OPEN ACCESS

Systematic review of the application of the plan–do–study–act method to improve quality in healthcare

Michael J Taylor,[1,2] Chris McNicholas,[2] Chris Nicolay,[1] Ara Darzi,[1] Derek Bell,[2] Julie E Reed[2]

▶ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/bmjqs-2013-001862).

[1]Department of Surgery and Cancer, Imperial College London, London, UK
[2]National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care (CLAHRC) for North-West London, London, UK

Correspondence to
Michael J Taylor, Academic Surgical Unit, 10th Floor, QEQM building, St Mary's Hospital, Paddington, London W2 1NY, UK; mtaylor3@imperial.ac.uk

Received 29 January 2013
Revised 25 June 2013
Accepted 4 July 2013
Published Online First 23 August 2013

▶ http://dx.doi.org/10.1136/bmjqs-2013-002703

To cite: Taylor MJ, McNicholas C, Nicolay C, *et al*. *BMJ Qual Saf* 2014;**23**:290–298.

ABSTRACT

Background Plan–do–study–act (PDSA) cycles provide a structure for iterative testing of changes to improve quality of systems. The method is widely accepted in healthcare improvement; however there is little overarching evaluation of how the method is applied. This paper proposes a theoretical framework for assessing the quality of application of PDSA cycles and explores the consistency with which the method has been applied in peer-reviewed literature against this framework.

Methods NHS Evidence and Cochrane databases were searched by three independent reviewers. Empirical studies were included that reported application of the PDSA method in healthcare. Application of PDSA cycles was assessed against key features of the method, including documentation characteristics, use of iterative cycles, prediction-based testing of change, initial small-scale testing and use of data over time.

Results 73 of 409 individual articles identified met the inclusion criteria. Of the 73 articles, 47 documented PDSA cycles in sufficient detail for full analysis against the whole framework. Many of these studies reported application of the PDSA method that failed to accord with primary features of the method. Less than 20% (14/73) fully documented the application of a sequence of iterative cycles. Furthermore, a lack of adherence to the notion of small-scale change is apparent and only 15% (7/47) reported the use of quantitative data at monthly or more frequent data intervals to inform progression of cycles.

Discussion To progress the development of the science of improvement, a greater understanding of the use of improvement methods, including PDSA, is essential to draw reliable conclusions about their effectiveness. This would be supported by the development of systematic and rigorous standards for the application and reporting of PDSAs.

INTRODUCTION

Delivering improvements in the quality and safety of healthcare remains an international challenge. In recent years, quality improvement (QI) methods such as plan–so–study–act (PDSA) cycles have been used in an attempt to drive such improvements. The method is widely used in healthcare improvement; however there is little overarching evaluation of how the method is applied. This paper proposes a theoretical framework for assessing the quality of application of PDSA cycles and explores the quality and consistency of PDSA cycle application against this framework as documented in peer-reviewed literature.

Use of PDSA cycles in healthcare

Despite increased investment in research into the improvement of healthcare, evidence of effective QI interventions remains mixed, with many systematic reviews concluding that such interventions are only effective in specific settings.[1–4] To make sense of these findings, it is necessary to understand that delivering improvements in healthcare requires the alteration of processes within complex social systems that change over time in predictable and unpredictable ways.[5] Research findings highlight the influential effect that local context can have on the success of an intervention[6 7] and, as such, 'single-bullet' interventions are not anticipated to deliver consistent improvements. Instead, effective interventions need to be complex and multi-faceted[8–11] and developed iteratively to adapt to the local context and respond to unforeseen obstacles and unintended effects.[12 13] Finding effective QI methods to support iterative development to test and evaluate

interventions to care is essential for delivery of high-quality and high-value care in a financially constrained environment.

PDSA cycles provide one such method for structuring iterative development of change, either as a standalone method or as part of wider QI approaches, such as the Model for Improvement (MFI), Total Quality Management, Continuous QI, Lean, Six Sigma or 'Quality Improvement Collaboratives'.[3 4 14] Despite increased use of QI methods, the evidence base for their effectiveness is poor and under-theorised.[15–17] PDSA cycles are often a central component of QI initiatives, however few formal objective evaluations of their effectiveness or application have been carried out.[18] Some PDSA approaches have been demonstrated to result in significant improvements in care and patient outcomes,[19] while others have demonstrated no improvement at all.[20–22]

Although at the surface level these results appear disheartening for those involved in QI, there is a need to explore the extent to which the PDSA method has been successfully deployed to draw conclusions from these studies. Rather than see the PDSA method as a 'black box' of QI,[23] it is important to understand that the use of PDSA cycles is, itself, a complex intervention made up of a series of interdependent steps and key principles that inform its application[5 24 25] and that this application is also affected by local context.[26] To interpret the results regarding the outcome(s) from the application of PDSA cycles (eg, whether processes or outcomes of care improved) and gauge the effectiveness of the method, it is necessary to understand how the method has been applied.

