

# MAKING EVALUATIONS USEFUL FOR HEALTHCARE LEADERSHIP DEVELOPMENT PROGRAMMES

## Author List:

Oscar Lyons<sup>1,2,3,4 \*</sup>

Juliette Catherine Phillipson<sup>4,5\*</sup>

Joao R Galante<sup>6,7</sup>

Timothy Jahn<sup>8</sup>

\*indicates corresponding authors

## Affiliations

<sup>1</sup>Green Templeton College, University of Oxford

<sup>2</sup>New College, University of Oxford

<sup>3</sup>Nuffield Department of Surgical Sciences, University of Oxford

<sup>4</sup>Thrum Leadership Ltd., Oxford

<sup>5</sup>London School of Hygiene and Tropical Medicine, London, UK

<sup>6</sup>Department of Medical Education, Oxford University Hospitals NHS Foundation Trust, Oxford, UK

<sup>7</sup>Department of Oncology, The Royal Marsden NHS Foundation Trust, London, UK

<sup>8</sup>Utah Chapter of Emergency Medicine, Salt Lake City, UT, US

## Correspondence:

Oscar Lyons

[oscar.lyons@gtc.ox.ac.uk](mailto:oscar.lyons@gtc.ox.ac.uk)

Juliette Phillipson

[Ju.phillipson@hotmail.com](mailto:Ju.phillipson@hotmail.com)

**Word Count: 3447**

**Figures: 1**

**Tables: 3**

# ABSTRACT

## BACKGROUND

Effective healthcare leadership has been linked to improved individual and organisational outcomes globally. However, evaluations of healthcare leadership development programmes has often been of low quality. This study investigates the evaluation and decision-making needs of stakeholders for the Oxford Emerging Leaders Programme (ELP) and aims to redesign its evaluation approach.

## METHODS

Drawing from Michael Quinn Patton's utilisation-focused evaluation approach, semi-structured interviews were conducted with 12 key programme stakeholders. Interviews were thematically analysed to identify key areas for useful and impactful evaluation.

## RESULTS

Three main themes were identified: impact on patients, impact on healthcare organisations, and individual outcomes. Individual outcomes were further divided into skills, and qualities. Stakeholders emphasised the importance of measuring improvements in organisational culture, as well as from the perspectives of patients and individual leaders. The need for a multifaceted and longitudinal evaluation approach was highlighted.

## CONCLUSIONS

The study underscores the importance of aligning evaluation methods with stakeholder needs. Tailoring evaluations to specific programme aims and incorporating both qualitative and quantitative measures can enhance their utility. These insights contribute to the broader literature on healthcare leadership development and programme evaluation.

## KEYWORDS

Healthcare, Leadership, Evaluation, Development, Qualitative Research, Oxford Emerging Leaders Programme.

## KEY MESSAGES:

- **What is already known on this topic:** Leadership training is linked to improved outcomes in healthcare, but evaluations of training programmes are often of low quality, with limited alignment to stakeholder needs.
- **What this study adds:** This study highlights the need for a tailored evaluation approach that includes both qualitative and quantitative measures, addressing the specific needs of stakeholders.
- **How this study might affect research, practice, or policy:** The findings advocate for a shift towards more utility-focused evaluations in healthcare leadership development, potentially influencing how future programmes are assessed and improved.

## BACKGROUND

Effective healthcare leadership has been increasingly linked to improved outcomes in healthcare systems worldwide[1–5]. Healthcare leadership development programmes have been shown to improve individual leadership skills and behaviours and in some cases to produce organisational outcomes[6–9].

Many authors have pointed out that evaluation of healthcare leadership programmes has to date been of low-quality. The Joanna Briggs Institute (JBI)[10] and Medical Education Research Study Quality Instrument (MERSQI)[11] tools are designed to assess methodological quality of individual studies. For both of these tools, a study scoring at or above two-thirds of the total possible score is generally considered of good quality. In their 2016 review, Geerts and colleagues found that only 28%(7/28) of published studies of medical leadership programmes met half of the MERSQI standards for high-quality evaluations[8]. In 2021, applying both the MERSQI and JBI tools, Lyons and colleagues found that only 25% of 117 published articles met half of the standards for high-quality evaluations in one of these tools[7]. Meeting only half the criteria would likely indicate average or moderate quality and might indicate that the study has significant methodological weaknesses[7,8,12].

