The evaluation of health care leadership development programmes

Burn, Emily; Waring, Justin

DOI:
10.1108/LHS-05-2022-0056

Document Version
Peer reviewed version

Citation for published version (Harvard):
Burn, E & Waring, J 2022, 'The evaluation of health care leadership development programmes: a scoping review of reviews', *Leadership in Health Services*. https://doi.org/10.1108/LHS-05-2022-0056

Leadership in Health Se

The Evaluation of Healthcare Leadership Development Programmes: A Scoping Review of Reviews

Journal:	*Leadership in Health Services*
Manuscript ID	LHS-05-2022-0056.R2
Manuscript Type:	Original Article
Keywords:	Health care, Leadership, Evaluation, Scoping review, Management development

The Evaluation of Healthcare Leadership Development Programmes: A Scoping Review of Reviews

Abstract

Purpose

There is continued investment in healthcare leadership development programmes (HLDPs), and despite growing calls for robust evaluations of their pedagogic design, delivery, and effectiveness, there are concerns regarding the quality of data associated with their evaluation. This scoping review of reviews investigated the reporting of HLDP evaluations to determine: 1. how the conceptualisation of leadership underpinning HLDPs influence their evaluation; 2. how the pedagogical approaches within HLDPs influence their evaluation; 3. the evaluation designs and measures used to assess HLDPs.

Design/methodology/approach

The scoping review was conducted on reviews of HLDPs. Searches were performed on four databases and on the grey literature. Data were extracted and a narrative synthesis was developed.

Findings

Thirty-one papers were included in the scoping review of reviews. A great deal of heterogeneity in HLDPs was identified. Evaluations of HLDPs were affected by poor data quality and there were limitations in the evidence about 'what works'. Leadership was conceptualised in different ways across HLDPs and consequently, there was a lack of consistency as to what is being evaluated and the methods used to assess HLDPs.

Originality/ value

This review of reviews summarises the current evidence on the evaluation of HLDPs. Evaluations of HLDPs need to explicitly account for the complexity of health systems, how this complexity impacts on the development and articulation of leadership practice, and how the underlying conceptualisation of leadership and associated theory of change articulate a set of assumptions about how HLDPs support leaders to affect change within complex systems.

Keywords: Healthcare, Leadership, Management Development, Evaluation, Review

Paper type: Literature review

Purpose

Effective leadership in healthcare has been associated with the maintenance and improvement of quality and safety in health services (West *et al.*, 2015; Sfantou *et al.*, 2017). Significant resources have

been invested in the development of healthcare leadership development programmes (HLDPs) for both clinical and non-clinical sections of the healthcare workforce (MacDonald, 2014; West *et al.*, 2015; Stoller, 2020). Yet, the contributions of HLDPs remain uncertain, especially in the context of changing organisational demands and shifts in emphasis from team or organisational leadership to systems leadership (Edmonstone, 2014; Curry *et al.*, 2020). Evaluations of HLDPs have often been concerned with establishing the scope, or availability of leadership development programmes and exploring 'what works' within interventions. Yet, multiple reviews highlight concerns about the quality of evaluations (Busari, 2012; Careau *et al.*, 2014; Jeyaraman *et al.*, 2018; Pizzirani *et al.*, 2020; Ahrari *et al.*, 2021). West *et al.* (2015, p. 13) summarise the main weaknesses in evaluation as relating to:

> *…small sample sizes; lack of underpinning theory; survey instruments with inadequate reliability and validity; failure to measure important control variables; cross sectional designs; reliance on self-report (e.g. for measuring patient safety); and poor measurement of leadership (not systematic).*

MacDonald (2014, p. 228) notes that a lack of robust evidence does not mean necessarily that HLDPs have no effect and do not improve participants' leadership skills. Instead, it is more that we are uncertain as to what contributes to the success of programmes and how we conceptualise the outcomes expected from HLDPs. Hence there is a need to review the existing literature on the evaluation of HLDPs to investigate the evaluation approaches taken, how these might influence the prevailing evidence-base, and what recommendations could be developed for future evaluations.

This paper reports on a scoping review of previously published reviews of evaluations of HLDPs with the aim of deriving recommendations for the future evaluation of HLDPs. The scoping review of reviews builds on Edmonstone's (2013) evaluation framework to consider, i) how leadership is conceptualised in these evaluations, ii) the attention given to appraising pedagogic design, and iii) the implications of these conceptualisations on the evaluation of HLDPs. Leadership can be conceptualised in multiple ways and interpretations of leadership will inform the methods used to convey the different theories of leadership underpinning HLDPs. Similarly, the evaluation design and methods used to assess HLDPs should be rooted in the programme's aims. Assessing the association between these constituent parts provides a rounded illustration of the current state of evaluation of HLDPs. Following from this aim, three research questions guide the scoping review of reviews.

1. *How does the conceptualisation of leadership underpinning HLDPs influence their evaluation?*

To evaluate a HLDP, it is necessary to understand how a given programme either implicitly or explicitly understands leadership. Exploring how HLDPs have conceptualised leadership provides understanding of the 'content' and 'purpose' of the HLDP on which to understand the strengths and limitations of their evaluation.

2. *How do the pedagogical approaches within HLDPs influence the evaluation of programmes?*

To evaluate a HLDP it is also necessary to understand the pedagogical approaches that inform the development of the anticipated leadership qualities and characteristics. Exploring pedagogical approaches brings to the fore the way in which the 'structure' and 'design' of a HLDP may influence its evaluation.

3. *What evaluation designs and measures are used to assess HLDPs?*

To evaluate a HLDP it is also necessary to understand how a given evaluation is itself designed, and how measures of the HLDP design, delivery, and outcomes are developed and operationalised. That is, what do evaluations intend to measure and analyse and are these consistent with the conceptualisation of leadership and pedagogic design.

Method

A scoping review of reviews was conducted with reference to guidance developed by Arksey and O'Malley (2005), Levac *et al.* (2010) and Peters *et al.* (2015). Following Peters *et al.* (2015), the review's scope was determined with reference to the population, concept, and context. The 'population' was defined as reported evaluations of HLDPs. The 'concept' was defined as evaluations of HLDPs in terms of structured interventions designed to develop leadership skills, knowledge, and behaviours. The 'context' refers to the health systems in which the HLDPs take place.

A search strategy was designed for four social science and humanities bibliographic databases: The Health Management and Policy Database, from The Healthcare Management Information Consortium (HMIC), Medline, EMBASE, and PsycINFO. The search strategy was developed in consultation with a

stakeholder group comprising UK-based educators, researchers, leadership practitioners, commissioners of HLDPs, and evidence review specialists. The agreed search strategy, terms and Boolean operators were adapted for each database. The search strategies are included in supplemental table 1 and comprised the following search terms: healthcare, leadership, management, programme, training programme, leadership development, evaluation, audit, programme evaluation, systematic reviews, synthesis, and review.

Further searches were performed on Open Grey to capture grey literature. Additional searches were conducted using Google Scholar and the first ten pages of results were screened for inclusion. The scoping review of reviews was limited to articles and publications which considered the evaluation of HLDPs. The primary basis of inclusion was whether the article reviewed the evaluation of HLDPs. The following inclusion and exclusion criteria informed the selection of reviews that were incorporated within the scoping review of reviews.

