Evaluating the design and implementation of an adapted Plan-Do-Study-Act (PDSA) approach to improve health system bottlenecks in Nigeria

*Ejemai Amaize Eboreime[1,2], John Olajide Olawepo[3,4], Aduragbemi Banke-Thomas[5], Rohit Ramaswamy[6]

*Corresponding Author: ejemaim@gmail.com

Authors' Affiliations:

1. Department of Medicine, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada
2. Department of Planning, Research and Statistics, National Primary Healthcare Development Agency, Abuja, Nigeria
3. Caritas Nigeria, Abuja, Nigeria
4. Department of Environmental and Occupational Health, School of Public Health, University of Nevada Las Vegas, USA
5. Department of Health Policy, London School of Economics and Political Science, London, UK
6. Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill NC, USA

Abstract

Background: The Plan-Do-Study-Act (PDSA) cycle is fundamental to many quality improvement (QI) models. For the approach to be effective in the real-world, variants must align with standard elements of the PDSA at the design stage. The adapted design must also be implemented with fidelity. Gaps at the adaptation or implementation stages result in design and implementation failures respectively. Using Taylor's framework, this study evaluates the alignment between theory, design and implementation of the DIVA

(Diagnose-Intervene-Verify-Adjust) model, a PDSA variant adapted for Nigeria's health system performance improvement.

Methods: An iterative consensus building approach was used to develop a scorecard evaluating new conceptual indices of design and implementation of QI interventions (design and implementation index, defects and gaps). The scorecard assigned quantitative scores that measured the extent to which the reported or documented design or implementation activity aligned with Taylor's theoretical framework.

Results: Design (adaptation) scores were across all standard features indicating that design was well adapted to the typical PDSA. Conversely, implementation fidelity scores were only optimal with two standard features: prediction-based test of change and the use of data over time. The other features, use of multiple iterative cycles and documentation, had implementation gaps of 17% and 50% respectively. This may imply that observed effectiveness deficits of DIVA may be attributable to implementation rather than design gaps.

Conclusions: The effectiveness of QI interventions depend of good contextual adaptation and implementation fidelity such that core features (theory and internal logic) must not be compromised. This study demonstrates how both adaptation and implementation are important for success of QI interventions. It also presents an approach for evaluating other QI models using Taylor's PDSA assessment framework as a guide, which might serve to strengthen the theory behind future QI models, and provide guidance on their appropriate use.

Introduction

The Plan-Do-Study-Act (PDSA) cycle (and its concept of iterative tests of change) is fundamental to both the fields of Quality Improvement (QI) and Implementation Science

(IS). Many popular QI models such as the Model for Improvement, Lean and Six Sigma(1–4) incorporate some version of PDSA in their methodologies to test improvement solutions on a small scale before scaling up. In the field of implementation science, PDSA is similarly used to adapt implementation strategies to local contexts (5)(6). A recently proposed model integrating QI and IS, called the Model for Improvement and Implementation(7) explicitly situates PDSA as a process at the intersection of both fields, necessary for both context specific intervention development (QI) and adaptation of IS strategies.

Moreover, PDSA cycles align well with models for decision-making in complex systems with non-linear processes that are unpredictable and that display emergent behaviour(11). In such systems, a "probe, sense, respond" approach is recommended, where decision makers initiate small experiments to test how the system behaves before engaging in large scale analysis and intervention(12). As organizations such the World Health Organization and the World Bank recommended QI to address larger scale problems such as health systems strengthening to improve global quality of care(13), the use of PDSA is likely to continue to expand.

In summary, the use of PDSA to find local solutions to solve complex improvement and implementation problems offers great promise (14,15), and this approach can be considered to be a best practice in developing quality improvement interventions. However, even though PDSA provides a rigorous experimental approach to improvement, it has not been always been used effectively to bring about outcomes(4). The reason for this is there is no standard protocol or formula for using PDSA to improve. This in itself is not a shortcoming – by definition, since PDSA cycles are intended to drive learning based on generating hypotheses about how to change a system, a PDSA approach should be deeply embedded in the systems in which implementation and improvement work takes place and its particular application should depend on the context of the system as several researchers

have pointed out (8–10). However this assumes that those engaged in improvement or implementation adhere to the primary principles of the PDSA such as creating a theory of change, developing hypothesis, designing multiple tests of increasing complexity, collecting and analysing data and making appropriate updates to the theory of change. As demonstrated by Taylor and colleagues(3) who used an evaluation framework to assess gaps in the use of PDSA cycles used to improve health care, this is not often the case. Despite the theoretical promise of PDSA, it often fails in execution.

This being said, however, the PDSA approach still remains the best theoretical model for designing quality improvement interventions, and therefore it is appropriate to test other models designed to improve quality against the PDSA approach as a standard. While any quality improvement approach needs to be implemented well, there is an additional barrier to success if the theory underlying the approach is untested or flawed. At the same time, as mentioned above, PDSA is not an inviolable formula, and therefore needs to be adapted for use in various settings. In this paper, we evaluate a quality improvement model called Diagnose-Intervene-Verify-Adjust (DIVA) adopted in Nigeria and in other low and middle-income countries (LMICs) to improve planning processes in health systems. This study is nested within a larger research series which sought to investigate and improve the implementation of health system strengthening initiatives in Nigeria.

