




BMJ Open Quality **Developing a tool to measure enactment of complex quality improvement interventions in healthcare**

Lauren MacEachern ¹, Liane R Ginsburg,² Matthias Hoben ^{3,4}, Malcolm Doupe,^{5,6} Adrian Wagg,⁷ Jennifer A Knopp-Sihota,⁸ Lisa Cranley,⁹ Yuting Song ^{4,10}, Carole A Estabrooks,⁴ Whitney Berta¹

To cite: MacEachern L, Ginsburg LR, Hoben M, *et al.* Developing a tool to measure enactment of complex quality improvement interventions in healthcare. *BMJ Open Quality* 2023;**12**:e002027. doi:10.1136/bmjopen-2022-002027

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2022-002027>).

Received 27 June 2022

Accepted 24 January 2023

ABSTRACT

Quality improvement (QI) projects are common in healthcare settings and often involve interdisciplinary teams working together towards a common goal. Many interventions and programmes have been introduced through research to convey QI skills and knowledge to healthcare workers, however, a few studies have attempted to differentiate between what individuals 'learn' or 'know' versus their capacity to apply their learnings in complex healthcare settings. Understanding and differentiating between delivery, receipt, and enactment of QI skills and knowledge is important because while enactment alone does not guarantee desired QI outcomes, it might be reasonably assumed that 'better enactment' is likely to lead to better outcomes. This paper describes the development, application and validation of a tool to measure enactment of core QI skills and knowledge of a complex QI intervention in a healthcare setting. Based on the Institute for Healthcare Improvement's Model for Improvement, existing QI assessment tools, literature on enactment fidelity and our research protocols, 10 indicators related to core QI skills and knowledge were determined. Definitions and assessment criteria were tested and refined in five iterative cycles. Qualitative data from four QI teams in long-term care homes were used to test and validate the tool. The final measurement tool contains 10 QI indicators and a five-point scale. Inter-rater reliability ranged from good to excellent. Usability and acceptability among raters were considered high. This measurement tool assists in identifying strengths and weaknesses of a QI team and allows for targeted feedback on core QI components. The indicators developed in our tool and the approach to tool development may be useful in other health related contexts where similar data are collected.

BACKGROUND

Quality improvement (QI) projects in healthcare settings are common and often require interdisciplinary teams to work cohesively towards a common goal. The Institute for Healthcare Improvement (IHI) has championed QI work for several decades and its foundational concepts have, often with adaptation, been widely adopted.^{1–6} Educational programmes, or interventions with educational components, are used to convey QI

concepts and the general skills and knowledge required to conduct QI activities (eg, development of clear aim statements, trialling change ideas using plan–do–study–act (PDSA) cycles).^{4–7} While QI knowledge comprehension of intervention participants is often assessed (eg, through participant testing at the conclusion of the intervention), few studies differentiate between what individuals learn or know versus their capacity to actually apply their learnt QI skills and knowledge in complex healthcare settings.^{8,9} The actual application of QI skills and knowledge to achieve meaningful practice change is highly dependent on contextual factors of the environment, particularly the supportive capacity of leadership, team dynamics and organisational culture.¹⁰

This distinction between knowledge and application aligns with the literature on implementation fidelity that differentiates among the delivery, receipt and enactment of QI skills and knowledge.^{11,12} Delivery refers to how the skills and knowledge are presented or made available and includes the mode, format, intensity and frequency of delivery to healthcare providers.^{13,14} Receipt of the intervention is described as providers' comprehension and mastery of the skills and knowledge delivered through training initiatives or other means.^{13,15} Enactment refers to the extent to which these skills and knowledge are observably applied or demonstrated by healthcare providers in the intended practice settings.¹³ Delivery and receipt are frequently assessed during or following educational interventions using self-reporting or participant observation.^{9,16} However, enactment is less frequently assessed and is arguably a better indicator of intervention effectiveness and likelihood of achieving the intended practice change.^{11,17}

On a trajectory from QI intervention delivery to realised outcomes (figure 1), enactment takes place after delivery and



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Lauren MacEachern;
lauren.maceachern@mail.utoronto.ca