Inclusion criteria

- Publications and articles on the topic of leadership development in healthcare.
- Publications and articles on the evaluation of HLDPs (including programmes targeted to a single profession e.g. nursing).
- Articles on the process of evaluating healthcare leadership development programmes.
- Articles evaluating HLDPs targeted to clinicians and administrators.
- Articles conducted internationally or nationally.
- Articles published from 1st January 2000 to 31st May 2021. This period saw an increased investment in healthcare leadership development.
- Available in English.

Exclusion criteria

- Articles that only consider leadership development at an undergraduate level.
- Primary evaluations of single HLDPs.
- Articles limited to the review of short courses intended to develop leadership skills.
- Reviews of leadership development in academic health centres.
- Reviews of team-based, task-oriented, leadership development when a team is focused on a specific procedure or process (e.g., resuscitation teams within a hospital setting).

- Articles where the full text article could not be accessed.

Primary evaluations of single HLDPs were excluded as the scoping review focused on exploring overarching themes in healthcare leadership development. Articles that review evaluations of undergraduate HLDPs were only included if the review also reported on postgraduate or practitioner levels. Academic health centres facilitate research and education and provide clinical care. Leadership in this setting encompasses both healthcare professionals as well as higher education faculty and so reviews of evaluations in this setting were excluded. In practice, evaluations varied in design, scope, and duration, with some including brief feedback appraisals of workshop design programmes, and others longitudinal studies of change in leadership knowledge, skills, and behaviour. Initial searches generated 6,696 entries and 1,190 duplicates were deleted. The titles and abstracts of 5,506 studies were reviewed and full-text reviews were conducted on 37 studies. The large reduction of sources was due to the initial search including studies which addressed issues around leading quality improvement and service management, along with individual evaluations of leadership development programmes (including programmes in settings other than healthcare). These publications were excluded. Any uncertainty as to whether a publication should be included was resolved with reference to a second opinion. In total, 31 records were included in the scoping review of reviews (as noted in figure 1).

Insert here: Figure 1: PRISMA flow diagram (Page *et al.*, 2021)

Extraction and analysis

Data from the 31 included reviews were extracted using a standardised template, including: aim of review, study setting and sector, study population, discussion of theory of change/ conceptualisation of leadership, objectives of programmes, evaluation design (including measures and indicators), pedagogical methods and reported challenges or limitations of evaluations. Data are summarised (table 1).

Insert here: Table 1: Summary of included studies

Findings

The included articles frequently referred to limitations in the evidence of 'what works' in HLDPs. The heterogeneity of HLDPs (in terms of content and structure) and low quality of evaluation design was suggested to restrict comparisons and limit the generalisations that could be drawn (Cummings *et al.*,

2008; Busari, 2012; Careau *et al.*, 2014; Jeyaraman *et al.*, 2018; Sadowski *et al.*, 2018; Flaig *et al.*, 2020). In response to the objectives of the scoping review of reviews, the conceptualisation of leadership, pedagogical methods, evaluation methods and metrics are discussed.

Multiple conceptualisations of leadership

The conceptualisation of leadership informing HLDPs should influence the focus or target of a given evaluation so that the right outcome measures that reflect the aims of the programme are used. As noted by Cummings (2008, p. 246), theoretical frameworks surface the 'ideas and variables' which inform programme design and are assessed through evaluation. Therefore, exploring the conceptualisation of leadership, or the theoretical underpinning of the programme reveals the clarity of the aims of the programme, as well as the appropriateness of the programme content and how it is delivered.

Several studies identified that evaluations should capture multiple perspectives of leadership (Frich *et al.*, 2015; Ahrari *et al.*, 2021). In other words, evaluations should consider an individual level of leader development, along with the 'effects' (or outcomes) of the programme at an organisational level. Kirkpatrick and Kirkpatrick's (2006) four-level evaluation model was often referred to as offering a way to conceptualise these different 'levels' of outcomes and to categorise the data required to evaluate HLDPs. Within Kirkpatrick's model, level 1 refers to reaction and refers to participants' level of satisfaction with the programme content and organisation. Level 2A captures a change in attitudes and level 2B relates to a change in knowledge or skills. Level 3A encapsulates self-reported behavioural change and level 3B captures behaviour change as recognised by a third party. Level 4A captures self-reported organisational outcomes and level 4B refers to 'verified' organisational outcomes as captured by a series of metrics.

Across the reviews, HLDPs were informed by multiple conceptualisations of leadership. Overall, there was no shared underpinning conceptualisation of leadership informing HLDPs, and hence evaluations assess different aspects of leadership. A number of reviews identified overarching conceptualisations of leadership (Careau *et al.*, 2014; Jeyaraman *et al.*, 2018; Johnson *et al.*, 2020). One prominent distinction in the literature is between individualised leader-development and socially situated leadership-development in which an emphasis is placed on the 'collective capacity' to enact change

(Frich *et al.*, 2015, p. 656), or the networks and the relationships formed between the components of a wider health system (Edmonstone, 2011).

A number of reviews identified that some evaluations did not discuss the theory of change or conceptualisation of leadership informing the HLDP (Mianda and Voce, 2018; Onyura *et al.*, 2019; Sultan *et al.*, 2019; Torres-Landa *et al.*, 2021). In a similar vein, Edmonstone (2013, p. 153) discusses the 'conceptual fuzziness' when there was a lack of discussion as to the 'guiding assumptions' informing the HLDPs. While a number of reviews did make a connection between the conceptualisation of leadership and evaluation (for example, Turner *et al.*, 2018), others did not (Mianda and Voce, 2018; Kumar *et al.*, 2020; Seidman *et al.*, 2020; Stoller, 2020; Lyons *et al.*, 2021). There was limited critical discussion about the conceptualisation of leadership and how this may affect the evaluation of HLDPs within these reviews. In other words, it was unclear as to what model or framework of leadership informed the content and approach of the HLDP (or whether the programmes were informed by such a model). Evaluation of HLDPs requires the elaboration of how the programme influences the knowledge and skills of the participants and would suggest that some discussion of the underlying theory or conceptualisation of leadership is outlined. Incorporating a conceptualisation, or theory, of leadership develops the rationale for the design of the evaluation and the outcomes data that is collected. The omission of a discussion of the conceptualisation of leadership limited the explanatory accounts of the association between HLDPs and identified outcomes.

Multiple pedagogical approaches

HLDPs are designed and structured according to pedagogical assumptions and approaches about how leadership can be developed within people or the wider workforce. In turn, evaluations of HLDPs should entail consideration and assessment of the relevance and utility of the pedagogy embedded in a programme. Across the reviews, there was a lack of exploration as to how a HLDP's conceptualisation of leadership may inform a programme's pedagogical approach and related evaluation design This disconnect represents a knowledge gap in our understanding of the evaluation of HLPDs.

