Evaluating quality improvement at scale: design considerations for routine reporting to executive board of a healthcare organisation in the UK National Health Service

Kia-Chong Chua (✉ kia-chong.chua@kcl.ac.uk)

King's College London <https://orcid.org/0000-0002-6693-6333>

Claire Henderson

Institute of Psychiatry, Psychology and Neuroscience, King's College London

Barbara Grey

South London and Maudsley NHS Foundation Trust

Michael Holland

South London and Maudsley NHS Foundation Trust

Nick Sevdalis

Institute of Psychiatry, Psychology and Neuroscience, King's College London

Research article

Keywords: quality improvement, executive board, routine reporting

Posted Date: November 9th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-103169/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background Quality improvement (QI) in healthcare is a cultural transformation process that requires long-term commitment from the executive board, a critical theme in emerging accounts of QI success in the UK National Health Service (NHS). To help sustain long-term commitment from the executive board, an organisation-wide picture of QI applications and their impact needs to be made routinely visible.

Method We developed a retrospective evaluation drawing inputs from the resident QI team of a healthcare organisation and academic colleagues in the field of implementation and improvement science, as well as peer-reviewed and grey literature on what constitutes success for QI in healthcare. Formative feedback on content relevance, acceptability, and feasibility issues were used to guide evaluation design. The evaluation was conducted as an online survey so that the data accrual process resembles routine reporting to help surface implementation challenges. A purposive sample of QI projects was identified to maximise contrast between projects that were or were not successful as determined by the resident QI team. To hone strategic focus in what should be reported, we also compared factors that might affect project outcomes. For understanding implementation issues, we reviewed data quality to surface challenges in the design and sustainability of routine reporting for the executive board.

Results Out of 52 QI projects, 10 led to a change in routine practice (henceforth referred to as adoption). Details of project outcomes were limited. Project team outcomes, indicative of capacity building, were not systematically documented. Service user involvement, quality of measurement plan, fidelity and documentation of plan-do-study-act (PDSA) cycles had a major impact on adoption. The proximal impact of these process factors on adoption was consistently more apparent than the distal impact of input and contextual factors.

Conclusions Designing a routine reporting framework is an iterative process involving continual dialogue with frontline staff and improvement specialists to navigate data accrual demands. A retrospective evaluation, as in this study, can yield empirical insights for dialogue about the routine visibility of QI applications and their organisation-wide impact, thereby honing the implementation science of QI in a healthcare organisation.

Background

A growing number of health care provider organisations in the UK National Health Service (NHS) are adopting quality improvement (QI) strategies across their organisations [1]. The Care Quality Commission (CQC), an independent regulator of all health and social care services in England, noted a trend after recent inspections that NHS provider organisations with outstanding CQC ratings have applied QI at scale [2, 3]. The systematic application of QI at scale in NHS provider organisations has required concerted investment to sustain a process of cultural transformation. To monitor the developmental stages of this process, an organisation-wide picture of QI applications and their impact is needed. This will help the organisation's executive board commit to a long-term perspective of investments made in QI

infrastructure, a critical theme in emerging accounts of QI success in NHS provider organisations for mental health [1, 4].

With prevailing pressures on NHS resources [5-7], exacerbated by the COVID-19 pandemic, sustaining long-term commitment of the executive board would require routine reports to increase visibility of the stage and progress in cultural transformation and QI success. This routine reporting needs systematic design to address a staff capacity building issue, particularly in developmental stages of embedding a culture of continuous improvement within an organisation. The level of technical skills needed for measurement, data collection and analysis is commonly underestimated in QI, and generally not sufficient in frontline NHS staff despite some training [8]. Furthermore, routinely collected data are often not as clean or well set up as originally anticipated, often requiring extensive effort to bring them up to a standard suitable for use [8]. This means that an organisation-wide picture of QI applications and their impact cannot be readily constructed by aggregating project reports of impact or success. Even for case studies with proper design and adequate analysis, they would not offer an organisation-wide picture of cultural transformation and QI success.