No formal criteria for evaluating the application or reporting of PDSA cycles currently exist. It is only in recent years, through SQUIRE guidelines, that frameworks for publication have been developed that explicitly consider description of PDSA application.[27 28] We consider that such criteria are necessary to support and assess the effective application of PDSA cycles and to increase their legitimacy as a scientific method for improvement. We revisited the origins and theory of the method to develop a theoretical framework to evaluate the application of the method.

The origins and theory of PDSA cycles

The PDSA method originates from industry and Walter Shewhart and Edward Deming's articulation of iterative processes which eventually became known as the four stages of PDSA.[25] PDCA (plan–do–check–act) terminology was developed following Deming's early teaching in Japan.[29] The terms PDSA and PDCA are often used interchangeably in reference to the method. This distinction is rarely referred to in the literature and for the purpose of this article we consider PDSA and PDCA but refer to the methodologies generally as 'PDSA' cycles unless otherwise stated.

Users of the PDSA method follow a prescribed four-stage cyclic learning approach to adapt changes aimed at improvement. In the 'plan' stage a change aimed at improvement is identified, the 'do' stage sees this change tested, the 'study' stage examines the success of the change and the 'act' stage identifies adaptations and next steps to inform a new cycle. The MFI[30] and FOCUS[31] (see figure 1) frameworks have been developed to precede the use of PDSA and PDCA cycles[30 31] respectively (table 1).

In comparison to more traditional healthcare research methods (such as randomised controlled trials in which the intervention is determined in advance and variation is attempted to be eliminated or controlled for), the PDSA cycle presents a pragmatic scientific method for testing changes in complex systems.[32] The four stages mirror the scientific experimental method[33] of formulating a hypothesis, collecting data to test this hypothesis, analysing and interpreting the results and making inferences to iterate the hypothesis.

The pragmatic principles of PDSA cycles promote the use of a small-scale, iterative approach to test interventions, as this enables rapid assessment and provides flexibility to adapt the change according to feedback to ensure fit-for-purpose solutions are developed.[10 12 13] Starting with small-scale tests provides users with freedom to act and learn; minimising risk to patients, the organisation and resources required and providing the opportunity to build evidence for change and engage stakeholders as confidence in the intervention increases.

In line with the scientific experimental method, the PDSA cycle promotes prediction of the outcome of a test of change and subsequent measurement over time (quantitative or qualitative) to assess the impact of an intervention on the process or outcomes of interest. Thus, learning is primarily achieved through interventional experiments designed to test a change. In recognition of working in complex settings with inherent variability, measurement of data over time helps understand natural variation in a system, increase awareness of other factors influencing processes or outcomes, and understand the impact of an intervention.

As with all scientific methods, documentation of each stage of the PDSA cycle is important to support scientific quality, local learning and reflection and to ensure knowledge is captured to support organisational memory and transferability of learning to other settings.

This review examines the application of PDSA cycles as determined by these principle features of the PDSA method described above. We recognise that a number of health and research related contextual factors may affect application of the method but these factors are beyond the scope of this review. The review intends to improve the understanding of