The reviews reveal the variety of pedagogical methods used within HLDPs. Reviews reported that didactic lectures were frequently used as a teaching method (Careau *et al.*, 2014; Frich *et al.*, 2015; Mianda and Voce, 2018; Onyura *et al.*, 2019; Johnson *et al.*, 2020; Kumar *et al.*, 2020; Ahrari *et al.*,

2021; Lyons *et al.*, 2021; Ravaghi *et al.*, 2021). Other pedagogical methods included reflective exercises (Careau *et al.*, 2014), small discussion groups (Careau *et al.*, 2014; Sultan *et al.*, 2019; Lyons *et al.*, 2021; Ravaghi *et al.*, 2021) and simulated cases (Careau *et al.*, 2014). Experiential learning was also built into HLDPs as participants completed leadership projects and action-based learning (Edmonstone, 2013; Frich *et al.*, 2015; Mianda and Voce, 2018; Sultan *et al.*, 2019; Johnson *et al.*, 2020; Ahrari *et al.*, 2021; Lyons *et al.*, 2021; Ravaghi *et al.*, 2021; Torres-Landa *et al.*, 2021). Experiential learning has been suggested to promote self-reflection and places an emphasis on problem solving (Bekas, 2015) and can build or reinforce a safe culture (Sonnino, 2016). Mentoring and coaching were also identified as a teaching and learning method within HLDPs (Curtis *et al.*, 2011; Edmonstone, 2013; Sonnino, 2016; Mianda and Voce, 2018; Sultan *et al.*, 2019; Johnson *et al.*, 2020; Ravaghi *et al.*, 2021; Torres-Landa *et al.*, 2021). A small number of reviews noted that HLDPs targeted to physicians often emphasised progressing 'conceptual knowledge' through the use of lectures and seminars (Frich *et al.*, 2015, p. 672) and that there was the potential for experiential methods to be developed for this target group (Blumenthal *et al.*, 2012; Frich *et al.*, 2015).

Overall, there was limited discussion of the association between pedagogical methods and the evaluation of HLDPs. Experiential learning, often including simulation and role playing was identified to be a way to support the translation of programme learning to the workplace (Onyura *et al.*, 2019; Pizzirani *et al.*, 2020; Lyons *et al.*, 2021; Torres-Landa *et al.*, 2021). While the reviews do not identify one 'best' pedagogical approach, there is some evidence to suggest that HLDPs should consider how to build in experiential learning and approaches which support the development of participants' self-awareness. Nevertheless, there was limited exploration of how decisions regarding the evaluation of HLDPs may be informed by their pedagogical methods. The developing evidence on the connection between pedagogy and the outcomes associated with HLDPs could be strengthened were a greater emphasis placed on assessing how evaluation designs respond to the different teaching methods found in HLDPs.

Evaluating HLDPs: Measuring and explaining HLDPs outcomes

Data measurement and appraisal within evaluation design and methods should be aligned with the programme's conceptualisation of leadership and pedagogic approach. A range of evaluation designs were identified across the publications. For example, Jeyaraman *et al.* (2018) identifies case studies, mixed methods studies, interrupted time series, longitudinal studies, pre- and post-test designs,

surveys, and retrospective studies. Edmonstone (2013) endorses the use of prospective designs that are built into the development of a HLDP but notes that resource limitations often result in post-hoc approaches. Evaluation designs were noted often to be short-term and lack a longitudinal focus that assess the implementation of learning gained via a HLDP (Johnson *et al.* 2021; Kumar *et al.* 2020).

Quantitative methods were often noted to provide a broad overview of the programme performance (as defined within the evaluation design). The range of identified quantitative methods included: the use of feedback questionnaires (Edmonstone, 2013; Careau *et al.*, 2014; Mianda and Voce, 2018; Sadowski *et al.*, 2018; Sultan *et al.*, 2019; Johnson *et al.*, 2020; Ahrari *et al.*, 2021; Lyons *et al.*, 2021), self-report questionnaires (Busari, 2012; Mianda and Voce, 2018; Sadowski *et al.*, 2018; Ahrari *et al.*, 2021), including validated questionnaires (Cummings *et al.*, 2008), and knowledge tests timed pre and post the intervention (Busari, 2012; Careau *et al.*, 2014; Sultan *et al.*, 2019; Johnson *et al.*, 2020; Ahrari *et al.*, 2021). The collection of quantitative data and 'formal statistics' (Geerts *et al.*, 2020, p. 8), or patient-centred outcomes (Careau *et al.*, 2014; Lyons *et al.*, 2021) were also identified, along with reflections on participants' behaviours by peers, subordinates, or managers (Lyons *et al.*, 2021).

Qualitative methods were often identified as a way of capturing programme participants' experiences in greater depth. Qualitative methods identified in the reviews included participant interviews or focus groups (Careau *et al.*, 2014; Mianda and Voce, 2018; Johnson *et al.*, 2020), interviews with stakeholders, assessments of participants' behaviour (Sadowski *et al.*, 2018; Johnson *et al.*, 2020), participants' reflections and observation of behaviour (Johnson *et al.*, 2020), and observation of participants' action learning sets (Mianda and Voce, 2018). Other identified evaluation methods included exploring participants' career progression (Johnson *et al.*, 2020).

Frequently, the studies identified that evaluations of HLDPs often placed greater emphasis on capturing participants' reactions and self-reports, rather than assessing organisational or systems-level outcomes (Careau *et al.*, 2014; Frich *et al.*, 2015; Sadowski *et al.*, 2018; De Brún *et al.*, 2019; Onyura *et al.*, 2019; Johnson *et al.*, 2020; Kumar *et al.*, 2020; Seidman *et al.*, 2020; Ahrari *et al.*, 2021; Ravaghi *et al.*, 2021). Kirkpatrick and Kirkpatrick's (2006) model was often used as a way to assess the robustness of evaluation data, with data at level 3b and 4 (behaviour change acknowledged by a third-party and organisational outcomes) viewed to be of a higher quality (Mianda and Voce, 2018;

Sadowski *et al.*, 2018; Sultan *et al.*, 2019; Flaig *et al.*, 2020; Geerts *et al.*, 2020; Lyons *et al.*, 2021; Ravaghi *et al.*, 2021).

The selected reviews did include evaluations that directly measured the effects of HLDPs at an organisational level. For example, a range of return on investment (ROI) indicators associated with HLDPs were also identified and represented both financial benefits but also wider outcomes of benefit to both patients and staff (Jeyaraman *et al.*, 2018). Reviews that did discuss organisational outcomes reported a series of metrics, including reporting quality indicators for specified conditions (Frich *et al.*, 2015) or reporting department-level outcomes (Stoller, 2020), or outcomes achieved within work-based projects incorporated within HLDPs (Johnson *et al.*, 2020; Kumar *et al.*, 2020; Lyons *et al.*, 2021). Other studies also noted patient feedback, or satisfaction as a further measure (Frich *et al.*, 2015), as well as more objective measures such as length of hospital stay, mortality rates and reduced clinical errors (Husebø and Akerjordet; 2016; Geerts *et al.*, 2020). While there was acknowledgement of the challenges of attributing organisational outcomes to HLDPs (De Brún *et al.*, 2019; Onyura *et al.*, 2019; Kumar *et al.*, 2020; Seidman *et al.*, 2020), such discussion could be quite limited. Therefore, the evaluation of HLDPs presents an incomplete picture which can lack a developed explanation for the identified changes. Again, there is a need for clear justifications for the outcomes collected as part of an evaluation and these justifications will revolve around the conceptualisation of leadership and the way that these conceptualisations have been embedded within the pedagogical approaches of the programmes.