There is a need for constructing a routine reporting framework that can help inform the executive board and sustain long-term commitments in QI. In this study we carried out a pilot evaluation to develop content of routine reporting. To hone strategic focus in what should be reported, we also compared factors that might affect project outcomes. For understanding implementation issues, we reviewed data quality to surface challenges in the design and sustainability of routine reporting for the executive board.

Methods

Setting

An NHS organisation that provides specialist care for mental health in South London established a resident QI team in 2016 with mandate to foster a continuous improvement culture. The resident team supports QI projects led by frontline staff through training and coaching on QI methodology (e.g., Model for Improvement, driver diagrams, Plan-Do-Study-Act (PDSA) cycles). In partnership with an academic institution, the NHS provider organisation also adopted a researcher-in-residence model that embeds an academic faculty in the resident QI team to support data and evaluation needs. With a steady rise of QI activities within the organisation between 2016 and 2018, the executive board requested an evaluation of these nascent developments.

Design

To explore what should be included for an executive board report on organisation-wide QI programme, we developed a pilot retrospective evaluation such that the data accrual process resembled routine reporting. Under these conditions, challenges in routine reporting might be made apparent to guide design considerations. A scoping exercise was first initiated by the researcher-in-residence to develop a proposal of evaluation content. This involved individual consultations with three improvement specialist

colleagues in the resident QI team, coupled with a rapid evidence scan of peer-reviewed and grey literature on what constitute success for QI in healthcare. Key concerns were raised with time resources, staff turnover, service user involvement, and outcomes like whether project ideas were adopted and spread. The literature also drew attention to the fidelity of PDSA applications and implementation of data or measurement plan. A list of items was generated and circulated among the entire team of six improvement specialists in the resident QI team for formative feedback on content relevance, acceptability and feasibility issues. This feedback was combined with inputs from two academic colleagues with expertise in implementation and improvement science. A revised version was then piloted with two colleagues back in the resident QI team before it was finalised as an online survey (see supplementary material).

Project outcomes

To offer an overview of a diverse range of QI projects, we chose to look at whether projects: achieved their aims; led to a change in routine practice (adoption); and triggered similar projects beyond the site of their original conception (spread). These outcomes were core themes that are relevant for all projects despite heterogeneous and highly localised aims. Such a focus enables regular reporting at scale and acts as an early signal for areas that needs attention.

Costs and benefits

To attribute costs and benefits of QI projects within the organisation, the evaluation captured resource use in terms of amount of contact with the resident QI team. Considerable variation in documentation precluded formal cost-benefit analysis in a retrospective evaluation, but a question was added on whether it was possible to quantify improvement in terms of cost savings and whether this had been attempted. QI is underpinned by an organisational management philosophy that recognises the critical need to empower frontline staff to learn and participate in continuous improvement in face of escalating complexity and change [9]. On this basis, we also considered aspects of skill development and capacity building which may be attributed to each quality improvement project. Such a focus aligns with the philosophy of QI to move beyond performance assurance and cultivate an organisational culture of learning regardless of project outcomes. As a measure of skill development, we inquired on the extent and form of dissemination efforts. We also tracked capacity building by inquiring whether project team members went on to develop new QI projects.

Contextual, input and process factors

To explore what may facilitate or impede QI project success, we inquired about contextual, input and process factors. Contextual factors refer to organisational conditions that are not within the influence of project teams. Besides setting (in-patient / community care), the resident QI team drew particular attention to contextual aspects like whether it was the team's first ever QI project, and whether protected time for QI activities was officially sanctioned. Input factors comprised team characteristics and staff

turnover. Process factors refer to actions or decisions of the project teams. They comprised stakeholder engagement, PDSA cycles and measurement plans implemented.

Sampling

With approximately 200 QI projects initiated between 2016-18, retrospective information retrieval was undertaken with a purposive sample for feasibility reasons. Each improvement specialist of the trust’s resident QI team was requested to identify up to five QI projects that they considered as ‘successful’ and up to another five that they considered as ‘unsuccessful’. Specifically, they were asked to rely on their own assessment of what did and did not work to maximise the gradient of contrast across the selection of QI projects. This in turn would also help surface insights on “work-as-done” rather than imposing a “work-as-imagined” criterion [10] in the absence of an established definition of ‘success’ of in QI [11].