DIVA was developed, introduced, tested and implemented in several LMICs by United Nations Children's Fund (UNICEF) (17). The model seeks to improve processes such as routine planning cycles and the use of data ,by local stakeholders and policy/decision makers as suggested by the Lancet Global Health Commission on High Quality Health Systems (HQSS) (17–19). Typically, operational planning occurs annually. DIVA strengthens the annual planning cycles by introducing quarterly iterations for monitoring and improvement of the annual plan. Further, outputs from DIVA (such as proposed

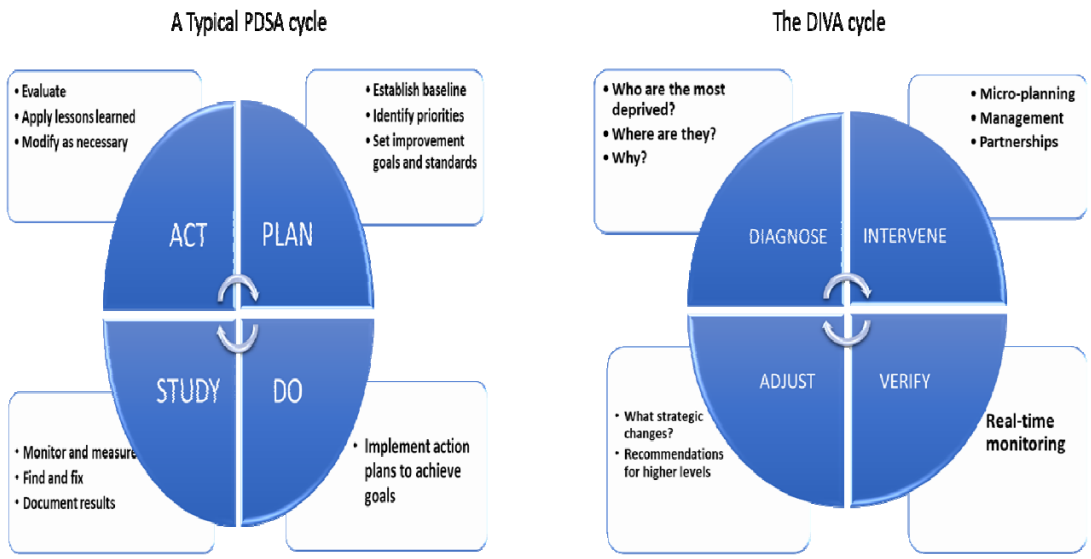
implementation strategies for health interventions) feed into subsequent operational planning cycles. DIVA also builds the capacity of district/local government health teams, communities and other local stakeholders to identify bottlenecks and barriers to the planning process. The model provides tools to address these barriers, take action and be accountable for service delivery within their jurisdiction towards improving effective coverage, quality and impact of health interventions (17,18).

Operationally, DIVA is a four-step model in which 'Diagnose' identifies bottlenecks to health interventions; 'Intervene' develops and implements action plans to resolve identified constraints; while 'Verify/Adjust' tracks the results of the action plans and recommends real-time data driven adjustments to the plan (Figure 1). Our research, published elsewhere (19,23–25) showed that the DIVA model can potentially improve health system performance, but that effective implementation is critical. Other evaluations in Nigeria, Ghana and Uganda have also demonstrated that the tools and methods of DIVA are effective in evidence-informed planning of health interventions (18–22) but outcomes have exhibited significant variability (19,23). To provide guidance to countries seeking to enhance the effectiveness of DIVA, it would be useful to assess both the theoretical underpinning and the implementation of DIVA. This will help practitioners and policy/decision makers evaluate whether the failure of DIVA to achieve outcomes comes from a failure in its design (i.e. the underlying premise of DIVA is flawed) or whether DIVA is a context-appropriate adaptation of the principles of PDSA, and failures in its performance can be attributed to poor implementation. If the latter is the case, strategies to improve the implementation of DIVA can be informed by the substantial body of knowledge that already exists about how to conduct PDSAs well. Specifically, in this paper, we seek to answer the following research questions:

1. To what extent can DIVA be considered to be an adaptation of PDSA?

2. To what extent does the implementation of DIVA in Nigeria exhibit fidelity to its design?

Figure 1: A typical PDSA cycle vs DIVA



Methods

Study context

Nigeria operates a decentralized three-tier health system governance structure. The federal level is responsible for overarching policy thrusts including programme design, the state ministries of health are largely responsible for subnational coordination and translation, while the local governments are mostly the implementers. DIVA was introduced nationwide in 2012, most states experienced implementation failure with only Kaduna sustaining implementation as at the time of this study.