It has been suggested that embedding comprehensive evaluation methods into programmes will allow for more effective understanding of factors which are reliably associated with high level programme outcome[13–15]. Evaluation is useful in both informing the improvement of an individual programme as well as for contributing to the literature as a whole[16].

The Oxford Emerging Leaders Programme (ELP) was developed in 2017 as part of a doctoral research programme[13] in response to a recognised demand for formal leadership development opportunities within a single UK hospital trust. Aligning with best practices in healthcare leadership development[7,8,17,18], the programme was based on a shared leadership model, and experiential learning - emphasising practical skill development through real-world projects. The programme ran three cohorts from 2017 to 2019. Across these cohorts, 54 healthcare professionals participated.

Each cohort featured a structured programme of workshops and project-based learning, with variations in duration and format (Table 1). Participants worked in multidisciplinary teams to develop and execute service improvement projects, which they presented at the final workshop. Monthly workshops incorporated facilitated project work, interactive teaching sessions, and invited speakers who shared insights on navigating leadership challenges within the organisation. Teaching content spanned personal, interpersonal, and organisational leadership domains (Table 2).

	<i>Duration</i>	<i>Format</i>	<i>Gender (n)</i>	<i>Participant roles (n)</i>
<i>Cohort 1</i>	4 months	6 half day workshops	Male (7) Female (5)	Junior (foundation level) doctors (13)
<i>Cohort 2</i>	8 months	8 half day workshops	Male (2) Female (13)	Doctors (9), Physiotherapists (2), Managers (1), Other multidisciplinary healthcare professionals (3)
<i>Cohort 3</i>	6 months	7 full day workshops	Male (13) Female (8)	Doctors (12), Nurses (7), Physiotherapists (1), Managers (1)

Table 1: Programme structure and participant demographics for Cohorts 1, 2, and 3 of the Oxford Emerging Leaders Programme

<i>Team Leadership and Project Management</i>	<i>Service Improvement</i>	<i>Organisational and Strategic Content</i>
<ul style="list-style-type: none"> <li>• Stakeholder analysis and engagement</li> <li>• Leadership styles</li> <li>• Personality and working styles</li> <li>• Conflict resolution and mediation skills</li> <li>• Team dynamics and team leadership</li> <li>• Negotiation, influencing and dialogue skills</li> </ul>	<ul style="list-style-type: none"> <li>• Analysing healthcare systems</li> <li>• Analysing clinical care pathways</li> <li>• Understanding and acting on patient experience</li> <li>• Improving quality and flow in care pathways</li> <li>• Estimating impact</li> </ul>	<ul style="list-style-type: none"> <li>• Trust strategy</li> <li>• Chief executive perspective</li> <li>• Business cases</li> <li>• Healthcare systems</li> <li>• Healthcare and research policy and finances</li> </ul>

Table 2: Programme content areas for Cohorts 1, 2 and 3 of the Oxford Emerging Leaders Programme

Multi-methods evaluation was embedded in the programme. However, it became clear after the completion of the first three cohorts that the evaluation was not in perfect alignment with the evaluation needs of stakeholders and decision-makers. Drawing from principles of Michael Quinn Patton’s utilisation-focused evaluation approach[15], we conducted this study to investigate the evaluation and decision-making needs of stakeholders for the ELP, and to re-design the evaluation approach. The programme has evolved iteratively, guided by feedback from participants and evaluations. Cohorts 10 and 11 are active at the time of publication.

## METHODS

We adhered to the Standards for Reporting Qualitative Research (SRQR) guidelines[19] throughout the study to ensure a rigorous and transparent approach to the design, execution, and reporting of our qualitative research. The purpose of evaluation as outlined by Patton[15] is to provide information to support decision-making. Using Patton’s utilisation-focussed evaluation approach, “a focus on *intended use by intended users* undergirds and informs every design decision in the evaluation”[20]. With this in mind, our first steps were to identify, organise, and engage primary intended users; conduct situation analysis with primary intended users; and identify primary intended uses by establishing the evaluation’s priority purposes[15,20–22].