Analysis

We first enumerated project outcomes in terms of whether they achieved their aims, introduced change ideas that were adopted in routine practice (adoption), and triggered similar projects at other sites (spread). We then compared projects that did and did not lead to a change in routine practice to see if they differ in terms of contextual, input and process factors. To compare the associations between these factors and adoption, we compared effect sizes based on odds ratios (with confidence intervals from logistic regression). An odds ratio (OR) smaller than 1.5 was considered to be a small effect size, whereas $OR > 5.0$ was considered to be a large effect size [12].

Results

The study sample included 52 QI projects across five boroughs of London, UK served by the NHS provider organisation (Table 1). Thirty projects were conceived by community mental health teams and the remaining (n=22) by inpatient care teams. Of the three themes of Quality Priorities,[13: patient safety, clinical effectiveness, and patient experience] improving clinical effectiveness was the most common focus in project aims (29 / 52). A small handful focused on multiple priorities at once (6 community mental health and 7 inpatient care projects).

Table 1. Study sample of 52 quality improvement projects in South London and Maudsley NHS Foundation Trust (2016 – 2018)

	Community Mental Health n = 24 (6)*	Inpatient n = 15 (7)*
Borough		
Croydon	8	10
Lambeth	7	4
Lewisham	7	4
Southwark	7	3
Wandsworth	1	1
Patient Safety		
Reduce use of restrictive interventions on service users	1 (2)	0 (3)
Safer staffing	0	1 (1)
Risk assessments	1 (1)	1 (1)
Clinical Effectiveness		
Physical healthcare screening	4 (1)	4 (3)
Care planning	10 (5)	6 (3)
Developing electronic systems to improve care delivery	3 (1)	2 (1)
Patient Experience		
Reducing number of acute out-of-area treatments	3 (2)	0 (2)
Carer's assessment and associated care plan	1 (3)	0 (1)
Quality of environment and food	1 (1)	1 (2)

* numbers in parentheses refers to number of projects that included multiple Quality Priorities in project aims

Project outcomes

In terms of project outcomes (Table 2), 18 of 52 projects (35%) reported a change in routine practice (adoption). However, only 10 of them reported formal project closure with aims achieved. Out of 7 of 52 projects (13%) that triggered similar projects in other sites (spread), only 3 reported formal project closure with aims achieved. A plausible explanation for these findings could be that some projects were organisation-wide initiatives that were adopted or spread across service sites regardless of the project outcome at specific sites. In light of this divergence between “work-as-imagined” and “work-as-done”,^[10] we decided to retain a more stringent definition by labelling projects as successful (n=10) only if they led to adoption after achieving their aims at formal closure. This offered a more interpretable benchmark for making comparisons with the remaining 42 projects.

Table 2. Project outcomes

		Adopted (n = 18)	Spread (n = 7)
Formal closure (n = 23)	Aims achieved (n = 12)	10	3
	Aims not achieved (n = 11)	3	2
Terminated (n = 29)	Aims achieved (n = 4)	4	1
	Aims not achieved (n = 25)	1	1

⁺ Project Status: Closed, Terminated at xP (Plan), xD (Do), xS (Study), xA (Act)

Adopted = Change idea adopted, Spread = Triggered similar projects

Costs and benefits

Among the 10 successful projects, half required six or more months (median = 6.0) for completion. Those that were not successful after formal closure (n = 13) showed large variation. Half were completed in under three months (median = 2.8) but some took up to 12 months (Figure 1). Among projects that did not reach formal closure, those that terminated at the Planning stage of PDSA (n = 18, median = 1.8 months) showed a shorter life span than those that terminated in more advanced stages (n = 11, median = 6.0 months). Meetings with the resident QI team staff took place typically on a monthly basis for successful projects (median = 1.1 meetings monthly), a slightly higher rate than for all others (median = 0.6 – 0.9 meetings monthly). Monthly correspondence between project and resident QI team staff (email/phone) shows a similar picture, with slightly higher activity levels in successful projects (median = 7.5 vs 3.8 – 4.8 monthly email/phone).