Reported limitations of HLDP: An under theorised exploration of causation

Across the included reviews, weaknesses in the evaluation of HLDPs were often noted based on evaluation design and data collection methods. Some studies make recommendations as to how the assessment of HLDPs could be strengthened. Nevertheless, the recommendations that emerge are limited as insufficient attention is paid to the complexity of health systems and the effect that such complexity has on the evaluation of HLDPs (although Edmonstone, 2013 and Careau *et al.*, 2014 do acknowledge this complexity). Several reviews (Mianda and Voce, 2018; Sadowski *et al.*, 2018; Sultan *et al.*, 2019; Flaig *et al.*, 2020; Geerts *et al.*, 2020; Lyons *et al.*, 2021; Ravaghi *et al.*, 2021) suggest that the evaluation of HLDPs could be improved by collecting data that fulfils the higher levels of Kirkpatrick and Kirkpatrick's model of training evaluation (levels 3B/4A/4B). In other words, evaluations should focus on collecting data on observed behaviour change and self-reported and observed organisational

outcomes. Furthermore, other reviews suggest that evaluations should incorporate the Medical Education Research Study Quality Instrument (MERSQI) (Kumar *et al.*, 2020; Lyons *et al.*, 2021) to strengthen the evaluation of HLDPs. Kirkpatrick and Kirkpatrick's model and the MERSQI encourage reflection on data collection and the reporting of findings and so may be helpful during the initial stages of evaluation design. However, it is questionable the extent to which the use of these frameworks alone will progress the evaluation of HLDPs. What is missing within these recommendations is an exploration of the association between the conceptualisation of leadership that is weaved throughout HLDPs and evaluation design. To put it another way, it is not just completing the checklist incorporated within MERSQI, or the collection of 'objective' outcome data that will strengthen evaluation. Instead, a greater emphasis should be placed on the underpinning theory informing HLDPs and how the conceptualisation of leadership feeds into the evaluation design.

Discussion

The aim of the review was to understand how HLDPs have been evaluated and the associated challenges of evaluating HLDPs, rather than assessing the programmes themselves, or the outcomes they achieve. The scoping review of reviews was designed to understand the linkages between how evaluation studies conceptualised leadership and the pedagogic design of leadership programmes as the primary focus of their enquiries, and to understand the different evaluation approaches and methods used, the different sources of evidence, and the range of reported limitations.

Conceptualising leadership and pedagogy

Findings from the scoping review of reviews reveal that there is no single conceptualisation of leadership that informs development programmes. Similarly, HLDPs incorporate a variety of pedagogical approaches and do not follow a prescribed structure. The diversity of ideas regarding what leadership 'is' and the pedagogical approaches to convey leadership development reflect the different contexts in which healthcare leadership operates. ~~Avolio et al. (2010) incorporate the length of programme within their analysis of the return on investment in leadership development. While the included reviews often noted that HLDPs were of different durations, there was minimal discussion as to whether the length of the programme informed the outcomes achieved.~~

HLDPs have different aims and, subsequently, the evaluation of HLDPs will have different targets. ~~— there is a lack of consistency as to what is being evaluated. Similarly, HLDPs incorporate a variety of pedagogical approaches. The diversity of approaches within leadership development and pedagogical~~

~~approaches reflects the different contexts in which healthcare leadership operates.~~ There is limited exploration as to how evaluation designs reflect ~~account for~~ the different content and organisation of HLDPs. ~~There are some suggestions that experiential approaches support the translation of learning into practice. However, this emerging evidence base would be strengthened if we understood more as to how evaluation designs are informed by the structure of the programme.~~ Across the reviews, t~~T~~here were some ~~are some~~ suggestions that experiential approaches support the translation of learning into practice. However, this emerging evidence base would be strengthened if we understood more as to how evaluation designs are informed by the structure of the programme. For instance, Avolio et al. (2010) incorporate the length of programme within their analysis of the return on investment in leadership development. While the included reviews often noted that HLDPs were of different durations, there was minimal discussion as to whether the length of the programme informed the outcomes achieved. The included articles often discussed the weaknesses of evaluation designs and methods that emphasise self-reported measures over exploring the impact of programmes at the organisational and systems level. ~~Finally, the included articles often discussed the weaknesses of evaluation designs and methods that emphasise self-reported measures over exploring the impact of programmes at the organisational and systems level.~~

The scoping review of reviews began with a discussion of how the conceptualisation of leadership should reverberate through the design and evaluation of HLDPs. However, this association was often under theorised within the included articles. While recommendations to improve the evaluation of HLDPs were presented, these were limited to discussions on methods. Instead, our understanding would be improved if there were greater exploration of the challenges of evaluating HLDPs. The evaluative limitations regarding the omission of a conceptualisation of leadership and theory of change become more apparent when we consider the complexity of health systems. It is this complexity and the implications for the evaluation of HLDPs that the review now turns.

Complex health systems: The challenge of attribution and evaluation

The complexity of the health system influences the evaluation of leadership development programmes. The health system has been described as a complex adaptive system comprising of components which interact with one another in an unpredictable way (Braithwaite *et al.*, 2013). Within this view, healthcare leaders are embedded within the health system and outcomes emerge from interactions between agents, rather than being the effects of isolated 'independent variables'

(Lichtenstein *et al.*, 2006, p. 2). Leaders facilitate 'the conditions within which the process occurs' and are not viewed to be a 'direct source of change' (Lichtenstein *et al.*, 2006, p. 3). This will affect the evaluation of HLPDs as the leadership 'success' will be informed by a number of external factors, some of which will be likely outside of the control of individuals. Therefore, given the complexity of health systems, it is difficult to isolate the 'effects' of HLDPs and attribute the achieved outcomes to the programmes (Edmonstone, 2013).

The evaluation of HLDPs is concerned with identifying the influence that the programme has had and the extent to which leadership development has led to identified outcomes. However, to view healthcare leaders as being embedded within the wider health system counsels us against viewing causation as a linear process (Edmonstone, 2013) and the evaluation of HLDPs should respond to this complexity. While a number of the included reviews do discuss the difficulty of attributing outcomes to HLDPs, this scoping review of reviews has revealed a research gap in terms of how the evaluation of HLDPs can respond to this complexity. Edmonstone's (2013) framework does reference the complexity of health systems, however, there is the opportunity to develop further evaluative approaches which can account for the contextual positioning of healthcare leaders. Avolio et al.'s (2010) work to estimate the return gained by investing in general leadership development notes that any return on investment analysis will be dependent on the quality of data and assumptions informing the analytical model. Due to the importance of assessing the contribution and value of leadership development, Avolio et al. make a case for working to refine the assumptions that feed into return on investment analysis. In other words, the complexity of HLDPs should not inhibit their evaluation, instead we need to find ways for evaluations to acknowledge and work with this complexity. Joseph-Richard and McCroy's (2022) review of evaluations of broader leadership development programmes identifies a series of qualitative evaluation models which they suggest could respond to the complexities inherent within a healthcare setting. This work is complemented by the current review of reviews by focusing on how leadership development in a healthcare setting has been evaluated in practice. Our understanding of the limitations of HLDP evaluation is developed as a result.