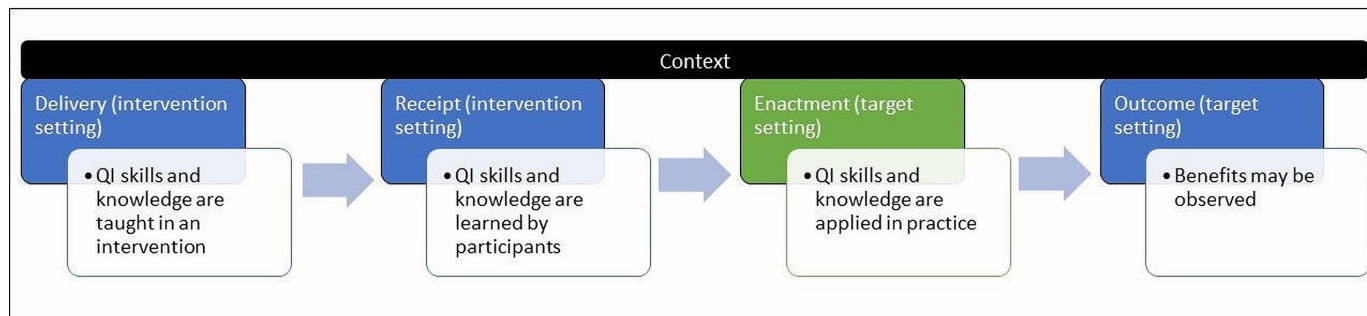


Figure 1 Trajectory from intervention delivery to realised outcomes. QI, quality improvement.

receipt but likely prior to realising outcomes where the benefits and potentially unintended consequences of the QI work become apparent.⁹ Enactment alone does not guarantee the achievement of desired QI intervention outcomes, as contextual factors (eg, poorly resourced work environments, emerging crises like influenza outbreaks) are known to influence the successful implementation and sustainability of projects.^{7 10 18 19} However, research suggests that ‘better enactment’ is likely to lead to (better) outcomes or QI success.²⁰ There is a growing body of work demonstrating that QI teams with observably higher levels of enactment achieve significantly greater outcome improvement when compared with teams with lower levels of enactment. Teams with higher enactment are also more likely to sustain these improvements over time.²¹

Purpose

The purpose of this paper is to describe the development and application of a tool for measuring the enactment of core QI components of a complex QI intervention. The tool was developed for Canadian long-term care homes. We include descriptions of the challenges encountered during the tool development and application processes, and outline the mitigation approaches that we took. We offer recommendations for other researchers who are considering adapting our tool for use in other QI interventions.

Situating the tool development and application: Safer Care for Older People in Residential Environments and Sustainment, Sustainability and Spread Study

The Sustainment, Sustainability and Spread Study (SSaSSy)²² is a postimplementation study of an evidence-informed, complex QI intervention called SCOPE (Safer Care for Older People in Residential Environments; ClinicalTrials.gov NCT03426072). The SCOPE intervention is education/facilitation-based and designed on IHI’s Breakthrough Series Collaborative model.^{3 23} In brief, the SCOPE study focused on developing the knowledge and skills of healthcare aides (HCAs) in long-term care homes to lead QI projects.^{23 24} In Canadian long-term care homes, HCAs provide 75%–80% of hands-on care for activities of daily living (eg, toileting, bathing).^{25 26} Thirty-nine long-term care homes from Western Canada were randomly selected from a longitudinal cohort and participated in

the SCOPE intervention between Spring 2017 and Spring 2019; each with one unit-based QI team.²⁷ Within each long-term care home, QI teams learnt and applied QI concepts over a 1-year intervention period. QI teams were led by HCAs and were supported by an external QI expert (QI Advisor) and internally by care home managers (unit managers and directors of care). Two years after SCOPE concluded, seven long-term care homes from Manitoba were invited to participate in SSaSSy.²²

In SSaSSy, we aim to understand the minimum support necessary to reinvigorate and sustain QI interventions over the longer term, postimplementation.²² SSaSSy is phasic and ongoing: we began working with QI teams in Manitoba in 2019, and are now implementing SSaSSy in long-term care homes in Alberta and British Columbia, Canada.²² In the first phase of SSaSSy, researchers ascertained the extent to which SCOPE QI skills and knowledge were sustained since SCOPE concluded in 2017.²⁸ Participating QI teams were then assigned to one of two ‘Booster’ conditions: a Low-Booster or High-Booster (see table 1).²² Both boosters included educational sessions and educational materials based on IHI’s Model for Improvement,³ as well as varying levels of external support from an experienced QI Advisor.