Retrospective estimates were requested for the number of service users and staff who directly benefitted from the undertaken QI projects. This proved problematic as indications were available for only a handful of projects (8 / 10 successful projects vs 11 / 42 for all others). Similarly, when inquired about whether it was possible to quantify improvement in terms of cost savings, more than half of the projects reported “not known”.

Among the 10 successful projects, seven disseminated publications of their findings (2 locally, 4 beyond local site/service, 1 not known). Among the remaining 42, five did so (2 locally, 2 beyond local site/service, 1 not known). For most projects, survey responses indicated “not known” in both respects. Two of the successful project teams went on to develop two new projects, whereas six in the latter group developed 10 new projects.

Contextual factors

We examined a range of factors that might be associated with QI project outcomes (Table 3). Seven (out of 30) QI projects in community mental health settings led to a change in routine practice. These odds were halved (Odds ratio, OR = 0.5, 95% CI: 0.1 – 2.3) for projects in inpatient care settings where three (out of 22) led to adoption, but the differences in odds did not attain statistical significance. The odds of adoption were twice higher for QI projects that required funding (OR = 2.1, 95% CI: 0.4 – 10.4), but statistical significance was not attained as well. Among the five projects for which funding was actually available, none achieved adoption (consequently, we could not calculate an OR for comparing odds). The odds of adoption were five times lower (OR = 0.2, 95% CI: >0.1 – 0.9) if it was the team’s first attempt at carrying out a QI project. These odds were five times higher if protected time for QI activities was officially sanctioned for the project (OR = 5.2, 95% CI: 1.2 – 23.4).

Table 3. Logistic regression models for association between project outcome of adoption (dependent variable) and contextual, input, and process factors (independent variables).

	n ₁	n ₂	Odds Ratio	95% CI
Contextual factors				
Inpatient	22	3	0.5	0.1 - 2.3
First QI project ⁺	27	2	0.2	>0.1 - 0.9
Time officially sanctioned ⁺	20	7	5.2	1.2 - 23.4
Funding available	5	0	#	#
Funding required	10	3	2.1	0.4 - 10.4
Input factors				
Made budget plan	5	1	1.1	0.1 - 10.6
Project lead: non-clinical staff	8	1	0.6	0.1 - 5.1
Project lead: non-managerial staff	18	3	0.8	0.2 - 3.4
Team size	-	-	1.4	0.7 - 2.7
Staff turnover	22	4	0.9	0.2 - 3.6
Process factors				
Engaged team leader	28	6	1.4	0.3 - 5.5
Engaged stakeholders	14	5	3.7	0.9 - 15.5
Engaged service users ⁺⁺	10	5	7.4	1.6 - 34.9
No. of outcome measures ⁺	-	-	3.2	1.1 - 9.8
No. of primary drivers ⁺	-	-	2.7	1.3 - 5.9
No. of secondary drivers ⁺	-	-	1.5	1.1 - 1.9
Aims quantified	35	10	#	#
Primary drivers tagged with measures ⁺⁺	13	6	7.5	1.7 - 33.7
Secondary drivers tagged with measures ⁺	11	5	6.0	1.3 - 27.2
Balancing measures ⁺⁺	10	5	7.4	1.6 - 34.9
No. of PDSA completed ⁺	-	-	1.5	1.1 - 2.2
PDSA that completed >1 cycle ⁺⁺	13	6	7.5	1.7 - 33.7
PDSA with documentation ⁺⁺	13	9	85.5	8.5 - 860.2
Data collected before implementing change idea ⁺⁺	9	5	9.5	1.9 - 47.6
Median value of random variation in outcome measures ⁺	12	5	5.0	1.1 - 22.0
Median value of random variation in process measures	4	4	#	#
Median value of random variation in balancing measures	2	2	#	#
Data collected after implementing change idea ⁺⁺	13	6	7.5	1.7 - 33.7

n₁: total number of projects that satisfy the condition described by the independent variable

n₂: total number of projects (in n₁) that led to a change in routine practice (adoption).

independent variables for which odds of project outcome could not be calculated

+ independent variables that show statistically significant odds ratio

++ independent variables for which conservative estimates (lower bound of 95%CI) show at least a moderate effect size (OR > 1.5, or in opposite direction: OR < 0.7)

Input factors

The odds of adoption increased slightly (OR = 1.4, 95% CI: 0.7 – 2.7) with team size but decreased if the project was led by non-clinical staff (OR = 0.6, 95% CI: 0.1 – 5.1) or non-managerial staff (OR = 0.8, 95% CI: 0.2 – 3.4). They were similar between QI projects that did or did not make a budget plan (OR = 1.1, 95% CI: 0.1 – 10.6). Staff turnover appeared slightly detrimental to the odds of adoption (OR = 0.9, 95% CI: 0.2 – 3.6). However, none of these differences in input factors attained statistical significance.