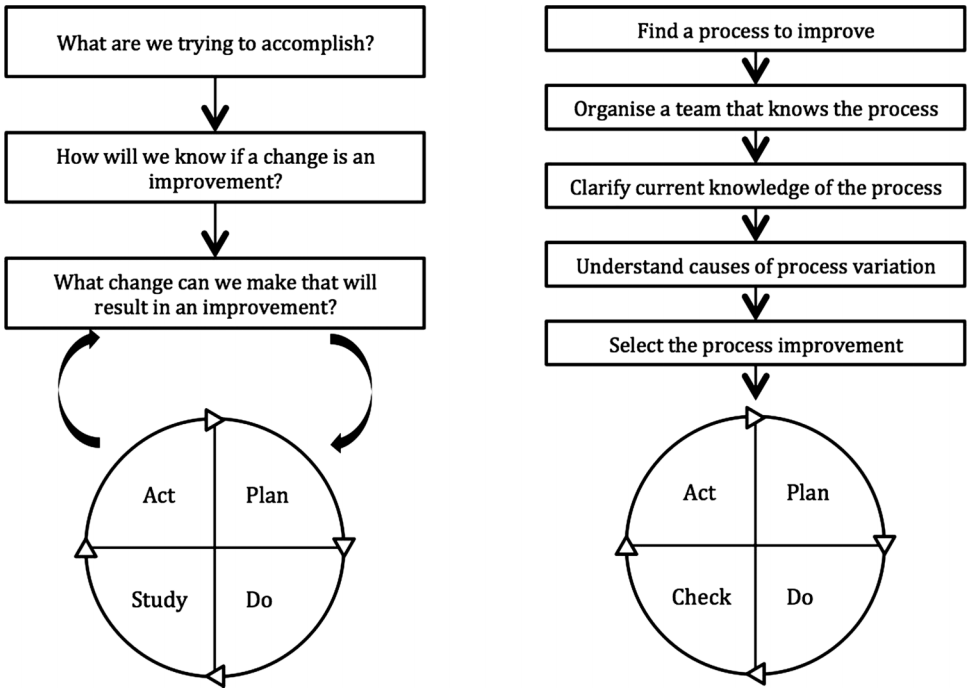

Figure 1 The Model for Improvement; FOCUS.

whether the PDSA method is being used and reported in line with the literature informed criteria and therefore inform the interpretation of studies that have used PDSA cycles to facilitate iterative development of an intervention.

METHODS

A systematic narrative review was conducted in adherence to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.[34]

Search

The search was designed to identify peer-reviewed publications describing empirical studies that applied the PDSA method. Taking into account the development of the method and terminology, the search terms used were 'PDSA', 'PDCA', 'Deming Cycle', 'Deming Circle', 'Deming Wheel' and 'Shewhart Cycle'. No year of publication restrictions were imposed.

Information sources

The following databases were searched for articles: Allied and Complementary Medicine Database (AMED; 1985 to present), British Nursing Index (BNI; 1985 to present), Cumulative Index to Nursing and Allied Health Literature (CINAHL; 1981 to present), Embase (1980 to present), Health Business Elite (EMBESCO Publishing, Ipswich, Massachusetts, USA), the Health Management Information Consortium (HMIC), MEDLINE from PubMed (1950 to present) and PsychINFO (1806 to present) using the NHS Evidence online library (REF), and the Cochrane Database of Systematic Reviews. The last search date was 25 September 2012.

Data collection process and study selection

Data were collected and tabulated independently by MJT, CM and CN in a manner guided by the Cochrane Handbook. Eligibility was decided independently, in a standardised manner and disagreements were resolved by consensus. If an abstract was not available from the database, the full-text reference was accessed.

Inclusion criteria for articles were as follows: published in peer-reviewed journal; describes PDSA method being applied to improve quality in a healthcare setting; published in English. Editorial letters, conference abstracts, opinion and audit articles were excluded from the study selection.

Data items

A theoretical framework was constructed by compartmentalising the key features of the PDSA method into observable variables for evaluation (table 2). This framework was developed in accordance with recommendations for PDSA use cited in the literature, describing the origins and theory of the method. Face validity of the framework was achieved through discussion among authors, with QI facilitators and at local research meetings.

Data were collected regarding application of the PDSA method in line with the theoretical framework. Other data collected included first author, year of publication, country, area of healthcare, use of PDSA or PDCA terminology, and use of MFI or FOCUS as

Table 1 Description of the plan–do–study–act (PDSA) cycle method according to developers and commentators

	Deming (1986)[25] Original description of the method relating to manufacturing	Langley (1996)[30] How the PDSA method may be adapted for use in healthcare contexts	Speroff and O'Connor (2004)[33] How the PDSA method is analogous to scientific methodology
Plan	Plan a change or test aimed at improvement	▶ Identify objective ▶ Identify questions and predictions ▶ Plan to carry out the cycle (who, when, where, when)	Formation of a hypothesis for improvement
Do	Carry out the change or test (preferably on a small scale)	▶ Execute the plan ▶ Document problems and unexpected observations ▶ Begin data analysis	Conduct study protocol with collection of data
Study	Examine the results. What did we learn? What went wrong?	▶ Complete the data analysis ▶ Compare data to predictions ▶ Summarise what was learnt	Analysis and interpretation of the results
Act	Adopt the change, abandon it or run through cycle again	▶ What changes are to be made? ▶ What will the next cycle entail?	Iteration for what to do next

supporting frameworks. Ratios were used to analyse the results regarding the majority of variables, and mean scores regarding data associated with length of study, length of PDSA cycle and sample size were also used for analysis. Data were analysed independently by MJT and CM. Discrepancies (which occurred in less than 3% of data items) were resolved by consensus.

Risk of bias in individual studies

The present review aimed to assess the reported application of the PDSA method and the results of individual studies were not analysed in this review.

Risk of bias across studies

Despite our review being focused on reported application, rather than success of interventions, it may still be possible that publication bias affected the results of this study. Research that used PDSA methodology, but did not yield successful results, may be less likely to get published than reports of successful PDSA interventions.