METHODS

Development of the measurement tool

We developed a tool to measure and understand any differentiable effects of the SSaSSy boosters on the Manitoba QI teams’ enactment of core QI processes and behaviours. This tool permitted us to measure participating QI team’s enactment of the core QI processes and behaviours originally conveyed by the SCOPE intervention and revisited with the SSaSSy QI teams (see table 1).

We developed this tool based on principles from: (1) IHI Model for Improvement,³ (2) existing QI assessment tools,^{29–31} (3) our research protocols^{22 23} and (4) literature on enactment fidelity.^{11 12 17} Table 2 summarises the SCOPE core processes and behaviours, the SSaSSy booster components designed to recall or reinforce them, and corresponding measurement indicators.

In an iterative process, we developed and refined a set of enactment indicators and their definitions based on QI competencies and core intervention components. Development discussions included nine researchers with

Table 1 SSaSSy booster components and description

High booster	Low booster
HCA-led team (QI team led by HCAs, supported by internal sponsors)	
Team and senior sponsors (Internal sponsors at unit and facility management levels)	
In-person Refresher Workshop (Half day workshop focusing on QI skills and processes, leadership skills, team-based approaches for decision-making)	
Educational binder (Resources on QI skills and processes, QI data collection and reporting templates, topic-specific resources on three clinical areas (responsive behaviours, pain, mobility))	
External QI advisor support for teams	
<ul style="list-style-type: none"> ▶ Three times after refresher workshop, at midpoint and prior to wrap up meeting (in-person) ▶ Continuous, on-demand support 	<ul style="list-style-type: none"> ▶ One time at midpoint of intervention period (in-person) ▶ Continuous, on-demand support
External QI advisor support for sponsors	
<ul style="list-style-type: none"> ▶ Three times after refresher workshop, at midpoint, and prior to wrap up meeting (teleconference) ▶ Continuous, on-demand support 	<ul style="list-style-type: none"> ▶ One time at midpoint of intervention period (teleconference) ▶ Continuous, on-demand support
HCA, healthcare aide; QI, quality improvement; SSaSSy, Sustainment, Sustainability and Spread Study.	

extensive practical and theoretical experience, several of whom had developed the original SCOPE intervention and two QI advisors with in-depth knowledge of the

Table 2 Situating scope and SSaSSy

SCOPE core processes/behaviours	SSaSSy booster component	Enactment evaluation indicator
<ul style="list-style-type: none"> ▶ HCA-led QI teams ▶ Adopting team-based approaches 	<ul style="list-style-type: none"> ▶ Refresher Workshop ▶ HCA-led QI team ▶ Ongoing QI advisor support 	<ol style="list-style-type: none"> 1. HCA empowerment 2. Leadership (ie, HCA-led) 3. Functioning as a team
<ul style="list-style-type: none"> ▶ Internal support of QI teams from sponsors (mentoring and coaching) 	<ul style="list-style-type: none"> ▶ Team sponsor/senior sponsor 	<ol style="list-style-type: none"> 1. Appropriate sponsor support
<ul style="list-style-type: none"> ▶ Learn QI skills and processes ▶ Apply QI skills and processes 	<ul style="list-style-type: none"> ▶ Refresher workshop ▶ Education binder ▶ Ongoing QI advisor support 	<ol style="list-style-type: none"> 1. PDSA cycles 2. Aim statements 3. Change ideas 4. Measurement 5. Engagement with best available evidence to inform QI project 6. Achievement of their stated aim
<ul style="list-style-type: none"> ▶ External support of QI teams and sponsors from QI advisors (coaching and guiding) 	<ul style="list-style-type: none"> ▶ Refresher workshop ▶ Ongoing QI Advisor support 	
HCA, healthcare aide; PDSA, plan-do-study-act; QI, quality improvement; SSaSSy, Sustainment, Sustainability and Spread Study.		

intervention. A detailed description of the evolution of the tool is included in online supplemental appendix A.

Validation of the measurement tool

Separately, a group of three and a group of four researchers each tested the tool following refinement iterations 2 and 3, respectively (see online supplemental appendix A). Data from two SSaSSy QI teams were used including: worksheets completed by the SSaSSy QI teams to document progress on their QI projects during the intervention, focus groups conducted with SSaSSy QI teams at the end of intervention period, meeting minutes from support calls between QI team Sponsors and SSaSSy QI Advisors, and diaries kept by SSaSSy QI Advisors regarding intervention enactment and barriers to enactment. Data collection activities for these data were developed as part of our research protocol, however, similar sources may be collected regularly as part of a QI project. See online supplemental appendix B for a detailed description of these data sources and ideas for similar sources. Group meetings were then held to discuss challenges encountered in applying the tool and necessary revisions to the tool, including refinement of the assessment scale, indicator definitions and clarification of assessment categories.