Process factors

The odds of adoption were higher if the project team engaged their team leader, stakeholders (e.g., staff members not in project team), and service users. Only service user engagement showed a statistically reliable impact, with a large effect size (OR = 7.4, 95% CI: 1.6 – 34.9).

The odds of adoption increased moderately with the number of outcome measures attached to the aim statement of the driver diagram produced by the project team (OR = 3.7, 95% CI: 1.1 – 9.8). This was also the case for the number of primary (OR = 2.7, 95% CI: 1.3 – 5.9) and secondary drivers (OR = 1.5, 95% CI: 1.1 – 1.9) in these diagrams.

Among 35 projects that quantified their target outcomes in the aims statement, 10 led to adoption. In the remaining 17 that did not quantify their target outcomes in the aims statement, none achieved adoption (consequently, we could not calculate an OR for comparing odds). The odds of adoption were much higher if measures were tagged to the primary and secondary drivers (OR = 7.5, 95% CI: 1.7 – 33.7 and OR = 6.0, 95% CI: 1.3 – 27.2 respectively). Projects that included balancing measures also had much higher odds of adoption (OR = 7.4, 95% CI: 1.6 – 34.9).

The odds of adoption increased slightly with the number of PDSAs carried out (OR = 1.5, 95% CI: 1.1 – 2.2). However, when the comparisons focused simply on whether projects did or did not report PDSA with more than one cycle, the odds were much higher (OR = 7.5, 95% CI: 1.7 – 33.7) if they did. The odds were even higher for projects that had PDSA documentation (OR = 85.5, 95% CI: 8.5 – 860.2). Despite the wide confidence interval, the lower bound interval estimate suggests that this latter aspect of PDSA had a major impact on project outcome.

Projects that collected data before implementing change ideas showed much higher odds of adoption (OR = 9.5, 95% CI: 1.9 – 47.6). This also applies to projects that established the median value of random variation in outcome measures (OR = 5.0, 95% CI: 1.1 – 22.0). We could not calculate OR for comparing odds for projects that established the median value of random variation in process and balancing measures because all that did so achieved adoption. Projects that collected data after implementing a change idea also show much higher odds of adoption (OR = 7.5, 95% CI: 1.7 – 33.7).

Discussion

This study developed a pilot evaluation to explore what should be reported to give an organisation-wide picture and sustain long-term commitments of the executive board in embedding a culture of continuous improvement. The evaluation yielded exploratory insights and lessons were learnt from the practical challenges that emerged. We reflect on these with the view of informing what is needed for a fuller development of routine reporting on cultural transformation and QI success in an NHS organisation providing specialist care in mental health.

We found that time invested in meetings, virtual correspondences, and overall project duration are data with minimal information retrieval burden. Across a diverse range of projects, time invested by staff is an opportunity cost that could serve as a common denominator for return on investment.

In contrast, we encountered major challenges in measuring benefits. Cost savings is often not an immediate focus of QI in mental healthcare. Consequently, it is usually not part of the measurement plan. Improvement in patient safety and experience needs a reporting framework that translates gains into cost savings. The onus of developing structured reports is likely to fall on the resident QI team to enable consistent measurement and coherent iterations. The level of skill development and capacity building in QI was not known for most projects in this retrospective evaluation. As with organisational benefits, monitoring project team outcomes (as opposed to project outcomes) is best carried out by the resident QI unit.