Theoretical Framework

To answer the first research question, we used the theoretical framework developed by Taylor and colleagues (3) mentioned earlier to assess the quality of implementation of PDSA cycles in 47 studies in healthcare settings. This framework uses five features to test the theory and the practice of the approach. They are: the use of Iterative cycles,

prediction-based test of change, small-scale testing, use of data over time, and documentation of processes and outcomes for learning as described in Table 1. Given the use of the DIVA implementation for adjusting the annual plan on a quarterly basis, one of these characteristics (small scale testing) was not relevant for DIVA, because the steps of the DIVA process were applied to the same plan every quarter. We adapted this framework first to evaluate whether the design guidelines for DIVA include these key features. We used our findings from this evaluation to develop the assessment questions to evaluate the fidelity of implementation.

Table 1: Taylor's theoretical framework based on key features of the plan–do–study–act (PDSA) cycle(Source: (3))

Feature of PDSA	Description of feature
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
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
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
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
Documentation	Documentation is crucial to support local learning and transferability of learning to other settings

Data collection

Data for the overall implementation research series of which this work is a component were obtained from content analyses of 39 policy documents, 15 in-depth interviews(19,23) and an embedded process observation between 2012 and 2016 to identify the design and implementation processes of DIVA in Nigeria (Supplementary material 1). The documents were identified from the researchers' knowledge (EAE was an

embedded researcher) and interviews of four national level subject matter experts who contributed to the initial design and implementation processes. The documents include policy guidelines, a DIVA implementation guide and implementation reports obtained at both the design phase (national) and implementation phases (subnational). This sub-study uses data related to the implementation of DIVA in Kaduna state, since this is the only state still sustaining the implementation of DIVA since its design and deployment in 2012(19,23).

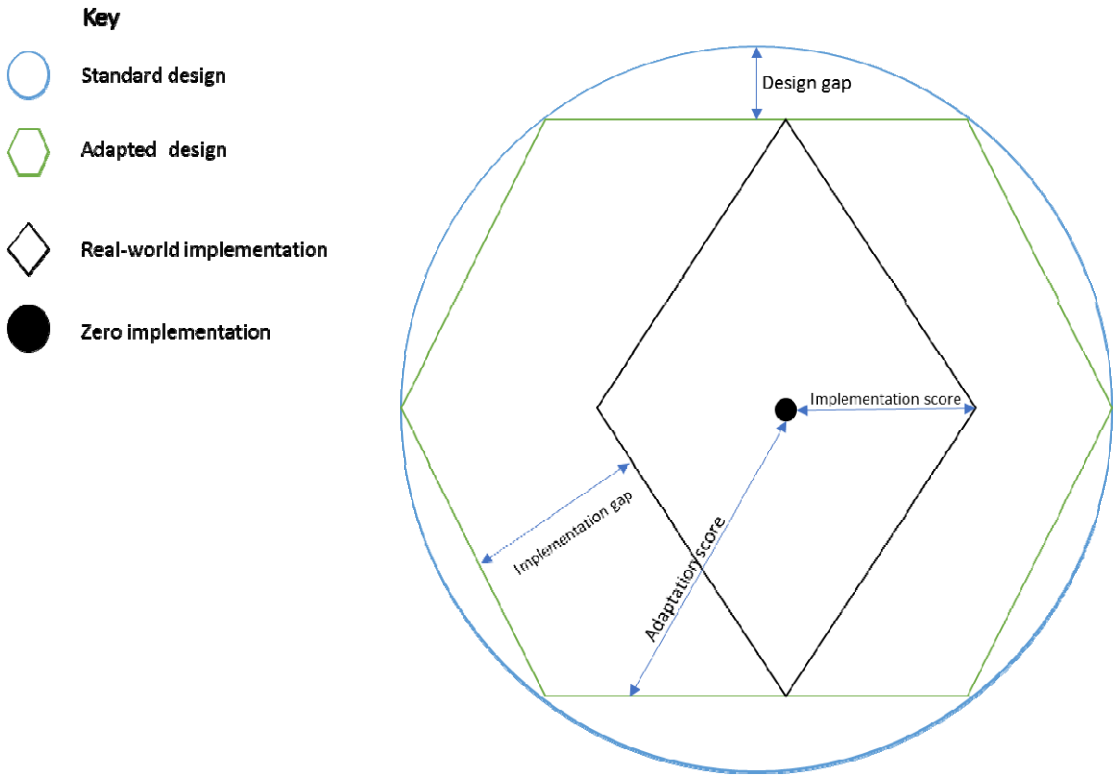
Analytic parameters

We created the following indicators to perform our evaluation:

1. *Adaptation score/index:* This is a measure of the extent to which the steps of DIVA is consistent with the applicable elements of PDSA. It assesses the conceptual similarity between the standard improvement model and its local adaptation.
2. *Implementation score/index:* This is a measure of the extent to which the adapted model was implemented in the real world. The implementation index may be considered to be a measure of adherence to how DIVA was intended to be used, which is a dimension of implementation fidelity (26).

Figure 2 is an illustration of these concepts.

Figure 2: Illustration of indicators



Taylor's framework used 13 questions to assess the implementation of PDSAs, which are paraphrased in Table 2. To apply this framework to the design of DIVA, an adapted guide of seven questions was created as shown in Table 3.