Because we did not have a large enough sample to apply the tool to new QI teams, we allowed sufficient time to pass (approximately 4 months) before five researchers with the most intimate knowledge of the core components of SCOPE/SSaSSy reviewed all the data for the 4 SSaSSy QI teams and applied the tool following refinement iteration 4. An analysis meeting was then held to discuss the researchers' independent assessments, and to discuss ratings where there were discrepancies. Based on these discussions, researchers were permitted to modify their ratings if a new interpretation was understood but consensus was not required. Scores from this meeting were used to conduct the inter-rater reliability assessment using the intraclass correlation coefficient (ICC) (see table 3). The ICC calculation is the most common calculation used for ordinal data and the score indicates how well the raters agreed on their ratings; with a higher score indicating higher estimated reliability.^{32 33} A two-way, mixed model approach was used for the ICC calculation in SPSS (V.28.0.0.0). The following ICC benchmarks were used for interpretation: poor <0.4; fair 0.4–0.59; good 0.6–0.74; excellent ≥ 0.75 .³²

RESULTS

The final indicators included in the measurement tool relate to enactment of: (1) PDSA cycles, (2) Aim statements, (3) Change ideas, (4) Measurement, (5) HCA empowerment*, (6) Engagement with best available evidence to inform QI project*, (7) Achievement of stated aim, (8) Appropriate Sponsor support*, (9) HCA Leadership* and (10) Functioning as a team. All indicators are theoretically informed and reflect critical components of

Table 3 ICC (total score and individual indicator)

Indicator	ICC (95% CI)
Total score	0.981 (0.920 to 0.999)
PDSA cycles	0.941 (0.761 to 0.996)
Aim statements	-0.714 (-18.926 to 0.895)
Change ideas	0.778 (0.129 to 0.984)
Measurement	0.720 (0.128 to 0.978)
HCA empowerment	0.947 (0.764 to 0.996)
Engagement with best available evidence to inform QI project	n/a*
Achievements of stated aim	0.909 (0.629 to 0.994)
Appropriate Sponsor support	0.985 (0.940 to 0.999)
HCA leadership	0.970 (0.865 to 0.998)
Functioning as a team	0.987 (0.946 to 0.999)

*ICC was not calculated due to missing data from rating exercise. HCA, healthcare aide; ICC, intraclass correlation coefficient; n/a, not available; PDSA, plan-do-study-act; QI, quality improvement.

QI work and those specific to SCOPE/SSaSSy are noted with an asterisk.^{22 23} Online supplemental appendix C describes the indicators and differentiates between those that were derived from the QI literature and those that are intervention-specific. The final assessment scale included both numbers and categories, with High enactment including scores of 5 or 4, Medium enactment including a score of 3, low/no enactment including scores of 2 or 1. The category 'n/a' was used when raters were 'unable to rate' due to insufficient information. The final measurement tool is presented as online supplemental table 3 to accommodate publication requirements.

We examined inter-rater reliability by calculating the ICC of each variable, as well as total score for each data set. The ICC on the total score across all data sets was 0.981, and the ICC for individual variables ranged from 0.720 to 0.987 (see table 3), apart from Aim Statement and Engagement with Evidence which are described separately below. Seven of the ten individual variables yielded an ICC of 0.75 or greater, indicating excellent agreement among raters. Of those seven indicators, six yielded a CI of 0.6 to 0.99, indicating that 95% of all samples will have an ICC of good or excellent estimated reliability.

The ICC for Aim Statement produced a negative number (ICC=-0.714; 95% CI -18.926 to 0.895), which can occur in ICC models when there is very low variance between raters and small sample size.³⁴ Online supplemental appendix D provides raw data for Aim Statement and each other indicator. For Engagement with Evidence, three of five raters were unable to rate at least one SSaSSy QI team using existing data sources. In these cases, raters reported that 'not enough information' was available to determine if evidence was used to inform the QI projects. Due to the missing data, the ICC was not calculated for this variable.