A fundamental goal of QI is to bring about a change in routine practice [14]. Of the 52 QI projects led by frontline NHS staff, 10 achieved this goal. Among factors that reliably increased the odds of effecting a change in routine practice, service user involvement showed at least a moderate to large impact. This finding is consistent with the wider literature on QI in mental health [15]. It also highlights the need for more granular data on the level of involvement of this stakeholder group, to guide efforts in optimising their potential contributions [16].

With the Model for Improvement as the dominant QI activity paradigm, the driver diagram is a common tool for problem recognition and theory of change analysis. If measures were tagged to primary drivers and balancing measures in the driver diagram, the odds of effecting a change in routine practice increased by at least a moderate to large effect size. This was also the case if data were collected before and after implementing a change. Our findings concur with existing improvement science literature. Measurement and the use of data is at the heart of the Model for Improvement [17]. A transparent, data-driven approach is paramount [18], otherwise project teams may get stuck in the 'Do' phase of a PDSA, or reach no actionable insight in the 'Study' phase.

As structured interventional experiments for testing changes, iterative cycles in PDSA are key for learning [19]. Improvement work is less likely to succeed if iterative cycles are too few [20]. In our evaluation, we found that the odds of effecting a change in routine practice increases by at least a moderate to large effect size if projects reported PDSA that completed more than one cycle. The task of PDSA

documentation by far show the strongest impact even by conservative estimates. Documentation of each stage of the PDSA cycle supports scientific quality, learning and reflection; even if PDSA cycles are well-executed, poor documentation would impede organisational memory and transferability of learning [20, 21]. Documentation is a critical part of fidelity in PDSA. It is not a simple task. Training should not overemphasise the conceptual simplicity of PDSA [19]. Achieving high fidelity in PDSA application as a QI method will require a gradual and negotiated process to explore different perspectives and encourage new ways of working [22]. Collecting routine data on PDSA fidelity can aid conversations in this effort. When PDSA cycles are performed and reported appropriately, knowledge is accumulated, which can be shared readily [20]. Of note, QI is not synonymous with improving quality [1]. Even if there was no improvement, PDSA cycles with rigorous measurement plans would still generate learning. This would also be considered a QI success [19].

Study limitations

This study set out to explore an organisation-wide picture of QI projects and their impact. While we defined adoption as bringing about a change in routine practice, the process of retrospective information retrieval posed major challenges in obtaining a wider range of indicators for illuminating on the impact. Furthermore, the absence of an established definition of 'success' of in QI [11] meant that the resident QI team had to rely on their own judgement when selecting a purposive sample. Coupled with retrospective recall difficulties, sampling bias is likely to have an impact on study findings. Notably however, our survey-based findings resonate with those from in-depth qualitative studies in terms of the process factors that matter for project outcomes. This suggests that the skilful application of QI can be routinely monitored as an indication of cultural transformation and QI success.

While several process factors showed an apparent impact on project outcomes, contextual and input factors generally showed little impact. This is likely to be due to limitations in the scope of our survey rather than evidence that these latter aspects are not important. It is well-established in improvement science literature that contextual factors play a prominent role [23]. However, the sheer number of variables and the unpredictability of their interactions make it hard to predict the distal impact of contextual factors on project outcomes [24]. Consequently, the need to minimise data accrual burden led to a narrow scope of contextual and input factors. Instead, we prioritise the scope of process factors (e.g. fidelity of data and measurement plans), to illuminate proximal influences that are potentially amenable to staff training and interventions.

Conclusion

To advance with a long-term perspective of cultural transformation and QI success, the executive board needs an organisation-wide picture that can be routinely made visible with minimal technical expertise. Setting up a routine monitoring framework can offer more timely vigilance than designing an elaborate retrospective evaluation, which some argued, is akin to attempting to "drive by looking in the rear-view mirror" [25]. Practically, routine monitoring can also improve data quality by alleviating information

retrieval difficulties and ensuring systematic data accrual. Operationally, respondent burden can be alleviated by spreading data collection over multiple occasions. Short and well-timed surveys can optimise relevance and thus the acceptability of data collection. Data can also be collected from different types of respondents (e.g., project team lead, sponsor, resident QI team). Well-targeted respondents will provide the most accurate data if only the most relevant questions are asked, thereby also minimising respondent burden.