RESULTS

Study selection

A search of the databases yielded 942 articles. After removal of duplicates, 409 remained; 216 and 120

Table 2 Theoretical framework based on key features of the plan–do–study–act (PDSA) cycle method

Feature of PDSA	Description of feature	How this was measured
Iterative cycles	To achieve an iterative approach, multiple PDSA cycles must occur. Lessons learned from one cycle link and inform cycles that follow. Depending on the knowledge gained from a PDSA cycle, the following cycle may seek to modify, expand, adopt or abandon a change that was tested	▶ Were multiple cycles used? ▶ Were multiple cycles linked to one another (ie, does the 'act' stage of one cycle inform the 'plan' stage of the cycle that follows)? ▶ When isolated cycles were used were future actions postulated in the 'act' stage?
Prediction-based test of change	A prediction of the outcome of a change is developed in the 'plan' stage of a cycle. This change is then tested and examined by comparison of results with the prediction	▶ Was a change tested? ▶ Was an explicit prediction articulated?
Small-scale testing	As certainty of success of a test of change is not guaranteed, PDSAs start small in scale and build in scale as confidence grows. This allows the change to be adapted according to feedback, minimises risk and facilitates rapid change and learning	▶ Sample size per cycle? ▶ Temporal duration of cycles? ▶ Number of changes tested per cycle? ▶ Did sequential cycles increase scale of testing?
Use of data over time	Data over time increases understanding regarding the variation inherent in a complex healthcare system. Use of data over time is necessary to understand the impact of a change on the process or outcome of interest	▶ Was data collected over time? ▶ Were statistics used to test the effect of changes and/or understand variation?
Documentation	Documentation is crucial to support local learning and transferability of learning to other settings	▶ How thoroughly was the application of the PDSA method detailed in the reports? ▶ Was each stage of the PDSA cycles documented?

were further discarded following review of abstracts and full texts, respectively. Excluded articles did not apply the PDSA method as part of an empirical study or coincidently used the acronyms PDSA or PDCA for different terms, or were abstracts for conferences or poster presentations. A total of 73 articles met the inclusion criteria and were included in the review (see figure 2).

General study characteristics

Country of study

The retrieved articles describe studies conducted in the USA (n=46), the UK (n=13), Canada (n=3) Australia (n=3), the Netherlands (n=2) and one each from six other countries (see online supplementary appendix A for complete synthesis of results).

Healthcare discipline to which method was applied

This varied across acute and community care and clinical and organisational settings. The most common settings were those of pain management and surgery (six articles each).

Method terminology

Of the 73 articles identified, 42 articles used 'PDSA' as terminology and 31 referred to the method as 'PDCA'. Eight of these reported using the MFI. Thirty-one articles used 'PDCA' terminology, with 20 using the preceding FOCUS framework. One article described use of FOCUS and MFI. Over time there was an increase in the prevalence of PDSA use with PDCA use diminishing (see online supplementary figure S1). The earliest reported use of PDCA and PDSA in healthcare was 1993 and 2000, respectively.

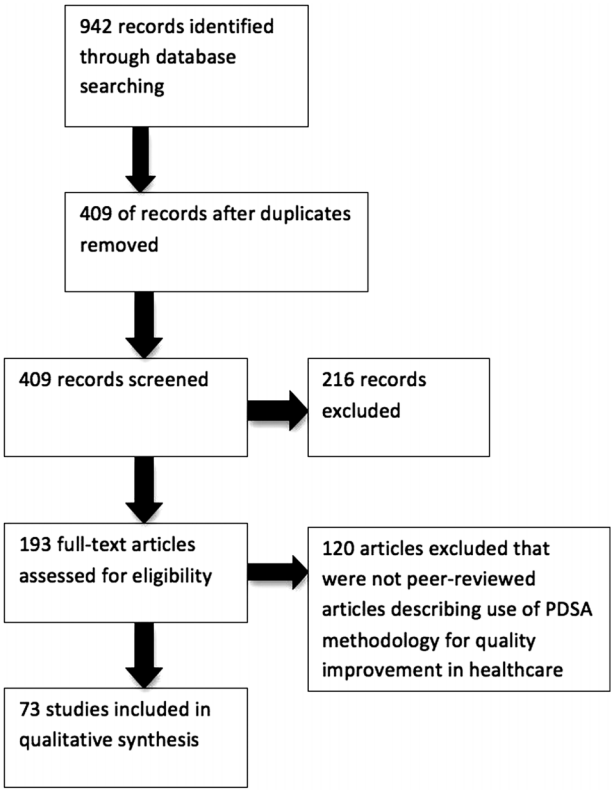

Figure 2 PRISMA diagram.

Documentation

The following four categories were used to describe the extent to which cycles were documented in articles (n=73): no detail of cycles (n=16); themes of cycles (but no additional details) (n=8); details of individual cycles, but not of stages within cycles (n=8); details of cycles including separated information on stages of cycles (n=41).