DISCUSSION

This tool presents a promising approach to measure the enactment of core, evidence-based QI activities introduced through complex interventions to QI teams. The inclusion of enactment evaluation criteria based on the IHI's Model for Improvement,³ existing QI assessment tools, and our research protocol makes this a useful tool for measuring enactment of QI components across a variety of projects. Most variables demonstrated good to excellent inter-rater reliability,³² with the exception of Aim Statement. Indicators were designed to be mutually exclusive and may be used as a stand-alone assessment of enactment for each specific QI component. As is mentioned in development work for similar tools,³⁰ the team-based nature of SSaSSy QI teams does not allow us to assess a team member's unique contributions or engagement with their QI project during the intervention period. However, our tool does acknowledge the different roles of individual team members in the intervention, including Sponsors and HCAs, and assesses their enactment of role-specific attributes (ie, Sponsor support, HCA Leadership). The tool was useful for identifying specific strengths and weaknesses of individual QI teams, as well as affording an overall sense of enactment with respect to intervention-specific criteria based on QI practices and principles.³⁰ In the future, this insight can be used by QI advisors in subsequent stages of the SSaSSy study or by coaches, managers, and facility administrators to guide and support teams engaged in similar QI initiatives.

Our approach to developing and refining our QI enactment tool was similar to that of comparable QI scoring rubrics.^{30 31} The iterative process of trialling and revising the language and criteria of the tool allowed for the natural assessment of usability and acceptability among raters,³⁰ which was considered high. Raters who applied our tool had a high level of experience and familiarity with the SCOPE/SSaSSy intervention but varied experience with the individual projects of SSaSSy QI teams. While our study did not pursue an assessment of external testing with a general population of raters, this is something that may be examined in future studies.

When engaging a general population of raters, standardised training around the SCOPE/SSaSSy intervention and QI principles generally would be required, including opportunities for discussion, familiarisation with concepts and practice examples.^{30 35 36} In such cases, inter-rater reliability may be more appropriately assessed using Cohen κ , which corrects for the chance agreement that may occur with a group of raters with lower training or familiarity.³⁵ While raters who applied our measurement tool were familiar with the SCOPE/SSaSSy intervention, there were challenges experienced with natural rating tendencies (eg, dove vs hawk). Future undertakings will incorporate a preassessment and calibrating exercise for all raters, including worked examples or vignettes that allows raters, regardless of their level of familiarity with the intervention, to identify and discuss their own natural biases.

The influence of context

Our tool was developed to isolate the enactment of specific intervention components, allowing for assessment and then targeted feedback or support for teams in particular QI areas. However, through our assessment, it became clear that assessing enactment is complicated by the fact that enactment is a function not only of delivery and receipt, but also of the real-world healthcare context in which knowledge is applied. Context is defined as factors that may influence responses¹⁹ and such factors have been described with respect to QI work as characteristics of the organisational setting, the individual, and their role in the organisation or QI project.^{18 19 37} Since enactment is susceptible to contextual factors that may affect outcomes, when evaluating QI projects, it is important to consider both enactment and context simultaneously to differentiate between what individuals are capable of doing and what they actually do within the constraints of their setting.¹⁸

In a systematic review by Kaplan *et al.*,¹⁰ contextual factors influencing QI success in healthcare organisations were explored. By their definition, QI success included the ‘extent of implementation of QI practices’,¹⁰ which may be considered closely related to our conceptualisation of enactment. Factors related to leadership from top management, strong team leadership, relationships and group climate have a positive association with measures of QI success.^{10 38} These factors are comparable to the Appropriate Sponsor Support and Functioning as a Team indicators.

Understanding the influence of context on enactment of QI practices is further complicated when considering factors that are both introduced and targeted by the intervention itself.^{9 18 39} Previous research on the SCOPE intervention by members of this research team have referred to this as a modifiable attribute of context.³⁸ For example, naming a senior and team sponsor was a requirement of the intervention.^{22 23} The leadership and supportive capacity of the sponsors are contextual factors that impact the enactment of QI skills and knowledge by QI team members. However, the SSaSSy intervention did not accept context but endeavoured to modify it by offering QI Advisor support for Sponsors to enhance their supportive capacity. To understand the role of context in QI work, Kaplan and Walsh¹⁸ recommend classifying influential factors as being part of the context, intervention or implementation. Our work demonstrates the importance of also identifying and targeting modifiable aspects of context in intervention design.³⁸ For the purpose of measuring enactment, our tool focused on specific activities related to the intervention. However, we recognise that relationships likely exist both between indicators and across the classifications of context, intervention and implementation.