Designing a routine reporting framework is an iterative process involving continual dialogue with frontline staff and improvement specialists to navigate data accrual demands [26]. Across diverse project aims, there is a need to identify themes that hold core relevance in QI. For an executive board, this framework should generate strategic insights that can inform resource commitments in QI. For staff and improvement specialists, routine reporting should offer a feedback loop to support discussion and engagement around finding the right focus [27] and applying QI with fidelity [19, 28]. Developing routine reporting can be an asset for improving practice [29], and honing the implementation science of QI in healthcare.

Declarations

Ethics approval and Consent to participate

According to the policy activities that constitute research at the South London & Maudsley NHS Foundation Trust, this work met criteria for operational improvement activities exempt from ethics review

Consent for publication

Not applicable

Availability of data and materials

The data that support the findings of this study are available from Dr. Kia-Chong Chua but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of Dr. Barbara Grey.

Competing interests

NS is the director of the London Safety and Training Solutions Ltd, which offers training in patient safety, implementation solutions and human factors to healthcare organisations. The other authors have no conflicts of interest to declare.

Funding

NS' research is supported by the National Institute for Health Research (NIHR) Applied Research Collaboration (ARC) South London at King's College Hospital NHS Foundation trust. NS and KC are

members of King's Improvement Science, which offers co-funding to the NIHR ARC South London and comprises a specialist team of improvement scientists and senior researchers based at King's College London. Its work is funded by King's Health Partners (Guy's and St Thomas' NHS Foundation trust, King's College Hospital NHS Foundation trust, King's College London and South London and Maudsley NHS Foundation trust), Guy's and St Thomas' Charity and the Maudsley Charity. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR, the charities, or the Department of Health and Social Care.

Authors' contributions

The evaluation was designed by KC, NS and BG. The analysis was conducted by KC. The manuscript was drafted by KC with critical revisions from CH, BG, MH and NS. All authors read and approved the final manuscript.

Acknowledgements

Colleagues in the QI resident team of the South London and Maudsley NHS Foundation Trust gave formative feedback on evaluation design and conducted project reviews to generate the study data. They include Rachael Leaton, Cassie Philpin, Polly Ragoobar, Sasha Persuad, Nirusha Nicholas, and Devon Elliott.

Authors' information

KC is a lecturer in applied health statistics in the Centre for Implementation Science in King's College London. He is also a researcher-in-residence in the QI resident team in the South London and Maudsley NHS Foundation Trust.

CH is a Reader in public mental health in the Centre for Implementation Science in King's College London. She is also a honorary consultant psychiatrist and the joint Academic Director for the clinical academic group for quality improvement and implementation science in the South London and Maudsley NHS Foundation Trust.

BG is the Director of quality improvement in the South London and Maudsley NHS Foundation Trust.

MH is the Medical Director (executive board member) of the South London and Maudsley NHS Foundation Trust.

NS is a Professor of implementation science and patient safety and Director of the Centre for Implementation Science in King's College London. He is also the Academic Director of the clinical academic group for quality improvement and implementation science in the South London and Maudsley NHS Foundation Trust.