Analysis of articles against the developed framework was dependent on the extent to which the application of PDSA cycles was reported. Articles that provided no details of cycles or only themes of cycles were insufficient for full review and excluded for analysis against all features. Articles that provided further details of cycles completed (n=49) were included for analysis against the remaining four features of the framework. A full breakdown of findings can be viewed in online supplementary appendix B.

Application of method

Iterative cycles (n=49)

Fourteen articles described a sequence of iterative cycles (two or more cycles with lessons learned from one cycle linking and informing a subsequent cycle), 33 described isolated cycles that are not linked, and 2 articles described cycles that used PDSA stages in the incorrect order (in one article, one plan, one do, two checks and three acts were described, PDACACA[35]; a further study did not report use of a 'check' stage; PDA[36]) and are excluded from further review. Of the 33 articles that described non-iterative cycles, 29 reported a single cycle being used, and 4 described multiple, isolated (non-sequential) cycles. Although future actions are often suggested in articles that reported a single cycle, only three explicitly mentioned the possibility of further cycles taking place. A total of 13.6% (3/22) of PDCA studies described the application of iterative cycles compared with 44% (11/25) of PDSA studies describing the application of iterative cycles (see figure 3).

Prediction-based testing of change (n=47)

The aims of the cycles adhered to one of two themes: tests of a change; and collection or review of data without a change made. Of the 33 articles with single cycles, 30 aimed to test a change while 3 used the PDSA method to collect or review data. Of the 14 articles demonstrating sequential cycle use, 8 solely used their cycles to test change whilse5 began with a cycle collecting or reviewing data followed by cycles testing change. One article described a mixture of cycles testing changes and cycles that involved collection/review of data. Four of the 47 studies contained an explicit prediction regarding the outcome of a

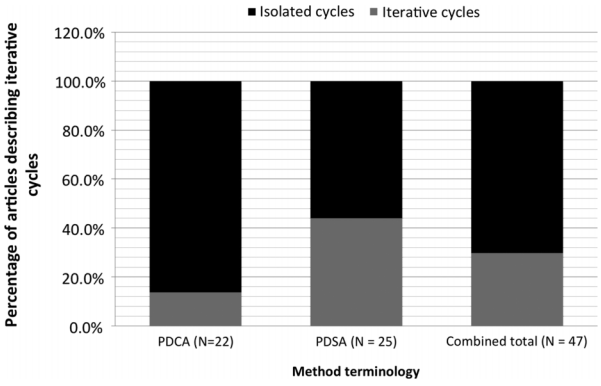

Figure 3 Iterative nature of cycles for all articles and split by plan–do–check–act and plan–do–study–act terminology.

change; all 4 aimed to test a change (see online supplementary table S1).

Small-scale testing (n=47)

Scale was assessed in three ways: sample size, duration and complexity. Sample size refers to quantity of observations used to measure the change; duration refers to the length of PDSA cycle application; and complexity refers to the quantity of changes administered per cycle.

Sample size

Patient data, staff data and case data were used as samples within PDSA cycles. Twenty-seven articles reported a sample size from at least one of their cycles. Twenty-one of these were isolated cycle studies with sample size ranging from 7 to 2079 (mean=323.33, SD=533.60). The remaining six studies reporting individual cycle sample sizes used iterative cycles; the sample size of the first cycles of these ranged from 1 to 34 (mean=16.75, SD=11.47). Two of these studies described the use of incremental sample sizes across cycles, three used non-incremental sample sizes across cycles, and one changed the type of sample. Of the eight iterative cycle articles that did not report individual cycle sample sizes, two did not differentiate sample sizes between cycles and instead gave an overall sample for the chain of cycles and six did not report sample size.

Duration

Reported study duration of isolated cycles ranged from 2 weeks to 5 years (mean=11.91 months, SD=12.81). Only five articles describing iterative cycles explicitly reported individual cycle duration. Individual cycle duration could be estimated from the total duration of the PDSA cycle chain and the number of cycles conducted, resulting in approximate cycle lengths ranging from three cycles in 1 day to one cycle in 16 months (mean=5.41 months, SD=4.80, see online supplementary figure S2). The total PDSA cycle duration for series of iterative cycles (first to last cycle of one chain) ranged from 1 day to 4 years (mean=20.38, SD=20.39 months).

Complexity

Twenty-two articles reported more than one change being tested within a single cycle. Of the articles describing iterative cycles, 42% administered more than one change per cycle compared with 48% of the articles describing non-iterative PDSA cycles.