Table 2: Taylor's PDSA assessment questions

Feature of PDSA	Assessment questions
Iterative cycles	• 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	• Was a change tested? • Was an explicit prediction articulated?
Small-scale testing	• 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	• Was data collected over time? • Were statistics used to test the effect of changes and/or understand variation?
Documentation	• How thoroughly was the application of the PDSA method detailed in the reports? • Was each stage of the PDSA cycles documented?

Table 3: Questions and response scale for design evaluation

Feature of PDSA	Analyses Criteria	Response scale for each question (Scores: 0/1/2)
Iterative cycles	• Does the DIVA model incorporate the principle of multiple cycles? • Are multiple cycles linked to one another (i.e., does the 'adjust' stage of one cycle inform the 'diagnose' stage of the cycle that follows)?	No/partly/yes
Prediction-based test of change	• Does DIVA depend on well-defined theory of change? • Does the use of the model require explicit predictions that are tested in the "intervene" step?	No/partly/yes
Small-scale testing	• Not applicable	
Use of data over time	• Does the DIVA model require the collection and display of data over time?	No/partly/yes
Documentation	• Does the design require detailed reporting of the application of the DIVA model in each quarter? • Does the design require detailed reporting of each step of the DIVA process in each quarter?	No/partly/yes

The design evaluation helped to clarify the key features of the DIVA model and the extent to which these features were aligned with the key steps of a PDSA approach. Based on the results, an implementation evaluation questionnaire was created. This questionnaire is shown in Table 4.

Table 4: Questions and response scale for implementation evaluation in Kaduna

Feature of PDSA	Analyses Criteria	Response scale for each question (Scores: 0/1/2)
Iterative cycles	• Were multiple cycles used? • Were multiple cycles linked to one another (i.e., does the 'adjust' stage of one cycle inform the 'diagnose' stage of the cycle that follows)?	No/partly/yes
Prediction-based test of change	• Was a theory of change articulated prior to the implementation? • Was an explicit prediction made in each step?	No/partly/yes
Use of data over time	• Was data collected over time?	No/partly/yes
Documentation	• Was the application of the DIVA method detailed in the reports? • Was each stage of the DIVA cycle(s) documented?	No/partly/yes

Data analysis

To assess the level of alignment between the data and standard features, we created a scorecard. The scorecard assigned quantitative scores that measured the extent to which the reported or documented design or implementation activity aligned with the theoretical framework. An iterative consensus building approach was used to discuss and finalize scores. This process mirrors the approach used in evaluating the deployment and installation processes of DIVA in the other sub-studies of this series(19,23).

Each question in the design and implementation was evaluated on a 3-point ordinal scale (No = 0, Partly = 1, Yes = 2). 'No' and 'Yes' were attributed to absolute non-compliance and absolute compliance respectively, whereas a 'partly' was assigned to partial compliance. Three evaluators with expertise on the DIVA implementation and knowledge of the PDSA independently reviewed the qualitative data and assigned a score to each question. Inter-rater variances were resolved using a consensus building approach. The rules were that disparities between each rater's scores on any question must not exceed one point. Thus, any rating disparity greater than one point was discussed and re-evaluated in the next

round. The underlying assumption was that a one-point difference may represent random variation which could be harmonized by calculating a mean. However, a two-point disparity may be indicative of a significant interpreter difference which may need more clarification(19,23). The mean was used as the final score for each question. Each question carried an equal weight per standard feature. The mean score for each feature was calculated. Having calculated the mean scores for each feature, composite scores for each feature were created by dividing by a factor of 2 expressed in percentage values ranging from 0% (Zero implementation) to 100% (Full implementation/Full design alignment).

Results

Table 5 presents the scores for design and implementation of DIVA on each evaluation criterion. Our design evaluation indicated that all the applicable characteristics of PDSA were part of the DIVA design, and therefore DIVA can be considered to be an adaptation of the PDSA model. We could therefore use the same questions we used to evaluate the design for the assessment of the implementation.

Table 6 presents the description of the adaptive design and implementation of DIVA in relation to the standard features of a typical PDSA. Table 6 also shows the gaps on each respective feature. Both tables highlight where design failure and/or implementation failure may exist, which have implications on the effectiveness of the intervention.

Conversely, implementation fidelity scores were only optimal with prediction-based test of change and the use of data over time. Multiple cycles were implemented but not appropriately linked to each other. Further the application of the PDSA methods, including the four stages, were not well documented in the reports. This may imply that any observed poor performance of DIVA may have been caused by implementation failure.