### SEMI-STRUCTURED INTERVIEWS

Using a four-quadrant stakeholder power and interest matrix[23–25], we identified and arranged semi-structured interviews with key stakeholders for evaluation of these leadership development programmes. Interviews were conducted by a medical doctor and doctoral researcher with expertise in programme evaluation research, over video-call or telephone. Each interview lasted for 30-60 minutes, and was recorded and professionally transcribed for analysis. A total of 12 interviews were

completed with 12 participants, as outlined in Table 3. Data saturation was reached when no new themes or insights emerged from the interviews, indicating that the data collected were sufficient to comprehensively address the research question.

<b>BACKGROUND</b>	<b>ROLE</b>	<b>REFERENCE*</b>
<b>NURSING</b>	Deputy Divisional Nurse	1
	Divisional Nurse	2
	Division Lead for Practice Development & Education, Surgical Division	3
	Divisional Director of Operations	4
<b>MEDICINE</b>	Chief Medical Officer	5
	Associate Dean	6
	Director of Quality Improvement and Deputy Chief Medical Officer	7
<b>MANAGEMENT</b>	Head of Leadership Development (interim), Culture and Leadership Service	8
	Divisional Lead for Education & Practice Development, Medical Division	9
	Director of Culture and Leadership, Culture and Leadership Service	10
<b>LAW</b>	Chief Finance Officer	11
	Chair of Board	12

Table 3: Description of interviewees. \*references are assigned to relevant quotes in the results section

## INTERVIEW ANALYSIS

Interviews were analysed through a process of reflexive thematic analysis as described by Braun and Clarke[26]. This involves: first, reading all the material to become familiar with the data; second, systematically reviewing the transcription line-by-line to identify meaningful codes; third, grouping and summarising groups of codes to initial themes; fourth, reviewing and refining themes; fifth, defining and naming themes; and sixth, interpreting and analysing identified themes. All of the authors contributed to the thematic analysis, with codes and themes refined through discussion. All authors critically reflected on potential personal biases in the thematic analysis process, ensuring that their interpretations were grounded in the participants' perspectives.

## RESEARCH ETHICS APPROVAL

As evaluation of a service delivery, these interviews were approved for an ethics exemption by the University of Oxford Clinical Trials Research Governance Committee.

## PUBLIC AND PATIENT INVOLVEMENT

No patients or the public were involved in this study.

## RESULTS

Three distinct themes were generated from the interviews as key areas to measure for successful evaluation. These were: impact for patients, for healthcare organisations, and for individuals. Within the theme of individual outcomes, sub-themes were generated relating to skills, qualities, and attitudes towards the programme. The relationship between these themes is outlined in Figure 1. A

full table of themes and relevant codes is included as supplementary material (Supplementary Table 1).

### IMPACT FOR PATIENTS

Interviewees emphasised the importance of measuring the impact of a leadership development programme on patients. Suggested measurable outcomes included improvement in patient care, patient experience and improvement in overall quality of service provision as a reflection of enhanced leadership.

*“Basic bottom line is that there has to come some good out of it for the individual and for patient care. If neither has happened, then we're wasting your time.” - Interviewee 5*

### IMPACT FOR THE ORGANISATION

Interviewees discussed the need for a change in culture, and highlighted that they saw leadership development as an important means to achieve this change.

*“If we can develop our leaders, that is one of the biggest aspects, isn't it, around changing our culture is to develop them so they're leading by example and embedding the behaviours and values that we're expecting.” – Interviewee 10*

Collaboration within the workforce and team performance were identified as indicative measures to determine impact on workplace culture. Improved relationships and networking between leaders and across disciplines were frequently discussed as potential outcomes from a leadership development programme. Interviewees suggested evaluating whether participants felt that they had better networking opportunities and whether there had been a shift from traditional hierarchical structures.

*“It should also potentially mean that we've got really effective multi-professional relationships, instead of being in separate groups. And that could have a real impact over a length of time.” - Interviewee 1*

*“It's about bringing those people together and creating those networks because as a leader, it can often be quite lonely. There's definitely the social aspect to the learning and the network element for our leaders, which is what we want to also create as part of leadership development.” – Interviewee*

10

Additionally, an increase in leadership diversity was seen as a core outcome for the organisation.