Data over time (n=47)

All studies used a form of qualitative and quantitative data to assess cycles. Studies were categorised according to four types of reporting quantitative data: regular (n=15), three or more data points with consistent time intervals; non-regular (n=16), before and after or per PDSA cycle; single data point (n=8), a single data point after PDSA cycle(s); and no quantitative data reported (n=8). Of the 15 articles that used regular data, only 7 used monthly or more frequent data intervals (see online supplementary figure S3 for full frequency of regular quantitative data reporting). No studies reported using statistical process control to analyse data collected from PDSA cycles. Eleven included analysis of data using inferential statistical tests (five of these studies collected isolated data, six involved continuous data collection).

Of the eight articles that did not report any quantitative data, two reported that quantitative analyses had taken place but did not present the findings and six described the use of qualitative feedback only (one non-regular, five single data point). Qualitative data were gathered through a range of mechanisms from informal staff or patient feedback to structured focus groups.

DISCUSSION

PDSA cycles offer a supporting mechanism for iterative development and scientific testing of improvements in complex healthcare systems. A review of the historic development and rationale behind PDSA cycles has informed the development of a theoretical framework to guide the evaluation of PDSA cycles against use of iterative cycles, initial small-scale testing, prediction-based testing of change, use of data over time and documentation.

Using these criteria to assess peer-reviewed publications of PDSA cycles demonstrates an inconsistent approach to the application and reporting of PDSA cycles and a lack of adherence to key principals of the method. Only 2/73[37 38] articles demonstrated compliance with criteria in all five principles. Assessment of compliance was problematic due to the marked variation in reporting of this method, which reflects a lack of standardised reporting requirements for the PDSA method.

From the articles that reported details of PDSA cycles it was possible to ascertain that variation is

inherent not just in reporting standards, but in the conduct of the method, implying that the key principles of the PDSA method are frequently not followed. Less than 20% (14/73) of reviewed articles reported the conduct of iterative cycles of change, and of these, only 15% (2/14) used initial small-scale tests with increasing scale as confidence in the intervention developed. These results suggest that the full benefits of the PDSA method would probably not have been realised in these studies. Without an iterative approach, learning from one cycle is not used to inform the next cycle, and therefore it is unlikely that interventions will be adapted and optimised for use in a particular setting. Furthermore, large-scale cycles risk significant resource investment in an intervention that has not been tested and optimised within that environment and risk producing 'false' negatives.

Only 14% (7/47) of articles reported use of regular data over time at monthly or more frequent intervals, indicating a lack of understanding around the use of the PDSA method to track change within a 'live' system, and limiting the ability to interpret the results from the study. Cycles that included an explicit prediction of outcomes were reported in only 9% (4/47) of articles, suggesting that PDSA cycles were not used as learning cycles to test and revise theory-based predictions.

Overall these results demonstrate poor compliance with key principles of the PDSA method, suggesting that it is not being used optimally. The increasing trend in using PDSA (as opposed to 'PDCA') cycles in recent years, however, does seem to have been accompanied by an increase in compliance with some key principles, such as use of iterative cycles. Deming was cautious over the use of the 'PDCA' terminology and warned it referred to an explicitly different process, referring to a quality control circle for dealing with faults within a system, rather than the PDSA process, which was intended for iterative learning and improvement of a product or a process.[39] This subtle difference in terminologies may help to explain the better compliance with key methodological principles in studies that refer to the method as 'PDSA'.

One of the articles identified in the search included comments by the authors that the PDSA method should be 'more realistically represented',[40] as ineffective cycles can be 'abandoned' early on, making it needless to go through all four stages in each iteration. These comments may provide insight into an important potential misunderstanding of the PDSA methodology. Ineffective changes will result in learning, which is a fundamental principle behind a PDSA cycle. However minor this abandoned trial may have been, it can still be usefully described as a PDSA cycle. A minor intervention may be planned (P) and put into practice (D). A barrier may be encountered (S), resulting in a decision being made to retract the intervention, and to do something differently (A).

The theoretical framework presented in this paper highlights the complexity of PDSA cycles and the underpinning knowledge required for correct application. The considerable variation in application observed in the reported literature suggests that caution should be taken in interpreting results from evaluations in which PDSAs are used in a controlled setting and as a 'black box' of QI. This review did not compare the effectiveness of use to reported outcomes and therefore this study does not conclude whether better application of the PDSA method results in better outcomes, but instead draws on theoretical principles of PDSAs to rationalise why this would be expected. Prospective mechanistic studies exploring the effective application of the method as well as study outcomes would be of greater use in drawing conclusions regarding the effectiveness of the method. The framework presented in this paper could act as a good starting point for such studies.

The fact that only peer-reviewed publications were assessed in this study means that results may be affected by publication bias. This is anticipated both in terms of what is accepted for publication but also the level and type of detail that is requested and allowed in typical publications (eg, before and after studies are more common than presenting data over time and this may make these types of studies easier to publish). Though QI work may be easier to publish now through recent changes in publication guidelines,[27] possible publication outlets continue to be relatively limited.