Theory-based evaluation (Stame, 2004; Brousselle and Buregeya, 2018) may offer a further way in which to challenge the 'black-box' (Scriven, 1999) character of programmes and the uncertainty as to the conditions that affect the outcomes achieved by HLDPs. Furthermore, our understanding of HLDPs and their evaluation would be progressed if we acknowledge the complexity of health systems and

how this affects the expression of leadership. The Medical Research Council's (Skivington *et al.*, 2021) *Framework for Developing and Evaluating Complex Interventions* offers a resource which could inform the evaluation of HLDPs. Westhorp (2012) identifies two approaches to the evaluation of complex systems, a realist approach and approaches which recognise complexity theory, and argues that these two approaches can be productively united within a single evaluation. There is the opportunity to consider how the evaluation of HLDPs can draw on these approaches and acknowledge the complexity inherent within healthcare leadership. While there is some direction and guidance as to how to approach the evaluation of HLDPs, we know relatively little about the preferences of those who conduct evaluations. To put it another way, why is it the case that designs informing the evaluation of HLDPs often do not take account of the complexity of leadership development? Practical considerations such as a lack of resource and time may encourage quick and dirty approaches to evaluation. Furthermore, post-hoc evaluation design may also prohibit the use of complexity informed approaches. Nevertheless, we know relatively little about how decisions are made during the development of evaluation designs. Exploring this topic in more detail through empirical research could help to inform guidance to address common barriers which can prohibit the use of more developed approaches to evaluation.

Limitations of the research

While the scoping review of reviews has identified gaps in our knowledge on HLDPs evaluation, these findings are weakened by the limitations of this review of reviews. The limitations of the incorporated articles will also be reflected within this scoping review of reviews. The content of a scoping review will be restricted by the boundaries of the included articles and the topics deemed to be of value within their discussions.

Several reviews noted that there was a risk of publication bias whereby positive findings may be more likely to be published. Only articles published in English were included in the review and this may have reduced the number of sources which could be called upon. Furthermore, searches of the grey literature did not return any sources and this suggests that the search terms were not sensitive enough to be used within these databases. Of note is the work of the Canadian Health Leadership Network (CHLNet) to produce a toolkit to inform the evaluation of HLDPs (CHLNet, 2017) which was not picked up when searching the grey literature. Similarly, the Leadership Development Evaluation and Research Group formed across the previous network of NHS leadership academies developed an evaluation

framework (NHS Leadership Academy, 2016). This framework was not found via the search. Other sources in non-academic publications may have been omitted by this review.

Directions for future research

Overall, this review of reviews has identified several areas of development, both practical and theoretical, which may progress the evaluation of HLDPs. A common theme across the included reviews is the suggestion that evaluations of HLDPs are often based on poor quality data. However, within the included articles, discussions on data quality appear to be limited as to whether data are subjective self-reports or objective assessments of behaviour change and associated organisational outcomes. The review of reviews has highlighted the complexity of leadership development in healthcare; however, complexity was not always recognised by the incorporated publications. Further reviews could explore how evaluation designs used to assess HLDPs have acknowledged the complexity of leadership in health systems.

Conclusion

The scoping review of reviews has considered reviews of the evaluation of HLDPs. Often, reviews were found to be descriptive in tone and there is greater scope to explore the theoretical underpinning that have informed the evaluation of HLDPs. The evaluation of HLDPs will not be progressed via an emphasis on methods alone. Clarity as to the theory of change that informs the HLDP offers a foundation on which to build the evaluation design. Future research should develop approaches to evaluation which account for the complexity of leadership in a health system.

References

Ahrari, A. *et al.* (2021) 'Leadership development programs for radiology residents: a literature review', C*anadian Association of Radiologists' Journal*, 72(4), pp. 669–677. Available at: https://journals.sagepub.com/doi/full/10.1177/0846537120986495 (Accessed: 9 November 2021).

Arksey, H. and O'Malley, L. (2005) 'Scoping studies: towards a methodological framework', *International Journal of Social Research Methodology*, 8(1), pp. 19–32. doi:10.1080 /1364557032000119616.

Avolio, B. J, Avey, J. B.and Quisenberry, D. (2010) 'Estimating return on leadership development investment', *The Leadership Quarterly,* 21(4), pp. 633-644. DOI: 10.1016/j.leaqua.2010.06.006

Bekas, S. (2015) 'Leadership development in UK medical training: Pedagogical theory and practice', *Teaching and Learning in Medicine*, 27(1), pp. 4–11. doi:10.1080/10401334.2014.979177.

Blumenthal, D.M. *et al.* (2012) 'Addressing the leadership gap in medicine: Residents' need for systematic leadership development training', *Academic Medicine*, 87(4), pp. 513–522. doi:10.1097/ACM.0b013e31824a0c47.

Braithwaite, J. *et al.* (2013) 'Health care as a complex adaptive system', in Hollnagel, E. *et al.* (eds) *Resilient Health Care*. Farnham: Taylor & Francis Group, pp. 57–73.

Brousselle, A. and Buregeya, A. (2019) 'Theory-based evaluations: Framing the existence of a new theory in evaluation and the rise of the 5th generation', *Evaluation*, 24(2), pp. 153–168. doi: 10.1177/1356389018765487.

Busari, J. (2012) 'Management and leadership development in healthcare and the challenges facing physician managers in clinical practice', *International Journal of Clinical Leadership*, 17, pp. 211–16.

Canadian Health Leadership Network (2017) *Development of a Return on Investment Tool for Healthcare Leadership Development. Evidence brief.* Available at https://chlnet.ca/wp-content/uploads/CHLNet-ROI-Evidence-Brief-Feb-2017.pdf (Accessed: 28 July 2022)

Careau, E. *et al.* (2014) 'Health leadership education programs, best practices, and impact on learners' knowledge, skills, attitudes, and behaviors and system change: A literature review', *Journal of Healthcare Leadership*, 2014, pp. 39–50. doi:10.2147/JHL.S61127.

Cummings, G. *et al.* (2008) 'Factors contributing to nursing leadership: A systematic review', *Journal of Health Services Research & Policy*, 13(4), pp. 240–248. doi:10.1258/jhsrp.2008.007154.

Curry, L.A., Ayedun, A.A., Cherlin, E.J., Allen, N. H. and Linnander, E. L. (2020) 'Leadership development in complex health systems: a qualitative study'. *BMJ Open*, Vol. 10, e035797. doi:10.1136/ bmjopen-2019-035797

Curtis, E.A., de Vries, J. and Sheerin, F.K. (2011) 'Developing leadership in nursing: exploring core factors', *British Journal of Nursing*, 20(5), pp. 306–309. doi:10.12968/bjon.2011.20.5.306.

De Brún, A., O'Donovan, R. and McAuliffe, E. (2019) 'Interventions to develop collectivistic leadership in healthcare settings: a systematic review', *BMC Health Services Research*, 19, 72. doi:10.1186/s12913-019-3883-x.

Edmonstone, J. (2011) 'Developing leaders and leadership in health care: a case for rebalancing?', *Leadership in Health Services,* 24(1), pp. 8–18. doi:10.1108/17511871111102490.

Edmonstone, J. (2013) 'Healthcare leadership: learning from evaluation', *Leadership in Health Services*, 26(2), pp. 148–158. doi:10.1108/17511871311319731.

Edmonstone, J. (2014) 'Whither the elephant: The continuing development of clinical leadership in the UK National Health Services', *The International Journal of Health Planning and Management*, 29(3), pp. 280–291. doi:10.1002/hpm.2245.

Flaig, J. *et al.* (2020) 'Examining how formal leadership development programs positively influence hospital leaders' individual competencies and organizational outcomes – an evidence-based literature review', *Journal of Healthcare Leadership*, 12, pp. 69–83. doi:10.2147/JHL.S239676.

Frich, J.C. *et al.* (2015) 'Leadership development programs for physicians: a systematic review', *Journal of General Internal Medicine,* 30(5), pp. 656–674. doi:10.1007/s11606-014-3141-1.