Table 5: Scores on analyses criteria

PDSA Feature	Analyses criteria	Response score (mean) for each question

		Design *(Adaptation Score)*	*Implementation* *(Fidelity score)*
Iterative cycles	1. Multiple cycles	2.00	2.00
	2. Cycles linked to one another	2.00	1.33
	Composite score (%)	***100***	***83***
Prediction-based test of change	3. Test of change	2	2
	4. Explicit prediction articulated	2	2
	Composite score (%)	***100***	***100***
Use of data over time	5. Data collected over time	2	2
	Composite score (%)	***100***	***100***
Documentation	6. Application of the PDSA method detailed in the reports	2	1
	7. Documentation of each stage of the PDSA cycles	1.67	1
	Composite score (%)	***92***	***50***

Table 6: Summary of findings and assessments

Feature of PDSA	DIVA Design	Design gap (%)	DIVA implementation	Implementation gap (%)	Assessment
Iterative cycles	DIVA is intended to be implemented using multiple iterative cycles which are interlinked such that the 'Verify/Adjust' phases of one cycle feed into the 'Diagnose' phase of the next.	0	DIVA is implemented using multiple (quarterly) cycles which are linked. However, in practice, we observed that while in many cycles the 'Verify/Adjust' phases of one cycle feed into the 'Diagnose' phase of the next, there were occasions where such interlinks were interrupted. This is often due to contextual influences such as donor interests, leadership and administrative deficiencies(19) .	17	Optimal theory-design consonance with mild implementation gap
Prediction-based test of change	The model tests changes in six supply and demand side determinants of health system performance. It includes tools and measures for explicitly predicting change in these determinants(27)..	0	During the 'Diagnose/Intervene' phases, action plans to address health system bottlenecks are made which include quantified coverage improvement targets to be met following implementation. However, given that the intervention was introduced and implemented at scale ab initio, the sample size remains the same in every DIVA cycle. Thus, implementation aligns with the adapted design but not with the PDSA theory.	0	Optimal theory-design-implementation consonance
Use of data over time	The design of DIVA is to build on existing data systems. Thus, routine data from the Health Management Information System is used over time to analyse change. Given the capacity of implementers (mostly district health managers and other stakeholders from the local community), DIVA guidelines recommend use of simple descriptive statistical evaluations rather than complex analytic processes such as statistical control charts	0	Data was collected over time and analysed using time series charts and bar graphs. The implementation wholly complies with design.	0	Zero design defect with optimal implementation fidelity
Documentation	The implementation guide contains processes and tools for documenting activities, plans and lessons learned from implementation for each stage of the implementation.	8	There was detailed documentation of the 'Diagnose/Intervene' phases. This was less so with the 'Verify/Adjust' phases which were observed to be the weakest link in the implementation chain.	50	Good theory-design consonance with poor implementation fidelity

Discussion

This study interrogates the conceptual underpinnings of DIVA and its implementation in Nigeria's health system. The implementation science literature recognizes that interventions are successful only, when adapted to the local contexts(28,29). However, the process of modifying an intervention must not compromise its core features (theory and internal logic), which are most likely to produce the main expected outcomes. The adaptation therefore must preserve the mechanism through which the desired change in outcomes is achieved, while at the same time assuring that the interventions is a good fit to the problem being addressed.(28,30,31). Design gaps occur when the core components of the intervention are compromised during the adaptation. Implementation defects occur when implementers fail to adhere to the execution of the design as specified by guidelines. This gap in implementation fidelity commonly results from unstable implementation environment such as evolving policies and discretionary action of implementers, among others(19,23).

Our principal finding in this study is that DIVA remains true to the principles of the PDSA approach, while excluding some components that are not germane to the objective of strengthening the planning process. One such component is the PDSA hallmark of conducting small tests of change. This is mainly because typical PDSAs are used within organizational context such as to improve services within a clinic. DIVA, however, was designed for higher (systems) level improvement such as complex district health systems which coordinate governance and services across several communities and facilities located within defined geographical boundaries(4,17,32). Typically, in organizational contexts, PDSAs start small in scale and expand scale as confidence grows. However DIVA, in Nigeria, was designed to review healthcare processes that had already been implemented and changes in each cycle were intended to address reallocation of resources or priorities rather than making changes to the processes themselves. In fact, small tests of change may disrupt existing

routine systems, compromising the goal of integrated health system strengthening with implications on sustainability of the intervention (17,19,24).

Another variation in DIVA compared to PDSA is the method of analysing data over time. While temporal trends are reviewed the method of analysis is descriptive, primarily using before and after bar charts. This prevents a more rigorous analysis of process variation. The adaptation was necessary because DIVA is designed to be an easy-to-use tool for decision making by health care managers/policymakers who are not trained in rules for interpreting run chart rules.

We also found that the implementation of DIVA varied partly from the design. Complete implementation fidelity (adherence to design) was observed in 2 of the 4 features of PDSA (prediction-based tests of change and the use of data over time). Implementation gaps were observed in the remaining 2 features: iterative cycles and documentation. Typical PDSA cycles are iterative, with lessons learned from one cycle feeding into the 'Plan' phase of cycle that follows. While this is also the design of DIVA, in practice, interruptions to these cycle linkages were observed.

In an another article in this study series(19), using Kaplan's Model for Understanding Success in Quality (MUSIQ)(37), we described how contextual factors such as weak political will, administrative inefficiencies and donor interests partly compromised the implementation of DIVA, particularly the 'Verity/Adjust' phases. These deficiencies had implications on linkages between cycles on some occasions. The detailed documentation observed with the 'Diagnose/Intervene' phases, was not replicated with the 'Verify/Adjust' phases, thus reinforcing the untoward effect of these contextual factors with implications on effectiveness.