To support systematic reporting and encourage appropriate usage, we suggest that reporting guidelines be produced for users of the PDSA method to increase transparency as to the issues that were encountered and how they were resolved. While PDSA is analogous to a scientific method, it appears to be rarely used or reported with scientific rigour, which in turn, inhibits perceptions of PDSA as a scientific method. Such guidelines are essential to increase the scientific legitimacy of the PDSA method as well as to improve scientific rigour or application and reporting. Although the SQUIRE guidelines make reference to the potential use of PDSA cycles, further support to users and teachers, and publication of this improvement method seems necessary. Consistent reporting of PDSA structure would allow meta-evaluation and systematic reviews to further build the knowledge of how to use such methods effectively and the principles to apply to increase chances of success.

It is clear from these findings that there is much room for improvement in the application and use of the PDSA method. Previous studies have discussed the influence of different context factors on the use of QI methods, such as motivation, data support infrastructure and leadership[20 22 41–43] Understanding how high-quality usage can be promoted and supported needs to become the focus of further research if such

QI methods are going to be used effectively in mainstream healthcare.

CONCLUSIONS

There is varied application and reporting of PDSAs and lack of compliance with the principles that underpin its design as a pragmatic scientific method. The varied practice compromises its effectiveness as a method for improvement and cautions against studies that view QI or PDSA as a 'black box' intervention.

There is an urgent need for greater scientific rigour in the application and reporting of these methods to advance the understanding of the science of improvement and efficacy of the PDSA method. The PDSA method should be applied with greater consistency and with greater accordance to guidelines provided by founders and commentators[25 30 44 45]

Acknowledgements The authors would like to thank Dr Thomas Woodcock for his valuable input into the theoretical framework and data analysis.

Contributors All listed authors qualify for authorship based on making one or more of the substantial contributions to the intellectual content: conceptual design (MJT, CM, CN, DB, AD and JR), acquisition of data (MJT, CM and CN) and/or analysis and interpretation of data (MJT, CM, CN and JR). Furthermore all authors participated in drafting the manuscript (MJT, CM, CN, DB, AD and JR) and critical revision of the manuscript for important intellectual content (MJT, CM, CN, DB, AD and JR).

Disclaimer This article presents independent research commissioned by the National Institute for Health Research (NIHR) under the Collaborations for Leadership in Applied Health Research and Care (CLAHRC) programme for North West London. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Competing interests The authors declare no conflict of interest.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Øvretveit J. *Does improving quality save money. A review of evidence of which improvements to quality reduce costs to health service providers*. London: The Health Foundation, 2009.
2 Walshe K, Freeman T. Effectiveness of quality improvement: learning from evaluations. *Qual Saf Health Care* 2002;11:85–7.
3 Schouten LMT, Hulscher MEJL, van Everdingen JJE, *et al*. Evidence for the impact of quality improvement collaboratives: systematic review. *BMJ* 2008;336:1491–4.
4 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.
5 Berwick DM. Developing and testing changes in delivery of care. *Ann Intern Med* 1998;128:651–6.
6 McCormack B, Kitson A, Harvey G, *et al*. Getting evidence into practice: the meaning of context. *J Adv Nurs* 2002;38:94–104.
7 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.
8 Oxman AD, Thomson MA, Davis DA, *et al*. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *CMAJ* 1995;153:1423.
9 Department of Health. *Report of the High Level Group (HLG) on clinical effectiveness*. London: Department of Health, 2007.
10 Greenhalgh T, Robert G, Macfarlane F, *et al*. Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Q* 2004;82:581–629.
11 Plsek PE, Wilson T. Complexity science: complexity, leadership, and management in healthcare organisations. *BMJ* 2001;323:746.
12 Damschroder LJ, Aron DC, Keith RE, *et al*. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci* 2009;4:50.
13 Powell AE, Rushmer RK, Davies HTO. *A systematic narrative review of quality improvement models in health care: NHS Quality Improvement Scotland*. 2009. Report No. 1844045242.
14 Boaden R, Harvey J, Moxham C, *et al*. *Quality improvement: theory and practice in healthcare*. NHS Institute for Innovation and Improvement, 2008.
15 Walshe K. Understanding what works—and why—in quality improvement: the need for theory-driven evaluation. *Int J Qual Health Care* 2007;19:57–9.
16 Shojania KG, Grimshaw JM. Evidence-based quality improvement: the state of the science. *Health Aff* 2005;24:138–50.
17 Auerbach AD, Landefeld CS, Shojania KG. The tension between needing to improve care and knowing how to do it. *N Engl J Med* 2007;357:608–13.
18 Ting HH, Shojania KG, Montori VM, *et al*. Quality improvement science and action. *Circulation* 2009;119:1962–74.
19 Pronovost P, Needham D, Berenholtz S, *et al*. An intervention to decrease catheter-related bloodstream infections in the ICU. *N Engl J Med* 2006;355:2725–32.
20 Benning A, Ghaleb M, Suokas A, *et al*. Large scale organisational intervention to improve patient safety in four UK hospitals: mixed method evaluation. *BMJ* 2011;342:d195.
21 Landon BE, Wilson IB, McInnes K, *et al*. Effects of a quality improvement collaborative on the outcome of care of patients with HIV infection: the EQHIV study. *Ann Intern Med* 2004;140:887–96.
22 Vos L, Duckers ML, Wagner C, *et al*. Applying the quality improvement collaborative method to process redesign: a multiple case study. *Implement Sci* 2010;5:19.
23 Grol R, Baker R, Moss F. Quality improvement research: understanding the science of change in health care. *Qual Saf Health Care* 2002;11:110–11.
24 Walshe K. Pseudoinnovation: the development and spread of healthcare quality improvement methodologies. *Int J Qual Health Care* 2009;21:153–9.
25 Deming WE. *Out of the crisis, 1986*. Cambridge, MA: Massachusetts Institute of Technology Center for Advanced Engineering Study xiii, 1991;507.
26 Øvretveit J. Understanding the conditions for improvement: research to discover which context influences affect improvement success. *BMJ Qual Saf* 2011;20(Suppl 1):i18–23.