Geerts, J.M., Goodall, A.H. and Agius, S. (2020) 'Evidence-based leadership development for physicians: A systematic literature review', *Social Science & Medicine* (1982), 246, p. 112709. doi:10.1016/j.socscimed.2019.112709.

Husebø, S.E. and Akerjordet, K. (2016) 'Quantitative systematic review of multi-professional teamwork and leadership training to optimize patient outcomes in acute hospital settings', *Journal of Advanced Nursing*, 72(12), pp. 2980–3000. doi: 10.1111/jan.13035.

Jeyaraman, M.M. *et al.* (2018) 'Return on investment in healthcare leadership development programs', *Leadership in Health Services*, 31(1), pp. 77–97. doi:10.1108/LHS-02-2017-0005.

Johnson, O. *et al.* (2020) 'Interventions to strengthen the leadership capabilities of health professionals in Sub-Saharan Africa: a scoping review', *Health Policy and Planning*, 36(1), pp. 117–133. doi:10.1093/heapol/czaa078.

Joseph-Richard, P. and McCray, J. (2022): 'Evaluating leadership development in a changing world? Alternative models and approaches for healthcare organisations', *Human Resource Development International*. doi: 10.1080/13678868.2022.204308

Kirkpatrick, D.L. and Kirkpatrick, J.D. (2006) *Evaluating Training Programs: The Four Levels*. 3rd edition. San Francisco, CA: Berrett-Koehler Publishers.

Kumar, B., Swee, M.L. and Suneja, M. (2020) 'Leadership training programs in graduate medical education: a systematic review', *BMC Medical Education*, 20(175). doi:10.1186/s12909-020-02089-2.

Levac, D., Colquhoun, H. and O'Brien, K.K. (2010) 'Scoping studies: Advancing the methodology', *Implementation Science*, 5(69). doi:10.1186/1748-5908-5-69.

Lichtenstein, B.B. *et al.* (2006) 'Complexity leadership theory: An interactive perspective on leading in complex adaptive systems', *Emergence: Complexity & Organization*, 8(4), pp. 2–12.

Lyons, O. *et al.* (2021) 'Evidence-based medical leadership development: a systematic review', *BMJ Leader*, 5(3). doi:10.1136/leader-2020-000360.

MacDonald, R. (2014) 'Leadership and leadership development in healthcare settings – a simplistic solution to complex problems?', *International Journal of Health Policy and Management*, 3(4), pp. 227–229. doi:10.15171/ijhpm.2014.101.

Mianda, S. and Voce, A. (2018) 'Developing and evaluating clinical leadership interventions for frontline healthcare providers: a review of the literature', *BMC Health Services Research*, 18(747). doi:10.1186/s12913-018-3561-4.

NHS Leadership Academy (2016) *The Leadership Development Evaluation Framework.* Available at: https://www.leadershipacademy.nhs.uk/wpcontent/uploads/dlm_uploads/2018/12/Evaluation-Framework.pdf (Accessed 22 July 2022).

Onyura, B. *et al.* (2019) 'Is postgraduate leadership education a match for the wicked problems of health systems leadership? A critical systematic review', *Perspectives on Medical Education*, 8(3), pp. 133–142. doi:10.1007/s40037-019-0517-2.

Page, M.J. *et al.* (2021) 'The PRISMA 2020 statement: An updated guideline for reporting systematic reviews', *BMJ*, 372(71). doi:10.1136/bmj.n71.

Peters, M.D.J. *et al.* (2015) 'Guidance for conducting systematic scoping reviews', *International Journal of Evidence-Based Healthcare*, 13(3), pp. 141–146. doi:10.1097/XEB.0000000000000050.

Pizzirani, B. *et al.* (2020) 'Clinical leadership development in Australian healthcare: A systematic review', *Internal Medicine Journal*, 50(12), pp. 1451–1456. doi:10.1111/imj.14713.

Ravaghi, H. *et al.* (2021) 'Effectiveness of training and educational programs for hospital managers: A systematic review', *Health Services Management Research*, 34(2), pp. 113–126. doi:10.1177/0951484820971460.

Sadowski, B. *et al.* (2018) 'Leadership training in graduate medical education: A systematic review', *Journal of Graduate Medical Education,* 10(2), pp. 134–148. doi:10.4300/JGME-D-17-00194.1.

Scriven, M. (1999) 'The fine line between evaluation and explanation', *Research on Social Work Practice*, 9(4), pp. 521–524. doi: 10.1177/104973159900900407.

Seidman, G., Pascal, L. and McDonough, J. (2020) 'What benefits do healthcare organisations receive from leadership and management development programmes? A systematic review of the evidence', *BMJ Leader*, 4(1), pp. 21–36. doi:10.1136/leader-2019-000141.

Sfantou, D.F. *et al.* (2017) 'Importance of Leadership Style towards Quality of Care Measures in Healthcare Settings: A Systematic Review', *Healthcare*, 5(4), p. 73. doi:10.3390/healthcare5040073.

Skivington, K. et al. (2021) 'A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance', *BMJ*, 374: n2061. doi: 10.1136/bmj.n2061.

Sonnino, R.E. (2016) 'Health care leadership development and training: progress and pitfalls', *Journal of Healthcare Leadership*, 8, pp. 19–29. doi:10.2147/JHL.S68068.

Stame, N. (2004) 'Theory-based evaluation and types of complexity', *Evaluation*, 10(1), pp. 58-76. doi: 10.1177/1356389004043135.

Stoller, J.K. (2020) 'Reflections on leadership in the time of COVID-19', *BMJ Leader,* 4(2), pp. 77–79. doi:10.1136/leader-2020-000244.

Sultan, N. *et al.* (2019) 'Leadership development in postgraduate medical education: A systematic review of the literature', *Academic Medicine,* 94(3), pp. 440–449. doi:10.1097 /ACM.0000000000002503.

Torres-Landa, S. *et al.* (2021) 'Evaluation of leadership curricula in general surgery residency programs', *American Journal of Surgery,* 222(5), pp. 916-921. doi:10.1016/j.amjsurg.2021.05.012.

Turner, S. *et al.* (2018) 'Discipline-specific competency-based curricula for leadership learning in medical specialty training: A critical review of the literature', *Leadership in Health Services*, 31(2), pp. 152-166. doi:10.1108/LHS-08-2017-0048.

West, M. *et al.* (2015) Leadership and leadership development in health care. Available at: https://www.kingsfund.org.uk/publications/leadership-and-leadership-development-health-care (Accessed: 9 November 2021).

Westhorp, G. (2012) 'Using complexity-consistent theory for evaluating complex systems', *Evaluation*, 18(4), pp. 405–420. doi:10.1177/1356389012460963.