As noted in previous studies(19,24,32), DIVA is potentially effective as it was able to improve isolated performance indicators of vertical programmes such as malaria management and measles vaccination, but did not significantly improve integrated health system performance due to poor implementation of the model. This study provides further insights into possible reasons for suboptimal effectiveness of DIVA. First, a minor 17% gap in implementing iterative cycles, then a

major 50% gap in documentation. Poor attention paid to documentation could compromise the quality of other processes given that lessons learned from one cycle may not completely feed into subsequent cycles. This phenomenon may potentially create some form of cyclical redundancies, compromising the effectiveness of the intervention. Thus, to improve the effectiveness of DIVA in this context, more attention must be paid to documentation of processes, outcomes and lessons learned. Beyond the context in which this study was carried out, this study demonstrates that the standard principles of quality improvement (and other) interventions should inform both the design and implementation of the interventions in order to ensure effective outcomes.

The design of this study may subject the findings to some interpreter bias and subjectivity. Despite these apparent biases, we view the measures and approaches used here as practical proxies for evaluating the alignment of theory, design and implementation in the real-world.

Conclusion

This study demonstrates how both adaptation and implementation are important for success of QI interventions (43–46). It also presents an approach for evaluating other QI models using Taylor's PDSA assessment framework as a guide, which might serve to strengthen the theory behind future QI models, and provide guidance on their appropriate use. This will be particularly important for improvement models like DIVA, designed for multi-contextual improvements, where different adaptations may be necessary in different countries and where implementation challenges may vary by setting.

References

1. Norman CL, Nola TW, Moen R, Provost LP, Nolan KM, Langley GJ. The improvement guide: a practical approach to enhancing organizational performance [Internet]. Jossey-Bass; 2013 [cited 2018 Nov 7]. Available from: https://books.google.ca/books/about/The_Improvement_Guide.html?id=kE4aEnZgBO8C&printsec=frontcover&source=kp_read_button&redir_esc=y#v=onepage&q&f=false
2. Schroeder RG, Linderman K, Liedtke C, Choo AS. Six Sigma: Definition and underlying theory. J Oper Manag [Internet]. 2008 Jul 1 [cited 2018 Nov 7];26(4):536–54. Available from:

https://www.sciencedirect.com/science/article/abs/pii/S0272696307000897

3. 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. 2013;14(23):290–8.

4. Reed JE, Card AJ. The problem with Plan-Do-Study-Act cycles. BMJ Qual Saf. 2016;25:147–52.

5. Powell BJ, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. Implement Sci [Internet]. 2015 Dec 12 [cited 2019 Jun 29];10(1):21. Available from: http://implementationscience.biomedcentral.com/articles/10.1186/s13012-015-0209-1

6. Fixsen DL, Naoom SF, Blase KA, Friedman RM, Wallace F. Implementation Research: A Synthesis of the literature [Internet]. The National Implementation Research Network,: University of South Florida, Louis de la Parte Florida Mental Health Institute.; 2005 [cited 2019 Jun 29]. Available from: https://nirn.fpg.unc.edu/sites/nirn.fpg.unc.edu/files/resources/NIRN-MonographFull-01-2005.pdf

7. Ramaswamy R, Johnson JK, Hirschhorn LR. Intergrating implementation science approaches into continuos quality improvement. In: Johnson JK, Sollecito WA, editors. McLaughlin & Kaluzny's Continuous Quality Improvement in Health Care. 5th ed. Burlington, MA: Jones & Bartlett Learning; 2018.

8. Leis JA, Shojania KG. A primer on PDSA: executing plan-do-study-act cycles in practice, not just in name. BMJ Qual Saf [Internet]. 2017 Jul 1 [cited 2019 Jun 29];26(7):572–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27986900

9. Ramaswamy R, Reed J, Livesley N, Boguslavsky V, Garcia-Elorrio E, Sax S, et al. Unpacking the black box of improvement. Int J Qual Heal care J Int Soc Qual Heal Care [Internet]. 2018 Apr 20 [cited 2019 Jun 29];30(suppl_1):15–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29462325

10. Øvretveit J. Understanding the conditions for improvement: research to discover which context influences affect improvement success. BMJ Qual Saf [Internet]. 2011 Apr 1 [cited 2019 Jun 29];20 Suppl 1(Suppl 1):i18-23. Available from: http://qualitysafety.bmj.com/lookup/doi/10.1136/bmjqs.2010.045955

11. WHO. Systems Thinking for Health Systems Strengthening [Internet]. Geneva; 2009 [cited 2019 Apr 8]. Available from: https://apps.who.int/iris/bitstream/handle/10665/44204/9789241563895_eng.pdf;jsessionid=25F32F936EDA7D35E0BAF578E7510123?sequence=1

12. Snowden DJ, Boone ME. A leader's framework for decision making. A leader's framework for decision making. Harv Bus Rev [Internet]. 2007 Nov [cited 2019 Jun 29];85(11):68–76, 149. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18159787

13. Kieny M-P, Evans TG, Scarpetta S, Kelley ET, Klazinga N, Forde I, et al. Delivering quality health services: a global imperative for universal health coverage [Internet]. Geneva, Switzerland: World Health Organization, OECD, and International Bank for Reconstruction and Development/The World Bank; 2018 [cited 2019 Apr 8]. Available from: http://apps.who.int/bookorders.