27 Davidoff F, Batalden P, Stevens D, *et al*. Publication guidelines for quality improvement in health care: evolution of the SQUIRE project. *Qual Saf Health Care* 2008;17(Suppl 1): i3–9.
28 Ogrinc G, Mooney S, Estrada C, *et al*. The SQUIRE (Standards for Quality Improvement Reporting Excellence) guidelines for quality improvement reporting: explanation and elaboration. *Qual Saf Health Care* 2008;17(Suppl 1):i13–32.
29 Imai M. *The key to Japan's competitive success*. New York: McGraw-Hill, 1986.
30 Langley GJ. *The improvement guide: a practical approach to enhancing organizational performance*. 1st edn. San Francisco: Jossey-Bass Publishers, 1996.
31 Batalden P. *Building knowledge for improvement-an introductory guide to the use of FOCUS-PDCA*. Nashville, TN: Quality Resource Group, Hospital Corporation of America, 1992.
32 Moen R, Norman C. Evolution of the PDCA cycle. 2006.
33 Speroff T, O'Connor GT. Study designs for PDSA quality improvement research. *Qual Manag Health Care* 2004;13:17–32.
34 Liberati A, Altman DG, Tetzlaff J, *et al*. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009;6:e1000100.
35 Bader MK, Palmer S, Stalcup C, *et al*. Using a FOCUS-PDCA quality improvement model for applying the severe traumatic brain injury guidelines to practice: process and outcomes. *Reflect Nurs Leadersh* 2002;28:34–5.
36 Reid D, Glascott G, Woods D. Improving referral information in community mental health. *Nurs Times* 2005;101:34–5.
37 Lynch-Jordan AM, Kashikar-Zuck S, Crosby LE, *et al*. Applying quality improvement methods to implement a measurement system for chronic pain-related disability. *J Pediatr Psychol* 2010;35:32–41.
38 Varkey P, Sathananthan A, Scheifer A, *et al*. Using quality-improvement techniques to enhance patient education and counselling of diagnosis and management. *Qual Prim Care* 2009;17:205–13.
39 Moen R, Norman C. Circling back: clearing up the myths about the Deming cycle and seeing how it keeps evolving. *Qual Progress* 2010;42:23–8.
40 Tomolo AM, Lawrence RH, Aron DC. A case study of translating ACGME practice-based learning and improvement requirements into reality: systems quality improvement projects as the key component to a comprehensive curriculum. *Postgrad Med J* 2009;85:530–7.
41 Berwick DM. Developing and testing changes in delivery of care. *Ann Intern Med* 1998;128:651.
42 Benn J, Burnett S, Parand A, *et al*. Perceptions of the impact of a large-scale collaborative improvement programme: experience in the UK Safer Patients Initiative. *J Eval Clin Pract* 2009;15:524–40.
43 Parand A, Burnett S, Benn J, *et al*. Medical engagement in organisation-wide safety and quality-improvement programmes: experience in the UK Safer Patients Initiative. *Qual Saf Health Care* 2010;19:e44.
44 Berwick D. Broadening the view of evidence-based medicine. *Qual Saf Health Care* 2005;14:315–16.
45 Speroff T, James BC, Nelson EC, *et al*. Guidelines for appraisal and publication of PDSA quality improvement. *Qual Manag Health Care* 2004;13:33–9.