*“I don't want for the trust to have a blueprint leader. What I want for the trust is to have a range of leaders who can do different things and are sufficiently aware of that and that they'll do them well. And that we're sufficiently aware of, that we know who to call on when we're doing a particular type of challenge.” – Interviewee 12*

Interviewees stressed the importance of evaluating organisational efficiency as a measure of programme success. The outcomes of quality improvement projects undertaken by programme participants was highlighted as a key measure of this.

*“It's probably most worth capturing the real world impact on skills and the real world impact on service, because those things are going to have the most impact in terms of ongoing commissioning of the course and demonstration of benefit.” - Interviewee 3*

A particular focus was placed on evaluating staff retention, recruitment and opportunities for career advancement. Other quantifiable outcomes included financial savings and time savings.

*“It will be really nice to see that those people are staying with us and are going to be part of our future leadership team, in whatever capacity.” - Interviewee 1*

*“It would be interesting to see whether there's an impact on the careers of the people who got the course. So are they promoted? Do they move up? Do they change role? Because it's not just about promotion, sometimes it's about finding the role that suits you.” - Interviewee 3*

### INDIVIDUAL OUTCOMES: KNOWLEDGE AND SKILLS

The importance of measuring individual changes in leadership knowledge and skills was highlighted by interviewees. It was felt that programme participants need not only to learn knowledge and skills but also to feel that they have the confidence to apply these in day-to-day work and life.

*“Because you teach me a skill, I've got to be confident to use it. And it doesn't matter whether I'm using it at the end of the programme, I've got to be confident to go forward and do it again.” - Interviewee 7*

Interviewees particularly emphasised the importance of learning and applying communication skills and coaching skills, pointing out that an evaluation should consider how a programme enhances participants' ability to interact and convey information effectively as well as developing and supporting others within the organisation. Leading with care and taking responsibility for behaviour were highlighted as other aspects for evaluation, encompassing the compassionate side of leadership that's crucial in healthcare settings.

*“There's a set of things about confidence. So, do the people who've been through the programme feel more confident about their leadership. That doesn't necessarily mean competence in the sense that they now think they know it all. And in some ways that would be an indication that it failed. But it is a feeling that they are better equipped to deal with the challenges.” - Interviewee 12*

Knowledge of quality improvement principles was also discussed, reflecting the need to evaluate participants' capabilities to initiate and manage positive change within healthcare settings. Moreover, interviewees saw an increased understanding of broader healthcare landscape and forces as being an important outcome.

*“What I would really like to see courses, groups like this, evolve into, is to produce the clinicians of the future, who will be able to understand all the nuances of running a health service, whether it's a GP surgery, whether it's a department in a hospital, whether it's an education program, whether it's a larger commissioning group, or on the board of a secondary care organization; with all the backgrounds, to make them be able to do that job effectively.” - Interviewee 6*

### INDIVIDUAL OUTCOMES: QUALITIES AND ATTITUDES

Beyond, the acquisition of knowledge and skills, interviewees emphasised the importance of developing self-awareness and understanding of leadership qualities, values and attitudes.

*“This is about people, and kind of how people are going to approach what then the rest of their career potentially. And it's very difficult to measure, because the things that they're going to learn, can be potentially life-changing.” - Interviewee 7*

This included developing an understanding of their own strengths, weaknesses, and to consider and reflect on the impact of their actions.

*“Great leaders are able to lead in ways that are not particularly comfortable for them when it's appropriate. So you need to know what works for you and what doesn't work for you, which is not abandoning things that don't work for you, but sort of working out when that type of leadership is required. Can you do it even though it's a bit against the grain or actually do you need partners to do that with, for whom it's their strength. So it's that sort of combination of self-awareness and the toolkit that leaders can pull upon.” - Interviewee 12*

Interviewees also discussed the possible effects on resilience, empowerment and independence. Some were interested in whether the programme might support participants' development of a sense of autonomy and self-efficacy.