Data sources

To apply a tool of this nature, process data describing QI activities must be collected. Our study used several

qualitative data sources (see online supplemental appendix B) to triangulate our findings whenever possible. We acknowledge that these data were not originally collected with the intent of measuring enactment; this presented us with challenges in data analysis.^{40–43}

Raters experienced instances where additional information was required/desired or when evidence from different sources was conflicting.⁴⁴ An example of this is the Engagement with Evidence indicator, where raters were unable to determine if the SSaSSy QI teams used evidence to inform the QI projects based on the data provided. Going forward, we have incorporated a more focused query on the use of research evidence by SSaSSy QI teams, on the part of QI advisors and members of the research team alike.

The group conversations between raters, the first author and project PI that facilitated the development of these guidelines were considered critical to capture the complexity of the measurement process. Given that these are complex interventions that have elicited complex responses from our participants, we feel that they merit detailed consideration and thus we intend to continue our collaborative approach to measuring enactment in the next phases of the SSaSSy study.

Future work and limitations

This paper presents preliminary reliability evidence for our enactment measurement tool for QI projects implemented in long-term care homes. One limitation was the small sample size used in testing our tool, which necessitated the reuse of data sets in the final testing stage. While researchers allowed a 4-month period to elapse between analysis of the data sets, it is possible that some memory of previous test cycles was retained. Future work by this team will include further testing with larger sample sizes.

This study was conducted in long-term care homes, however, we feel strongly that the indicators in our tool and approach to tool development may be useful in other health related contexts where similar data are collected. Similar data sources may include QI tracking sheets and documentation collected throughout a QI project, notes from QI team meetings, and both formal and informal conversations between an evaluator and the QI team members to reflect ‘facts’ and ‘impressions’. By focusing on the enactment of QI activities in situ, our tool does not require prospective data collection. Our tool differs from most existing tools in the literature, which have largely included written or oral evaluation components to assess the receipt of skills and knowledge by participants at the conclusion of educational interventions.^{29–31 35 36 45}

This tool was developed based on the assumption of a positive association between enactment and QI success. Existing research in the intersections of QI, healthcare and context, have generated valuable hypotheses for the relationships that may exist between the indicators included in our tool and QI success.^{10 39} Going forward,



we intend to explore the relationships between our assessment of QI teams' enactment of core QI processes and behaviours, and the outcomes that the teams ultimately achieve.

CONCLUSION

This study demonstrates the development and application of a tool to assess the enactment of core, evidence-based QI activities by QI teams. Our focus on QI team enactment and consideration for context-sensitive indicators stands to enhance the measurement of improvement practices in healthcare settings. This research reinforces the suggestion that measuring engagement in QI methods is challenging⁴¹ and that many components may contribute to the enactment of QI knowledge and skills. However, the tool was designed to assess each indicator individually, thus allowing for targeted feedback on specific QI activities. Our tool demonstrated a high level of inter-rater reliability across most indicators and may be suitable for use in QI projects where process data have been collected, thus eliminating the need for traditional knowledge tests for delivery and receipt.

Author affiliations

¹Institute for Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada

²Health Policy & Management, York University, Toronto, Ontario, Canada

³School of Health Policy and Management, York University, Toronto, Ontario, Canada

⁴Faculty of Nursing, University of Alberta, Edmonton, Alberta, Canada

⁵Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, Manitoba, Canada

⁶Centre for Care Research, Western Norway University of Applied Sciences, Bergen, Norway

⁷Department of Medicine, University of Alberta, Edmonton, Alberta, Canada

⁸Faculty of Health Disciplines, Athabasca University, Athabasca, Alberta, Canada

⁹Lawrence S Bloomberg Faculty of Nursing, University of Toronto, Toronto, Ontario, Canada

¹⁰School of Nursing, Qingdao University, Edmonton, Shandong, China

Twitter Lauren MacEachern @lmaceachern

Contributors LM, WB, LRG, MH, MD, AW, JAK-S, LC and YS contributed substantially to the conception and design of the tool. JAK-S, YS, LC and LRG were involved in testing the tool. LRG, AW, MD, MH and LC were involved in the application and testing of the tool. LM and WB contributed extensively to drafting the manuscript. All authors contributed extensive feedback to multiple drafts of the manuscript. All authors read and approved the final manuscript.

Funding Funding for this study has been provided by the Canadian Institutes for Health Research, Grant # 201803PJT-400653-KTR-CEAA-107836.

