AI tools as for promoting their use.

Assessment redesign. Assessment has been the most consequential stream. Institutional assessment reform, consistent with sector-wide two-lane models (Bridgeman & Liu, 2025) and national guidance on assessment in the age of AI (Lodge et al., 2023), requires programs to be explicit about where assessment is secured and supervised and where students may legitimately work with AI. Our challenge is to comply with this reform while preserving the integrity of a programmatic assessment system built on many distributed, often unsupervised, low-stakes data points (van der Vleuten et al., 2012). Mapping every assessment against the two lanes, deciding which judgements of competence must be made under secured conditions, and redesigning open assessments so that AI use becomes a learning opportunity rather than an integrity risk is a multi-year undertaking that is still in progress.

Lessons and implications

Several lessons from our experience may be transferable to other contexts or programs. First, the working group model has provided a useful forum for a problem that is fast moving, cross-cutting and contested. A standing, cross-functional group with a monthly meeting schedule and a direct reporting line to the School Executive can move faster than annual curriculum cycles while remaining connected to formal governance. Second, the deliberate diversity of membership, academic, clinical, professional, library and student, has broadened the perspectives represented, distributing both expertise and legitimacy. Third, the group's stance has mattered as much as its structure: we have consciously sought to enable colleagues to change rather than to drive change onto them, prioritising relationships, capability building and psychological safety to experiment. Fourth, resistance from staff and students should be expected rather than dismissed. Engaged with respectfully, it surfaces legitimate concerns; compassion for those unsettled by rapid change is a condition of progress, not a courtesy. Finally, external requirements such as accreditation expectations around digital health, regulator guidance on assessment and institutional policy on AI-ready graduates are most useful when framed as leverage for locally owned change rather than as compliance burdens.

Conclusion

Integrating AI into a medical program is generally framed as a technological challenge. Our experience suggests it is better understood as a human one: a sustained exercise in connection, negotiation and trust across a large and varied educational community. A cross-functional working group, convened with authority, meeting regularly and committed to enabling rather than mandating, has helped one graduate medical program coordinate work on AI literacy, teaching innovation and assessment reform while holding onto the relationships and the programmatic assessment philosophy that define the program. The work is far from finished, and evaluation is the priority for the next phase. A mixed-methods evaluation of the virtual patient trial is in development, a study of student and staff attitudes to AI in education and clinical practice is in preparation, and we intend to conduct a staff development needs analysis alongside indicators drawn from the assessment mapping. We expect to be able to report student and staff outcomes after the next full curriculum cycle. Until then we offer the model itself, and the lessons above, as a starting point for programs facing similar challenges.

References

Akkerman, S. F., & Bakker, A. (2011). Boundary crossing and boundary objects. *Review of Educational Research*, *81*(2), 132-169. https://doi.org/10.3102/0034654311404435

Artsi, Y., Sorin, V., Glicksberg, B. S., Korfiatis, P., Nadkarni, G. N., & Klang, E. (2025). Large language models in real-world clinical workflows: A systematic review of applications and implementation. *Frontiers in Digital Health*, *7*, 1659134. https://doi.org/10.3389/fdgth.2025.1659134

Australian Medical Council, & Australian Digital Health Agency. (2021). *The digital health in medicine capability framework*. Australian Medical Council. https://www.amc.org.au/about-the-amc/strategic-projects/digital-capability-framework-for-medicine/

Bridgeman, A., & Liu, D. (2025). The Sydney Assessment Framework. *Teaching@Sydney*. Retrieved 2025/03/18, from https://educational-innovation.sydney.edu.au/teaching@sydney/the-sydney-assessment-framework/

Cervantes, J., Smith, B., Ramadoss, T., D'Amario, V., Shoja, M. M., & Rajput, V. (2024). Decoding medical educators' perceptions on generative artificial intelligence in medical education. *Journal of Investigative Medicine*, *72*(7), 633-639. https://doi.org/10.1177/10815589241257215

Lodge, J., Howard, S., Bearman, M. L., & Dawson, P. (2023). *Assessment reform for the age of artificial intelligence*. https://www.teqsa.gov.au/sites/default/files/2023-09/assessment-reform-age-artificial-intelligence-discussion-paper.pdf

McCoy, L., Ganesan, N., Rajagopalan, V., McKell, D., Niño, D. F., & Swaim, M. C. (2025). A training needs analysis for AI and generative AI in medical education: Perspectives of faculty and students. *Journal of Medical Education and Curricular Development*, *12*, 23821205251339226. https://doi.org/10.1177/23821205251339226

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, *2*, 100041. https://doi.org/10.1016/j.caeai.2021.100041

Thirunavukarasu, A. J., Ting, D. S. J., Elangovan, K., Gutierrez, L., Tan, T. F., & Ting, D. S. W. (2023). Large language models in medicine. *Nature Medicine*, *29*(8), 1930-1940. https://doi.org/10.1038/s41591-023-02448-8

Tolentino, R., Baradaran, A., Gore, G., Pluye, P., & Abbasgholizadeh-Rahimi, S. (2024). Curriculum frameworks and educational programs in AI for medical students, residents, and practicing physicians: Scoping review. *JMIR Medical Education*, *10*(1), e54793. https://doi.org/10.2196/54793

van der Vleuten, C. P., Schuwirth, L., Driessen, E., Dijkstra, J., Tigelaar, D., Baartman, L., & Van Tartwijk, J. (2012). A model for programmatic assessment fit for purpose. *Medical Teacher*, *34*(3), 205-214. https://doi.org/10.3109/0142159X.2012.652239

Hart, J., Adams, P., Ayesa, S., Boscolo, A., Bowles, C., Charles, K., Dening, K., Hampshire, L., Hou, K., Joshi, P., Roth, P., Shu, M., Symon, D., Todd, I., Wong, A., & Barnes, C. (2026). Enabling AI-ready medical graduates: A working group approach to curriculum and assessment change in a graduate medical program. In Cochrane, T., Narayan, V., Poole, T., Lai, P., Lake, P., Moore, O., & Li, L. (Eds.), The Future is Connection: Collaboration, Creativity and Compassion in a Digital Society. Proceedings ASCILITE 2026. Brisbane (pp. 621–625). https://doi.org/10.65106/apubs.2026.3300

Note: All published papers are refereed, having undergone a double-blind peer-review process.
The author(s) assign a Creative Commons by attribution license enabling others to distribute, remix, tweak, and build upon their work, even commercially, as long as credit is given to the author(s) for the original creation.

© Hart, J., Adams, P., Ayesa, S., Boscolo, A., Bowles, C., Charles, K., Dening, K., Hampshire, L., Hou, K., Joshi, P., Roth, P., Shu, M., Symon, D., Todd, I., Wong, A., & Barnes, C. 2026