*“Having that opportunity to hear a personal reflection about where the program is taking them. And being able to collect in some way, the impact that it's having on them. Is it empowering them?” - Interviewee 9*

Other beneficial areas for assessment included critical thinking abilities, and whether participants had developed a broader approach to thinking or a willingness to challenge established ways of thinking and acting

*“You want to challenge the way that they think things are done in their area, and just see that you can still do that, but perhaps bring other things in and not feel so threatened by people that do things differently, but to embrace it.” - Interviewee 1*

Changes in attitudes towards future leadership development was also identified as a valuable measure, recognising that effective leadership involves continual development and willingness to adapt.

*“There's something about open-endedness. So finding out whether people are more or less likely to do leadership development in the future. Find out whether people feel this is the end. They've done it now, and they don't need to do it anymore, which would be a negative indicator for me. Or whether, actually what this has opened up is the realisation that this is something that needs to be refreshed or nurtured in some way.” - Interviewee 12*

Empathy for colleagues was often mentioned, underscoring the relational aspect of leadership.

*“It was about having an appreciation of what your colleagues are going through in the workplace, and learning from them, and I really believe that is an element of allowing you to reflect” - Interviewee 9*

Additionally, the wellbeing of both leaders and those they lead was deemed important for understanding the effects of the leadership programme on the workplace environment.

*“I know that you can see the impact of bad leadership in the well-being or morale of the teams people lead. This really is a bit ambitious, but I wonder if, in addition to trying to measure the impact on the well-being of individual leaders who we were developing, if it were ever possible to measure the impact on the teams they lead” - Interviewee 11*

## INDIVIDUAL OUTCOMES: ATTITUDES TOWARDS THE PROGRAMME ITSELF

Interviewees discussed the need for continual improvement of the programme itself, and therefore the need to understand how participants interacted with the programme. They pointed out the importance of understanding participant demographics, attendance and completion, as well as their reasons for choosing the programme, the perceived reputation of the programme, and participants’ recommendations for programme improvement.

*“I think it is very important to demonstrate that the course is useful to people from a variety of backgrounds.”- Interviewee 3*

*“It’s about how we can keep evolving what we’re offering so it doesn’t get stale, it doesn’t stop making a difference. And so yeah, there’s got to be a purpose. I wouldn’t just want to evaluate just for the sake of it. There’s got to be a really clear purpose about seeing what’s working and what you’re going to do differently from it. What will you keep? What will you change?” - Interviewee 10*

Teaching faculty satisfaction and participant feedback on teaching faculty performance were also raised as important outcomes from the programme.

## APPROACHES TO EVALUATION

Interviewees suggested a variety of methods and ways to improve the quality of evaluations, which tie into all of the themes above. Their methodological suggestions included: pre- and post- intervention measures, 360 feedback, long term evaluation, and conducting a needs assessment in order to determine the desired outcomes of the programme.

*“It’d be quite interesting to just see a little bit of a follow-up, I suppose. Does anything ever actually happen? What relationships did you end up keeping and why and things like that. That’s a kind of a longer term thing, but that’d be quite interesting because I think that’s going to be a real key part to any leadership development is that bringing together and the cross-divisional aspect.” - Interviewee 10*

## DISCUSSION

This study provides insights into outcome measures we should endeavour to capture in order to conduct a successful evaluation of the impact of a healthcare leadership development programme. The key domains for evaluation generated from this thematic analysis related to patient outcomes, organisational outcomes, and individual outcomes. Individual outcomes were further classified into those relating to skills or knowledge, those relating to qualities or attitudes, and those relating to the experience of the programme itself. Interviewees also highlighted the perceived value of a mixed methods evaluation, alongside incorporation of pre- and post-measures, 360 feedback, and a longitudinal evaluation.

Our findings support existing evidence that evaluation of leadership development programmes necessitates a multi-faceted approach[7]. There is general consensus that cross-sectional evaluation

designs in leadership development are insufficient; there is a recognised need for longitudinal studies with measurements at multiple time points[7,27–30]. Russon highlighted issues with relying solely on data collected from programme participants[30], and Solansky quoted a number of studies cautioning against exclusive use of self-reported data “which are prone to many kinds of response bias”[31].