Disclaimer The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by University of Toronto Research Ethics Board Protocol #:18364, University of Alberta Research Ethics Board Study ID Pro00098546. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been

peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Lauren MacEachern <http://orcid.org/0000-0001-8316-1878>

Matthias Hoben <http://orcid.org/0000-0003-3465-315X>

Yuting Song <http://orcid.org/0000-0001-6597-0567>

REFERENCES

- Bourn SS, Mackey KE, Redlener MD. Measuring, and improving quality. In: Cone DC, Brice JH, Delbridge TR, et al, eds. *Emergency medical services*. Hoboken, NJ, USA: John Wiley & Sons, Ltd, 2021: 432–45.
- Günay A, Ward J, Clarkson PJ. Mapping a systems approach to existing healthcare design, delivery and improvement methods. *Proc Des Soc* 2021;1:1997–2006.
- Institute for healthcare improvement. the breakthrough series: IHI's collaborative model for achieving breakthrough improvement; IHI innovation series [White paper]. Boston, USA: Institute for Healthcare Improvement; 2003. Available: <http://www.ihl.org:80/resources/Pages/IHIWhitePapers/TheBreakthroughSeriesIHICollaborativeModelforAchievingBreakthroughImprovement.aspx> [Accessed 29 Mar 2022].
- Nadeem E, Olin SS, Hill LC, et al. Understanding the components of quality improvement collaboratives: a systematic literature review. *Milbank Q* 2013;91:354–94.
- Nicolay CR, Purkayastha S, Greenhalgh A, et al. Systematic review of the application of quality improvement methodologies from the manufacturing industry to surgical healthcare. *Br J Surg* 2012;99:324–35.
- Taylor MJ, McNicholas C, Nicolay C, et al. Systematic review of the application of the plan-do-study-act method to improve quality in healthcare. *BMJ Qual Saf* 2014;23:290–8.
- Nembhard I, Learn A. All teach, all learn, all improve?: the role of interorganizational learning in quality improvement collaboratives. *Health Care Manage Rev* 2012;37:154–64.
- Barr H, Koppel I, Reeves S, et al. Effective interprofessional education – argument, assumption and evidence. *Int J Health Care Qual Assur* 2006;19:ijhcaq.
- Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J Public Health Dent* 2011;71:S52–63.
- Kaplan HC, Brady PW, Dritz MC, et al. The influence of context on quality improvement success in health care: a systematic review of the literature. *Milbank Q* 2010;88:500–59.
- Richards DA, Hallberg IR, eds. *Complex interventions in health: an overview of research methods*. New York, NY, USA: Taylor & Francis Routledge, 2015.
- Zvoch K. How does fidelity of implementation matter? using multilevel models to detect relationships between participant outcomes and the delivery and receipt of treatment. *Am J Eval* 2012;33:547–65.
- Bellg AJ, Borrelli B, Resnick B, et al. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH behavior change consortium. *Health Psychol* 2004;23:443–51.
- Gearing RE, El-Bassel N, Ghesquiere A, et al. Major ingredients of fidelity: a review and scientific guide to improving quality of intervention research implementation. *Clin Psychol Rev* 2011;31:79–88.
- Palsola M, Renko E, Kostamo K, et al. Thematic analysis of acceptability and fidelity of engagement for behaviour change interventions: the let's move it intervention interview study. *Br J Health Psychol* 2020;25:772–89.