Figure 1: PRISMA flow diagram (Page *et al.*, 2021)

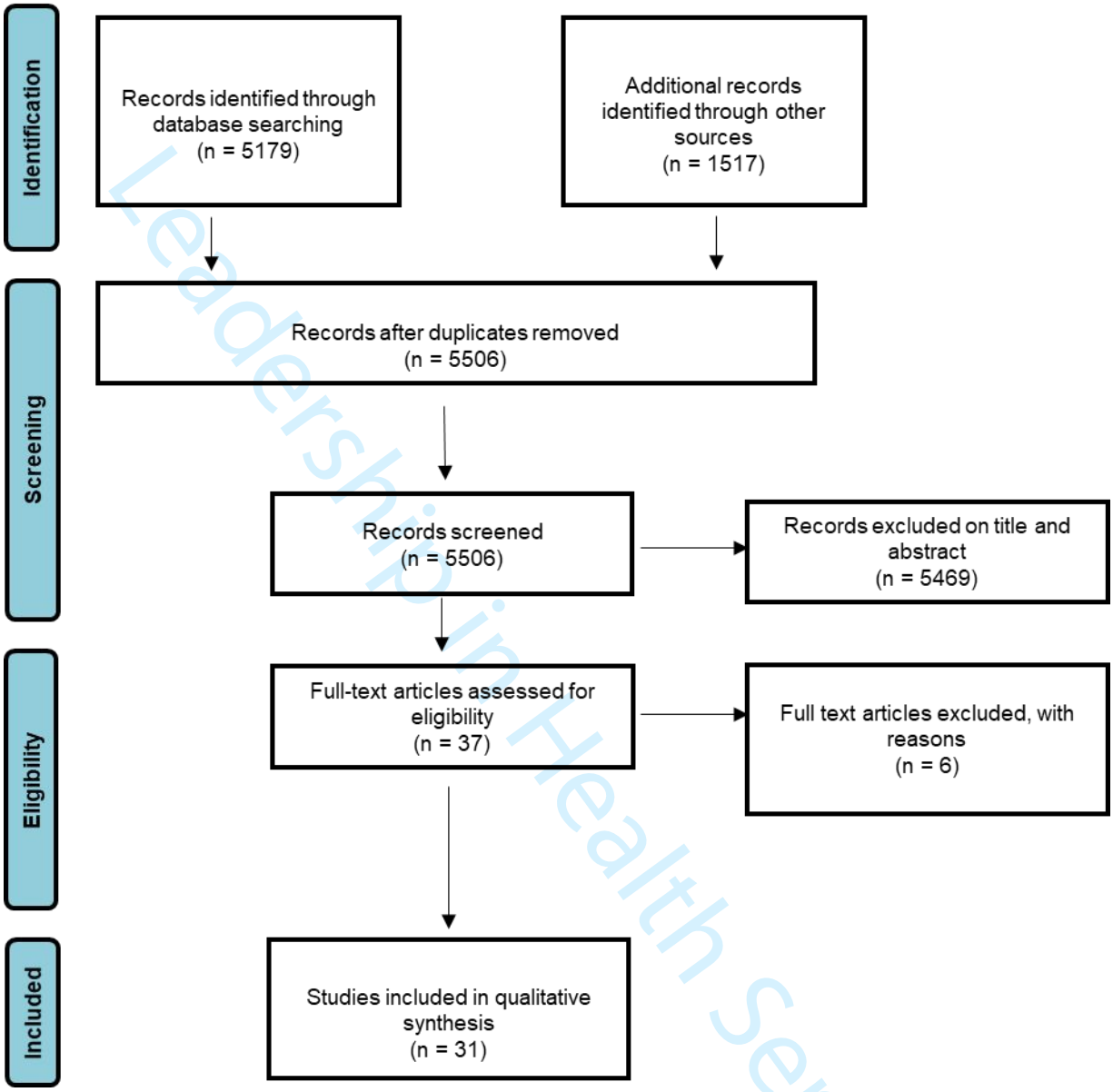

Table 1: Summary of included studies

First author and year of publication	Intervention targets	Aims	Number of included publications	Summary of findings
Ahrari *et al.* (2021)	Radiology residents	Describe leadership training programmes (structure, content, and evaluation) in radiology residencies.	6	Lack of published findings on HLDPs. Programmes vary in terms of structure, content and evaluation design.
Bekas (2015)	Postgraduate medical training	Explore leadership development in postgraduate medical training.	N/A	Debate how to develop leadership in medical education is ongoing. Experiential learning, reflective practice, action learning and mentoring appear to be most appropriate but evaluation of HLDPs can be limited.
Blumenthal *et al.* (2012)	Postgraduate medical training	Develop the argument for healthcare leadership development in postgraduate medical training. Explore barriers affecting healthcare leadership development.	N/A	Argues that leadership skills can be developed and identifies best practice for leadership development.
Busari *et al.* (2011)	Postgraduate medical training	Identify the skills and competencies required by physicians with a management role and assess whether developed curricula meet these needs. Explore the availability of development programmes.	40	Management education curricula evaluated positively, however studies have a low number of participants. There is little consensus as to the content and structure of programmes. Variety of evaluation designs were identified.

First author and year of publication	Intervention targets	Aims	Number of included publications	Summary of findings
Careau *et al.* (2014)	Health education leadership programmes (all levels of learner)	Explore the content and effectiveness of programmes teaching collaborative leadership.	250	Four models of leadership identified across programmes. Limited evaluation of effects of programmes on patient centred outcomes or systems change identified.
Cummings *et al.* (2008)	Leadership development in nursing	Explore the factors that inform nursing leadership and HLDPs targeting this group.	26	Studies that examined HLDPs reported positive outcomes. However, there is limited evidence of the specific factors influencing nursing leadership.
Curtis *et al.* (2011)	Leadership development in nursing	Explore the different forms of nursing leadership education.	N/A	Leadership development should be integrated within wider professional development.
De Brún *et al.* (2019)	Development of collectivistic leadership in healthcare settings	Explore evaluation designs and outcomes of programmes intended to develop collectivistic leadership.	21	Interventions indicated positive outcomes; however, the rigour of evaluations could be developed.
Edmonstone (2011)	N/A	Outlines a critique of healthcare leader development.	N/A	Distinguishes between leader development and leadership development. Leadership development in healthcare should place a greater emphasis on context and relationships.

First author and year of publication	Intervention targets	Aims	Number of included publications	Summary of findings
Edmonstone (2013)	Explores eight evaluations of leadership and management development programmes across the UK National Health Service	Identify areas of common learning in the design and evaluation of programmes.	8	Establishes the difficulty of evaluating HLDPs. Develops an evaluation framework for HLDPs.
Edmonstone (2014)	N/A	Update of Edmonstone (2009).	N/A	Limited conceptual clarity with regards to leadership. Distinguishes between the development of leaders and leadership.
Flaig *et al.* (2020)	Programmes targeted to leadership and management in a hospital setting	Explore the impact of hospital-based HLDPs on competencies and organisational outcomes.	23	HLDPs do demonstrate positive outcomes however further research is needed to explore the effects of HLDPs at an organisational level.
Frich *et al.* (2015)	Physician HLDPs	Explore physician leadership development programmes to describe the structure, content and learning outcomes achieved.	45	Positive outcomes were reported but often relied on self-reported measures. Limited number of studies explored organisational outcomes.