Traditional frameworks like Kirkpatrick’s model[32] and Avolio’s evaluation framework[27] offer valuable infrastructure but also have limitations. Many of the identified areas for evaluation in this study could be stratified into Kirkpatrick’s hierarchical framework of “reaction, learning, behaviour, and results”. Yet, these levels often underemphasise the soft skills and personal qualities developed such as resilience, empowerment, and collaboration. This framework also may not fully capture broader cultural shifts and long-term career impacts. Avolio’s approach has some additional benefits in that the “cognitive” domain may more effectively encapsulate aforementioned “soft skills”. However, his team nonetheless experienced difficulties categorising outcomes into their framework. Similar conclusions have been drawn in reviews which have attempted to categorise outcomes into a pre-existing framework[6–8,33–37]. Michael Quinn Patton’s emphasis on evaluation utility informed our study[15]. This stakeholder- and decision-focused approach ensures that the evaluation results are directly applicable to decision-making processes. This aligns with our interviews and the identified need for evaluations that inform and improve programme design and implementation.

Our interviewees highlighted several factors that had not been fully addressed in our previous evaluation approach, particularly outcomes related to organisational culture and personal qualities, which they viewed as more critical than we had initially anticipated. This likely reflects a growing awareness of the importance of culture in fostering job satisfaction, staff retention, and leadership diversity. These missed opportunities for evaluation align with Yardley and Dornan’s critique of Kirkpatrick’s framework, which emphasises the need to consider unexpected outcomes beyond traditional categories[38]. Stakeholders expressed a strong interest in measuring long-term organisational changes, such as workforce transformation, collaboration, and workplace culture—domains that were previously underrepresented in our framework.

While utilisation-focused evaluation is valuable for aligning evaluation protocols with stakeholder needs, its practical application requires balancing aspirational goals with realistic constraints. For example, measuring patient-level outcomes or attributing systemic changes directly to the programme may be unrealistic due to attribution challenges and resource limitations. Moving forward, the evaluation approach will incorporate pre- and post-programme measures, including assessments of staff retention, career progression, and longitudinal follow-up to capture individual and organisational impacts. At the same time, we will temper expectations around outcomes such as financial savings or patient experience shifts, acknowledging that these are influenced by broader systemic factors beyond the programme’s scope. By addressing these new domains and managing expectations, we aim to provide a richer, more actionable evaluation of the programme’s impact.

The generalisability of this study to other programmes and settings may be limited by the number of interviewees and the focus on a single programme. It would be useful in subsequent research to investigate the views of stakeholders in other programmes, as well as the insights of programme faculty and participants, both of whom are important decision-makers in a leadership development programme.

In order to effectively and usefully evaluate a leadership development programme, we found that there was considerable benefit to be gained from approaching real stakeholders to incorporate their needs into the evaluation design. In particular, we observed a need to consider outcomes which are not yet included in standard training programme evaluation frameworks. These included a broader range of qualitative and quantitative measures, as well as longitudinal follow-up. Finally, we concluded that the specific nature of stakeholder needs with respect to evaluation means that evaluations should ideally be focused on utilisation and tailored to each programme's specific aims. By doing so, we can ensure that healthcare leadership development programme evaluations are useful to programme stakeholders and also contribute to the wider literature regarding healthcare leadership development.

## CONTRIBUTORS

OL is responsible for the overall content (as guarantor). OL planned the study and conducted the interviews. All authors contributed to thematic analysis. OL and JP drafted the manuscript and revised it critically for intellectual content. All authors approved the final version of the manuscript. All authors agreed to be accountable for all aspects of the work, ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

## COMPETING INTERESTS

OL designed the ELP programme and volunteered his time as faculty. JG is a graduate of the first Emerging Leaders Programme, and contributed to cohorts 2 and 3 as graduate programme director. JP was supported in this research by Thrum Leadership Ltd., who continue to deliver the ELP programmes.

## FUNDING STATEMENT

OL completed this research with the support of a Rhodes Scholarship (Award/Grant number N/A).

## ACKNOWLEDGEMENTS

We would like to thank the many speakers, mentors and facilitators who volunteered their time to deliver the ELP. Thanks are also due to the Oxford University Hospitals Medical Education Department and Health Education England Thames Valley for providing catering for the programme. A particular thanks is also due to the programme graduates who contributed to the programme as teaching faculty, whose enthusiasm enabled the programme to continue.