- 16 Walton H, Spector A, Tombor I, *et al.* Measures of fidelity of delivery of, and engagement with, complex, face-to-face health behaviour change interventions: a systematic review of measure quality. *Br J Health Psychol* 2017;22:872–903.
- 17 Walton H, Spector A, Williamson M, *et al.* Developing quality fidelity and engagement measures for complex health interventions. *Br J Health Psychol* 2020;25:39–60.
- 18 Kaplan HC, Walsh KE. Context in implementation science. *Pediatrics* 2022;149(Suppl 3):e2020045948C.
- 19 Rousseau DM. Characteristics of departments, positions, and individuals: contexts for attitudes and behavior. *Administrative Science Quarterly* 1978;23:521.
- 20 Kislav R, Wilson P, Cummings G, *et al.* From research evidence to “evidence by proxy”? organizational enactment of evidence-based health care in four high-income countries. *Public Admin Rev* 2019;79:684–98.
- 21 Hoben M, Ginsburg LR, Norton PG, *et al.* Sustained effects of the inform cluster randomized trial: an observational post-intervention study. *Implement Sci* 2021;16:83.
- 22 Berta WB, Wagg A, Cranley L, *et al.* Sustainment, sustainability, and spread study (ssassy): protocol for a study of factors that contribute to the sustainment, sustainability, and spread of practice changes introduced through an evidence-based quality-improvement intervention in canadian nursing homes. *Implement Sci* 2019;14.
- 23 Cranley LA, Norton PG, Cummings GG, *et al.* Scope: safer care for older persons (in residential) environments: a study protocol. *Implement Sci* 2011;6:71.
- 24 Doupe M, Brunkert T, Wagg A, *et al.* SCOPE: safer care for older persons (in residential) environments—a pilot study to enhance care aide-led quality improvement in nursing homes. *Pilot Feasibility Stud* 2022;8.
- 25 Song Y, Thorne TE, Norton PG, *et al.* Rushing care by care aides associated with experiences of responsive behaviors from residents in nursing homes. *J Am Med Dir Assoc* 2022;23:954–61.
- 26 Estabrooks CA, Squires JE, Carleton HL, *et al.* Who is looking after mom and dad? unregulated workers in canadian long-term care homes. *Can J Aging* 2015;34:47–59.
- 27 Estabrooks CA, Squires JE, Cummings GG, *et al.* Study protocol for the translating research in elder care (TREC): building context-an organizational monitoring program in long-term care project (project one). *Implement Sci* 2009;4:52.
- 28 Song Y, MacEachern L, Doupe MB, *et al.* Influences of post-implementation factors on the sustainability, sustainment, and intra-organizational spread of complex interventions [accepted]. *BMC Health Serv Res* 2022;22:666.
- 29 Ogrinc G, Headrick LA, Morrison LJ, *et al.* Teaching and assessing resident competence in practice-based learning and improvement. *J Gen Intern Med* 2004;19.
- 30 Rosenbluth G, Burman NJ, Ranji SR, *et al.* Development of a multi-domain assessment tool for quality improvement projects. *J Grad Med Educ* 2017;9:473–8.
- 31 Singh MK, Ogrinc G, Cox KR, *et al.* The quality improvement knowledge application tool revised (QIKAT-R). *Acad Med* 2014;89:1386–91.
- 32 Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor Quant Methods Psychol* 2012;8:23–34.
- 33 Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15:155–63.
- 34 McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychological Methods* 1996;1:30–46.
- 35 Lawrence RH, Tomolo AM. Development and preliminary evaluation of a practice-based learning and improvement tool for assessing resident competence and guiding curriculum development. *J Grad Med Educ* 2011;3:41–8.
- 36 Steele EM, Butcher R, Carluzzo KL, *et al.* Development of a tool to assess trainees’ ability to design and conduct quality improvement projects. *Am J Med Qual* 2020;35:125–32.
- 37 Kaplan HC, Provost LP, Froehle CM, *et al.* The model for understanding success in quality (MUSIQ): building a theory of context in healthcare quality improvement. *BMJ Qual Saf* 2012;21:13–20.
- 38 Ginsburg L, Easterbrook A, Berta W, *et al.* Implementing frontline worker-led quality improvement in nursing homes: getting to “how.” *Jt Comm J Qual Patient Saf* 2018;44:526–35.
- 39 Coles E, Anderson J, Maxwell M, *et al.* The influence of contextual factors on healthcare quality improvement initiatives: a realist review. *Syst Rev* 2020;9.
- 40 Mazzali C, Paganoni AM, Ieva F, *et al.* Methodological issues on the use of administrative data in healthcare research: the case of heart failure hospitalizations in lombardy region, 2000 to 2012. *BMC Health Serv Res* 2016;16.
- 41 McNicholas C, Lennox L, Woodcock T, *et al.* Evolving quality improvement support strategies to improve plan-do-study-act cycle fidelity: a retrospective mixed-methods study. *BMJ Qual Saf* 2019;28:356–65.
- 42 Mendlowitz A, Croxford R, MacLagan L, *et al.* Usage of primary and administrative data to measure the economic impact of quality improvement projects. *BMJ Open Qual* 2020;9:e000712.
- 43 O’Connor S. Secondary data analysis in nursing research: a contemporary discussion. *Clin Nurs Res* 2020;29:279–84.
- 44 Wickham RJ. Secondary analysis research. *J Adv Pract Oncol* 2019;10:395–400.
- 45 Doupanik SK, Ziniel SI, Glissmeyer EW, *et al.* Validity and reliability of a tool to assess quality improvement knowledge and skills in pediatrics residents. *J Grad Med Educ* 2017;9:79–84.