14. Muczyk JP. A Systems Approach to Organizational Effectiveness: The Alignment of Critical

Organizational Dimensions with Selected Business/Competitive Strategies. J Comp Int Manag [Internet]. 2004;7(1). Available from: https://journals.lib.unb.ca/index.php/JCIM/article/view/454

15. Sales A, Smith J, Curran G, Kochevar L. Models, strategies, and tools. J Gen Intern Med [Internet]. 2006 Feb [cited 2018 Nov 16];21(S2):S43–9. Available from: http://link.springer.com/10.1007/s11606-006-0274-x

16. Roser C. The many Flavors of the PDCA [Internet]. 2016 [cited 2018 Nov 7]. Available from: https://www.allaboutlean.com/pdca-variants/

17. UNICEF. The guidebook: Strengthening district management capacity for planning, implementation and monitoring for results with equity. New York: UNICEF; 2012.

18. O'Connell TS, Sharkey A. Reaching Universal Health Coverage through District Health System Strengthening: Using a Modified Tanahashi Model Sub-nationally to Attain Equitable and Effective Coverage. New York: UNICEF; 2013.

19. Eboreime EA, Nxumalo N, Ramaswamy R, Eyles J. Strengthening decentralized primary healthcare planning in Nigeria using a quality improvement model: how contexts and actors affect implementation. Health Policy Plan [Internet]. 2018;1–14. Available from: https://academic.oup.com/heapol/advance-article-abstract/doi/10.1093/heapol/czy042/4993753

20. Baker U, Peterson S, Marchant T, Mbaruku G, Temu S, Manzi F, et al. Identifying implementation bottlenecks for maternal and newborn health interventions in rural districts of the United Republic of Tanzania. Bull World Heal Organ. 2015;93:380–389.

21. Baker U, Okuga M, Waiswa P, Manzi F, Peterson S, Hanson C. Bottlenecks in the implementation of essential screening tests in antenatal care: Syphilis, HIV, and anemia testing in rural Tanzania and Uganda. Int J Gynecol Obstet. 2015;130:S43–50.

22. Henriksson DK, Fredriksson M, Waiswa P, Selling K, Peterson SS. Bottleneck analysis at district level to illustrate gaps within the district health system in Uganda. Glob Heal Action. 2017;10(1):1327256.

23. Eboreime EA, Eyles J, Nxumalo N, Eboreime O, Ramaswamy. R. Implementation process and quality of a primary healthcare system improvement initiative in a decentralized context: A retrospective appraisal using the Quality Implementation Framework. Int J Heal Plann Mgmt. 2018;

24. Eboreime E, Nxumalo N, Ramaswamy R, Ibisomi L, Eyles J. Implementing a Model for Improving Integrated Primary Healthcare Planning and Performance: An Effectiveness Evaluation of DIVA in Kaduna, Nigeria. In: Cochrane Database of Systematic Reviews [Internet]. 2017. Available from: http://www.cochranelibrary.com/dotAsset/7c625b5f-377d-4a7f-9468-6d09a1cce668.pdf

25. Eboreime EA, Abimbola S, Obi FA, Ebirim O, Olubajo O, Eyles J, et al. Evaluating the sub-national fidelity of national Initiatives in decentralized health systems: Integrated Primary Health Care Governance in Nigeria. BMC Health Serv Res [Internet]. 2017 Dec 21 [cited 2018 Nov 23];17(1):227. Available from: http://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-017-2179-2

26. James Bell Associates. Evaluation brief: Measuring Implementation Fidelity [Internet]. Arlington, VA; 2009 [cited 2018 Nov 10]. Available from: https://www.jbassoc.com/wp-content/uploads/2018/03/Measuring-Implementation-Fidelity.pdf

27. NPHCDA. Institutionalization of the Primary Health Care Reviews in Nigeria: Status and Progress. Abuja: National Primary Health Care Development Agency; 2013.

28. McKleroy VS, Galbraith JS, Cummings B, Jones P, Harshbarger C, Collins C, et al. Adapting evidence-based behavioral interventions for new settings and target populations. AIDS Educ Prev. 2006;

29. Galbraith JS, Stanton B, Boekeloo B, King W, Desmond S, Howard D, et al. Exploring implementation and fidelity of evidence-based behavioral interventions for HIV prevention: Lessons learned from the focus on kids diffusion case study. Heal Educ Behav. 2009;

30. Pérez D, Van der Stuyft P, Zabala M del C, Castro M, Lefèvre P. A modified theoretical framework to assess implementation fidelity of adaptive public health interventions. Implement Sci [Internet]. 2015 Dec 8 [cited 2018 Nov 16];11(1):91. Available from: http://implementationscience.biomedcentral.com/articles/10.1186/s13012-016-0457-8