Finally, we are grateful to the programme stakeholders who offered their time as interviewees to help us better understand how to evaluate healthcare leadership development programmes.

## REFERENCES

1. Tasi MC, Keswani A, Bozic KJ. Does physician leadership affect hospital quality, operational efficiency, and financial performance? *Health Care Manage Rev* [Internet]. 2017;44(3):256–

62. Available from:  
[https://journals.lww.com/hcmrjournal/fulltext/2019/07000/does\\_physician\\_leadership\\_affect\\_hospital\\_quality..8.aspx](https://journals.lww.com/hcmrjournal/fulltext/2019/07000/does_physician_leadership_affect_hospital_quality..8.aspx)
2. West M, Armit K, Loewenthal L, et al. Leadership and leadership development in health care: the evidence base. The King's Fund [Internet]. 2015 [cited 2024 Mar 14]. Available from: <https://www.kingsfund.org.uk/insight-and-analysis/reports/leadership-development-health-care>
  3. Goodall AH. Physician-leaders and hospital performance: is there an association? Soc Sci Med [Internet]. 2011;73(4):535–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/21802184/>
  4. Spurgeon P, Long P, Clark J, et al. Do we need medical leadership or medical engagement? Leadersh Health Serv (Bradf Engl) [Internet]. 2015;28(3):173–84. Available from: <https://pubmed.ncbi.nlm.nih.gov/26083633/>
  5. Bohmer R. The instrumental value of medical leadership: engaging doctors in improving services. [Internet]. 2012. Available from: <https://www.england.nhs.uk/wp-content/uploads/2022/05/engaging-doctors-improving-services.pdf>
  6. Frich JC, Brewster AL, Cherlin EJ, et al. Leadership development programs for physicians: a systematic review. J Gen Intern Med [Internet]. 2015;30(5):656–74. Available from: <https://link.springer.com/content/pdf/10.1007/s11606-014-3141-1.pdf>
  7. Lyons O, George R, Galante JR, et al. Evidence-based medical leadership development: a systematic review. BMJ Leader [Internet]. 2021;5(3):206–13. Available from: <https://bmjleader.bmj.com/content/5/3/206>
  8. Geerts JM, Goodall AH, Agius S. Evidence-based leadership development for physicians: a systematic literature review. Soc Sci Med. 2020;246:112709. doi:10.1016/j.socscimed.2020.112709
  9. Lacerenza CN, Reyes DL, Marlow SL, et al. Leadership training design, delivery, and implementation: a meta-analysis. J Appl Psychol [Internet]. 2017 [cited 2024 Mar 1]. Available from: <http://dx.doi.org/10.1037/apl0000241>
  10. Lockwood C, Munn Z, Porritt K. Qualitative research synthesis. Int J Evid Based Healthc [Internet]. 2015;13(3):179–87. doi:10.1097/XEB.000000000000062
  11. Reed DA, Cook DA, Beckman TJ, et al. Association between funding and quality of published medical education research. JAMA [Internet]. 2007;298(9):1002–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/17785645/>
  12. Joanna Briggs Institute. Chapter 7: Systematic reviews of etiology and risk. JBI Manual for Evidence Synthesis. 2020.
  13. Lyons O. Evaluating medical leadership development programmes [doctoral thesis]. University of Oxford; 2021.
  14. Donaldson SI, Patton MQ, Fetterman DM, et al. The 2009 Claremont debates: the promise and pitfalls of utilization-focused and empowerment evaluation. J Multidiscip Eval. 2010;6(14):159–65.
  15. Patton MQ. Utilization-focused evaluation. 4th ed. Thousand Oaks, CA: SAGE; 2008.
  16. Hannum K, Craig SB. Introduction to special issue on leadership development evaluation. Leadersh Q. 2010;21(4):581–2. doi:10.1016/j.leaqua.2010.06.007
  17. Rabin R. Blended learning for leadership: the CCL approach [Internet]. 2014. Available from: <https://www.ccl.org/articles/white-papers/blended-learning-for-leadership/>