First author and year of publication	Intervention targets	Aims	Number of included publications	Summary of findings
Geerts *et al.* (2020)	Physician leadership development	Explore the elements of HLDPs associated with positive outcomes at the individual and organisational level.	25	Identifies positive outcomes at both the individual and organisational level. Notes that baseline and retrospective objective outcome data should be collected.
Husebø and Akerjordet (2016)	Multi-professional teamwork and leadership training in acute hospital settings	Identify the components of multi-professional teamwork and HLDPs that impact patient outcomes in hospitals.	12	Impact on patient outcomes was identified, however the effect of HLDPs is uncertain. Design and evaluation of programmes should be improved.
Jeyaraman *et al.* (2018)	Leadership development interventions targeted to healthcare professionals.	Identify evidence on healthcare outcomes and return on investment indicators associated with leadership quality, and leadership development programmes.	223	Summarises the range of return-on-investment metrics, classified as patient-oriented, staff-oriented, or organisational outcomes. Leadership development programmes varied widely. While enhanced leadership skills and improved outcomes were identified, these were often self-reported.
Johnson *et al.* (2021)	Interventions to strengthen the leadership capabilities of	Explore leadership development in low- and middle-income countries.	28	Identified HLDPs were diverse and lacked consistency in the conceptualisation of

First author and year of publication	Intervention targets	Aims	Number of included publications	Summary of findings
	healthcare professionals in Sub-Saharan Africa			leadership. Evaluation designs often found to be poor quality.
Kumar *et al.* (2020)	Post-graduate medical education	Identify the characteristics and outcomes of medical education HLDPs.	15	Evaluation designs often lacked methodological rigour. However, studies that used pre and post surveys demonstrated improvement in self-perceived knowledge and leadership skills.
Lyons *et al.* (2020)	HLDPs targeted to physicians (practising doctors post-qualification)	Explore the content and structure of HLDPs associated with physician leadership development.	117	Pedagogical methods were viewed to have a greater effect on outcomes than programme content. However, there remains uncertainty surrounding the impact of HLDPs and there is scope to improve evaluation designs.
Mianda and Voce (2018)	HLDPs targeted to frontline healthcare workers	Describe the content and evaluation HLDPs. Aim to inform clinical leadership development in low- and middle-income countries, particularly maternal and perinatal care in South Africa.	24	Identifies a gap in the literature on leadership development in lower and middle-income settings. Evaluation of expected outcomes should be planned during programme development. Pre- and post-measures and 360-degree

First author and year of publication	Intervention targets	Aims	Number of included publications	Summary of findings
				feedback are preferred approaches for evaluation.
Onyura *et al.* (2019)	Postgraduate medical education	Synthesise evidence on the characteristics and evaluation of HLDPs.	31	Postgraduate leadership education is often focused on developing individual competencies. Some evidence for the generation of positive individual-level outcomes.
Pizzirani *et al.* (2019)	Healthcare professionals in Australia	Explore the content, structure and evaluation of healthcare HLDPs in Australia.	4	Limited number of studies exploring healthcare leadership development in Australia. Limited use of theoretical underpinning in programme design. Evaluation of programmes are often of poor quality.
Ravaghi *et al.* (2020)	Hospital managers and administrators	Explore the impact of training programmes for hospital managers and administrators.	9	Programmes identified to have a moderately positive effect on the skills, knowledge, and competencies of hospital managers and administrators.
Sadowski *et al.* (2018)	Postgraduate medical education	Identify areas of commonality, impact and gaps in postgraduate medical education leadership.	52	Programmes are often varied in content and structure. Studies were often low-quality. There is scope for the development of longitudinal evaluations.

First author and year of publication	Intervention targets	Aims	Number of included publications	Summary of findings
Seidman *et al.* (2020)	Healthcare professionals	Identify forms of evidence demonstrating benefit of HLDPs to healthcare organisations.	55	Identifies four categories of organisational outcomes. However, the evidence-base for these findings is limited and future studies should use more rigorous methods.
Sonnino (2016)	Healthcare professionals	Explore healthcare leadership development and identify the benefits and obstacles of implementing HLDPs.	N/A	There is little uniformity with regards to HLDPs' content and structure.
Stoller (2020)	HLDPs targeted to physicians (practising doctors post-qualification)	Review the impact of HLDPs.	3	Notes that there is a lack of robust evaluation of HLDPs.
Sultan *et al.* (2019)	Postgraduate medical trainees	Describe leadership development interventions within postgraduate medical education and aim to assess the effectiveness of interventions.	21	HLDPs often lack theoretical grounding and evaluation designs are often poor.
Torres-Landa *et al.* (2021)	General surgery residents	Describe the structure and content of general surgery residency leadership programmes.	7	A variety of evaluation plans were identified. There was no single definition of leadership across the studies. Evaluation should be planned within programme design and should have clear, measurable outcomes.

First author and year of publication	Intervention targets	Aims	Number of included publications	Summary of findings
Turner *et al.* (2018)	Postgraduate medical trainees	Describe the literature on leadership development incorporated into postgraduate medical specialty training.	7	There is limited specialty specific guidance on leadership development.
West *et al.* (2015)	Leadership in healthcare	Summarise the evidence base on leadership and leadership development.	N/A	Summarises leadership theory and research. Discusses the challenge of evaluating leader development programmes.

Supplemental table 1: Search strategies

Database	Search strategy
Medline	exp health care AND exp leadership OR leader$.ti,ab. OR health services manage$.ti,ab. OR capabilit$.ti,ab. AND exp program OR exp training program OR program$.ti,ab. OR leader$ develop$.ti,ab. OR exp staff development OR training.ti,ab. OR exp management training AND evaluation$.ti,ab. OR audit$.ti,ab. OR exp program evaluation OR exp systematic reviews/ OR review$.ti,ab. OR synthesis$.ti,ab. OR exp review/
EMBASE	exp health care AND exp leadership skills OR exp leadership OR exp health service management OR exp management role OR leader$.ti,ab. OR health service manage$.ti,ab. OR capabilit$.ti,ab. AND exp professional development OR exp management OR exp development OR exp programmes OR exp educational programmes OR exp training OR exp training programmes OR program$.ti,ab. OR leader$ develop$.ti,ab. OR training.ti,ab. OR exp management training AND exp evaluation OR exp assessment methods OR exp evaluation methods OR exp audit OR evaluation$.ti,ab. OR audit$.ti,ab. OR exp effectiveness AND exp literature reviews OR exp systematic reviews OR review$.ti,ab. OR synthesis$.ti,ab.
PsycINFO	exp health care AND exp leadership skills OR exp leadership OR exp health service management OR exp management role OR leader$.ti,ab. OR health service manage$.ti,ab. OR capabilit$.ti,ab. AND exp professional development OR exp management OR exp development OR exp programmes OR exp educational programmes OR exp training OR exp training programmes OR program$.ti,ab. OR leader$ develop$.ti,ab. OR training.ti,ab. OR exp management training AND exp evaluation OR exp assessment methods OR exp evaluation methods OR exp audit OR evaluation$.ti,ab. OR audit$.ti,ab. OR exp effectiveness OR exp literature reviews OR exp systematic reviews OR review$.ti,ab. OR synthesis$.ti,ab.
Health management and policy database	exp health care AND exp leadership skills/ OR exp leadership/ OR exp health service management/ OR exp management role/ OR leader$.ti,ab OR manage$.ti.ab. OR health service manage$.ti,ab. AND capabilit$.ti,ab. OR exp professional development/ OR exp management/ OR exp development/ OR exp programmes/ OR exp educational programmes/ OR exp training/ OR exp training programmes/ OR program$ti,ab. OR leader$ development. ti,ab. OR training.ti,ab. OR education.ti,ab. OR exp management training/ AND exp evaluation/ OR exp assessment methods/ OR exp evaluation methods/ OR exp audit/ OR evaluation$.ti,ab. OR audit$.ti,ab. OR exp effectiveness/ OR assess$.ti,ab. OR effective$.ti,ab. AND exp literature reviews/OR exp systematic reviews/ OR review$.ti,ab. OR synthesis$.ti,ab.