31. Sundell K, Beelmann A, Hasson H, von Thiele Schwarz U. Novel Programs, International Adoptions, or Contextual Adaptations? Meta-Analytical Results From German and Swedish Intervention Research. J Clin Child Adolesc Psychol [Internet]. 2016 Nov 11 [cited 2018 Nov 16];45(6):784–96. Available from: https://www.tandfonline.com/doi/full/10.1080/15374416.2015.1020540

32. Eboreime EA, Nxumalo N, Ramaswamy R, Ibisomi L, Ihebuzor N, Eyles J. Effectiveness of the Diagnose-Intervene-Verify-Adjust (DIVA) model for integrated primary healthcare planning and performance improvement: An embedded mixed methods evaluation in Kaduna state, Nigeria. BMJ Open. 2019;In Press.

33. Gilson L. Health Policy and Systems Research: A Methodology Reader. Geneva: Alliance for Health Policy and Systems Research, World Health Organization; 2012.

34. Harn B, Parisi D, Stoolmiller M. Balancing Fidelity with Flexibility and Fit: What Do We Really Know about Fidelity of Implementation in Schools? Except Child [Internet]. 2013 Jan 1 [cited 2018 Nov 14];79(3):181–93. Available from: http://journals.sagepub.com/doi/10.1177/001440291307900204

35. Castro FG, Barrera M, Martinez CR. The cultural adaptation of prevention interventions: resolving tensions between fidelity and fit. Prev Sci [Internet]. 2004 Mar [cited 2018 Nov 16];5(1):41–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15058911

36. Provost LP, Murray SK. Understanding Variation Using Run Charts. In: The Health Care Data Guide. San Fransisco, CA: Jossey-Bass; 2011. p. 25–66.

37. 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.

38. Lipsky M. Street-Level Bureaucracy. Classics of public administration. 1980. Ch. 40.

39. Lipsky M. Street Level Bureaucracy: Dilemmas of the Individual in Public Services [Internet]. The Russel Sage Foundation. 2010. 1–11 p. Available from: https://steveadami.myefolio.com/Uploads/Street-Level Bureaucracy.pdf%5Cnhttp://www.jstor.org/stable/10.7758/9781610446631

40. Sumah AM, Baatiema L, Abimbola S. The Impacts of Decentralization on Health-Related Equity: A Systematic Review of the Evidence. Health Policy (New York). 2016;120(10):1183–92.

41. Hogwood B, Gunn L. Policy Analysis for the Real World . Oxford: Oxford University Press; 1984.

42. Sabatier P, Mazmanian D. The Conditions of Effective Implementation: Guide to Accomplishing Policy Objectives. Policy Anal. 1979;

43. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. Implement Sci [Internet]. 2007 Dec 30 [cited 2018 Nov 16];2(1):40. Available from: http://implementationscience.biomedcentral.com/articles/10.1186/1748-5908-2-40

44. Borrelli B, Sepinwall D, Ernst D, Bellg AJ, Czajkowski S, Breger R, et al. A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. J Consult Clin Psychol [Internet]. 2005 Oct [cited 2018 Nov 16];73(5):852–60. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16287385

45. Stirman SW, Miller CJ, Toder K, Calloway A. Development of a framework and coding system for modifications and adaptations of evidence-based interventions. Implement Sci [Internet]. 2013 Dec 10 [cited 2018 Nov 16];8(1):65. Available from: http://implementationscience.biomedcentral.com/articles/10.1186/1748-5908-8-65

46. El-Sadr WM, Philip NM, Justman J. Letting HIV Transform Academia — Embracing Implementation Science. N Engl J Med [Internet]. 2014 May [cited 2018 Nov 16];370(18):1679–81. Available from: http://www.nejm.org/doi/10.1056/NEJMp1314777

47. Lahariya C, Menabde N. Evidence to implementation continuum for universal health coverage. Lancet Infect Dis [Internet]. 2015;15(3):250–1. Available from: http://dx.doi.org/10.1016/S1473-3099(14)71090-4

48. Kruk ME, Gage AD, Arsenault C, Jordan K, Leslie HH, Roder-DeWan S, et al. High-quality health systems in the Sustainable Development Goals era: time for a revolution. Lancet Glob Heal [Internet]. 2018 Nov 1 [cited 2018 Nov 7];6(11):e1196–252. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30196093

49. Fineberg H V. A Successful and Sustainable Health System — How to Get There from Here. N Engl J Med [Internet]. 2012 Mar 15 [cited 2018 Nov 16];366(11):1020–7. Available from: http://www.nejm.org/doi/abs/10.1056/NEJMsa1114777

50. Brooks H, Harris K, Bee P, Lovell K, Rogers A, Drake R. Exploring the potential implementation of a tool to enhance shared decision making (SDM) in mental health services in the United Kingdom: a qualitative exploration of the views of service users, carers and professionals. Int J Ment Health Syst [Internet]. 2017 Dec 28 [cited 2018 Nov 16];11(1):42. Available from: http://ijmhs.biomedcentral.com/articles/10.1186/s13033-017-0149-z