18. Pendleton D, Furnham A. Leadership: all you need to know. 2nd ed. Basingstoke: Palgrave Macmillan; 2016.
19. O'Brien BC, Harris IB, Beckman TJ, et al. Standards for reporting qualitative research: a synthesis of recommendations. *Acad Med* [Internet]. 2014;89(9):1245–51. doi:10.1097/ACM.0000000000000388
20. Patton MQ. Qualitative research & evaluation methods. 3rd ed. Thousand Oaks, CA: SAGE; 2002.
21. Patton MQ. Utilization-focused evaluation (U-FE) checklist [Internet]. 2013 [cited 2024 Mar 12]. Available from: [www.wmich.edu/evalctr/checklists](http://www.wmich.edu/evalctr/checklists)
22. Erickson ASG, Noonan PM. The SAGE encyclopedia of educational research, measurement, and evaluation. Thousand Oaks, CA: SAGE; 2018.
23. Schmeer K. Stakeholder analysis guidelines. *Analysis* [Internet]. 1999;15:338–45. Available from: <https://www.who.int/workforcealliance/knowledge/toolkit/33.pdf>
24. Patton MQ. Stakeholders. In: The SAGE encyclopedia of educational research, measurement, and evaluation [Internet]. 2018 [cited 2020 Nov 1]. Available from: <https://dx.doi.org/10.4135/9781506326139>
25. NHS Improvement. Stakeholder analysis [Internet]. 2018. Available from: <https://www.england.nhs.uk/wp-content/uploads/2022/02/qsir-stakeholder-analysis.pdf>
26. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3(2):77–101. doi:10.1191/1478088706qp063oa
27. Avolio BJ, Reichard RJ, Hannah ST, et al. A meta-analytic review of leadership impact research: experimental and quasi-experimental studies. *Leadersh Q*. 2009;20(5):764–84. doi:10.1016/j.leaqua.2009.06.006
28. Day DV, Fleenor JW, Atwater LE, et al. Advances in leader and leadership development: a review of 25 years of research and theory. *Leadersh Q* [Internet]. 2014;25(1):63–82. doi:10.1016/j.leaqua.2013.11.004
29. Dinh JE, Lord RG, Gardner WL, et al. Leadership theory and research in the new millennium: current theoretical trends and changing perspectives. *Leadersh Q* [Internet]. 2014;25(1):36–62. doi:10.1016/j.leaqua.2013.11.005
30. Russon C, Reinelt C. The results of an evaluation scan of 55 leadership development programs. *J Leadersh Organ Stud*. 2004;10(3):104–7. doi:10.1177/107179190401000309
31. Solansky ST. The evaluation of two key leadership development program components: leadership skills assessment and leadership mentoring. *Leadersh Q* [Internet]. 2010;21(4):675–81. doi:10.1016/j.leaqua.2010.06.009
32. Kirkpatrick DL. Techniques for evaluating training programs. *Train Dev J*. 1979;178–92.
33. Onyura B, Crann S, Tannenbaum D, et al. Is postgraduate leadership education a match for the wicked problems of health systems leadership? A critical systematic review. *Perspect Med Educ*. 2019;8(3):133–42. doi:10.1007/s40037-019-0515-1
34. Sadowski B, Cantrell S, Barelski A, et al. Leadership training in graduate medical education: a systematic review. *J Grad Med Educ* [Internet]. 2018;10(2):134–48. doi:10.4300/JGME-D-17-00194.1
35. Steinert Y, Naismith L, Mann K. Faculty development initiatives designed to promote leadership in medical education: a BEME systematic review. *Med Teach* [Internet]. 2012;34(6):483–503. doi:10.3109/0142159X.2012.680937

36. Sultan N, Torti J, Haddara W, et al. Leadership development in postgraduate medical education: a systematic review of the literature. *Acad Med* [Internet]. 2019;94(3):440–9. doi:10.1097/ACM.0000000000002503
37. Webb A, Tsipis N, McClellan T, et al. A first step toward understanding best practices in leadership training in undergraduate medical education. *Acad Med* [Internet]. 2014;89(11):1563–70. doi:10.1097/ACM.0000000000000514
38. Yardley S, Dornan T. Kirkpatrick’s levels and education ‘evidence’. *Med Educ* [Internet]. 2012;46(1):97–106. doi:10.1111/j.1365-2923.2011.04076.x

Caption:

Figure 1: Important areas for evaluation of a healthcare leadership development programme.