References

1. Ross S, Naylor C: **Quality improvement in mental health.** In.: The King's Fund; 2017.
2. Care Quality Commission: **The state of care in mental health services 2014 to 2017.** In.; 2018.
3. Care Quality Commission: **Quality improvement in hospital trusts: Sharing learning from trusts on a journey of QI.** In.; 2018.
4. Shah A, Course S: **Building the business case for quality improvement: a framework for evaluating return on investment.** *Future Healthc J* 2018, **5**(2):132-137.
5. Dunn P, McKenna H, Murray R: **Deficits in the NHS 2016.** In.: The King's Fund; 2016.
6. Smithson R, Maguire D, Honeyman M, Jabbal J: **Workforce planning in the NHS.** In.: The King's Fund; 2015.
7. Gilbert H: **Mental health under pressure.** In.: The King's Fund; 2015.
8. Woodcock T, Liberati EG, Dixon-Woods M: **A mixed-methods study of challenges experienced by clinical teams in measuring improvement.** *BMJ Qual Saf* 2019.
9. Blumenthal D, Kilo CM: **A report card on continuous quality improvement.** *Milbank Q* 1998, **76**(4):625-648, 511.
10. Hollnagel E: **Why is work-as-imagined different from work-as-done?** In: *Resilient Health Care. Volume 2*, edn. Edited by Wears RL, Hollnagel E. London: CRC Press; 2017.
11. Morganti KG, Lovejoy S, Haviland AM, Haas AC, Farley DO: **Measuring success for health care quality improvement interventions.** *Med Care* 2012, **50**(12):1086-1092.
12. Chen H, Cohen P, Chen S: **How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies.** *Communications in Statistics - Simulation and Computation* 2010, **39**(4):860-864.
13. South London and Maudsley NHS Foundation Trust: **Quality Account 2016/2017.** In.; 2017.
14. Batalden PB, Davidoff F: **What is "quality improvement" and how can it transform healthcare?** *Qual Saf Health Care* 2007, **16**(1):2-3.
15. Robert G, Hardacre J, Locock L, Bate P, Glasby J: **Redesigning mental health services: lessons on user involvement from the Mental Health Collaborative.** *Health Expect* 2003, **6**(1):60-71.
16. Armstrong N, Herbert G, Aveling EL, Dixon-Woods M, Martin G: **Optimizing patient involvement in quality improvement.** *Health Expect* 2013, **16**(3):e36-47.
17. Boland B: **Quality improvement in mental health services.** *BJPsych Bull* 2020, **44**(1):30-35.
18. Birdas TJ, Rozycki GF, Dunnington GL, Stevens L, Liali V, Schmidt CM: **"Show Me the Data": A Recipe for Quality Improvement Success in an Academic Surgical Department.** *J Am Coll Surg* 2019, **228**(4):368-373.
19. Reed JE, Card AJ: **The problem with Plan-Do-Study-Act cycles.** *BMJ Qual Saf* 2016, **25**(3):147-152.
20. Ogrinc G, Shojania KG: **Building knowledge, asking questions.** *BMJ Qual Saf* 2014, **23**(4):265-267.
21. Taylor MJ, McNicholas C, Nicolay C, Darzi A, Bell D, Reed JE: **Systematic review of the application of the plan-do-study-act method to improve quality in healthcare.** *BMJ Qual Saf* 2014, **23**(4):290-298.

22. McNicholas C, Lennox L, Woodcock T, Bell D, Reed JE: **Evolving quality improvement support strategies to improve Plan-Do-Study-Act cycle fidelity: a retrospective mixed-methods study.** *BMJ Qual Saf* 2019, **28**(5):356-365.
23. Kaplan HC, Provost LP, Froehle CM, Margolis PA: **The Model for Understanding Success in Quality (MUSIQ): building a theory of context in healthcare quality improvement.** *BMJ Qual Saf* 2012, **21**(1):13-20.
24. Braithwaite J: **Changing how we think about healthcare improvement.** *BMJ* 2018, **361**:k2014.
25. Hamblin R, Shuker C: **Well spotted: but now you need to do something.** *BMJ Qual Saf* 2020.
26. Gardner K, Olney S, Dickinson H: **Getting smarter with data: understanding tensions in the use of data in assurance and improvement-oriented performance management systems to improve their implementation.** *Health Res Policy Syst* 2018, **16**(1):125.
27. Soong C, Cho HJ, Shojania KG: **Choosing quality problems wisely: identifying improvements worth developing and sustaining.** *BMJ Qual Saf* 2020.
28. Knudsen SV, Laursen HVB, Johnsen SP, Bartels PD, Ehlers LH, Mainz J: **Can quality improvement improve the quality of care? A systematic review of reported effects and methodological rigor in plan-do-study-act projects.** *BMC Health Serv Res* 2019, **19**(1):683.
29. Dixon-Woods M: **How to improve healthcare improvement-an essay by Mary Dixon-Woods.** *BMJ* 2019, **367**:l5514.

Figures

Figure 1

Duration (months) of projects that terminated in planning (xP), doing (xD), studying (xS) and Acting (xA) stage of PDSA cycle, as well as projects that had formal closure with and without a change in routine practice (xAdopt / Adopt).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [evalQI.pdf](#)