OXFORD

Quality in Practice

Learning from the design, development and implementation of the Medication Safety Thermometer

PARYANEH ROSTAMI[1], MAXINE POWER[2], ABIGAIL HARRISON[2], KURT BRAMFITT[2], STEVE D. WILLIAMS[1,3], YOGINI JANI[4,5], DARREN M. ASHCROFT[1,6], and MARY P. TULLY[1]

[1]Manchester Pharmacy School, University of Manchester, Manchester Academic Health Sciences Centre (MAHSC), Oxford Road, Manchester, UK, [2]Haelo, Salford Royal NHS Foundation Trust, Stott Lane, Salford, UK, [3]University Hospital of South Manchester NHS Foundation Trust, Southmoor Road, Wythenshawe, Manchester, UK, [4]Pharmacy Department, University College London Hospitals NHS Foundation Trust, Euston Road, London, UK, [5]UCL School of Pharmacy, Brunswick Square, London, UK, and [6]NIHR Manchester Primary Care Patient Safety Translational Research Centre, University of Manchester, Oxford Road, Greater Manchester, UK

Address reprint requests to: Paryaneh Rostami, Manchester Pharmacy School, University of Manchester, Manchester Academic Health Sciences Centre (MAHSC), Oxford Road, Manchester, UK. Tel: +44-0161-275-8363; Fax: +44-0161-275-2416; E-mail: paryaneh.rostami@manchester.ac.uk

Abstract

Quality issue: Approximately 10% of patients are harmed by healthcare, and of this harm 15% is thought to be medication related. Despite this, medication safety data used for improvement purposes are not often routinely collected by healthcare organizations over time.

Initial assessment: A need for a prospective medication safety measurement tool was identified.

Choice of solution: The aim was to develop a tool to allow measurement and aid improvement of medication safety over time. The methodology used for the National Health Service (NHS) Safety Thermometer was identified as an approach. The resulting tool was named the 'Medication Safety Thermometer'.

Implementation: The development of the Medication Safety Thermometer was facilitated by a multidisciplinary steering group using a Plan, Do, Study, Act (PDSA) method. Alpha and beta testing occurred over a period of 9 months. The tool was officially launched in October 2013 and continued to be improved until May 2016 using ongoing user feedback.

Evaluation: Feedback was gained through paper and online forms, and was discussed at regular steering group meetings. This resulted in 16 versions of the tool. The tool is now used nationally, with over 230 000 patients surveyed in over 100 NHS organizations. Data from these organizations are openly accessible on a dedicated website.

Lessons learned: Measuring harm from medication errors is complex and requires steps to measure individual errors, triggers of harm and actual harm. PDSA methodology can be effectively used to develop measurement systems. Measurement at the point of care is beneficial and a multidisciplinary approach is vital.

Key words: medication errors, harm, measurement, PDSA

Quality issue

Approximately 1 in 10 patients are harmed by healthcare [1–3]. It is thought that 15% of these harms are associated with medication-related incidents [3], which remain the single largest source of repetitive healthcare error [4]. Despite these statistics, there is a lack of tools to routinely measure medication safety in healthcare organizations over time.

Initial assessment

Previous research indicates that harm to patients involving medication is often preventable [5]. Therefore, interventions aimed at reducing medication errors have the potential to make a substantial difference to improving patient safety [3]. In order to prevent medication errors and reduce the risks of harm, organizations must detect and measure errors [6], and analyse the information collected to understand what is happening and why. Medication errors are currently under-reported, often because they are corrected before reaching the patient [7]. Nonetheless, the small proportion of errors that do reach the patient may potentially cause severe harm, including death [8].

Most medication safety data are obtained through either research studies or, more commonly, voluntary reporting. The latter has been the mainstay of learning from medication safety incidents within the UK's National Health Service (NHS). However, voluntary reporting underestimates error [8–12], and even though the number of reports has continually increased since the National Reporting and Learning System (NRLS) was established [13], the numbers and quality of reports from individual organizations remain variable [12]. Data collected for research studies are more reliable than voluntary reports and can be used for learning 'and' measuring. However, such data collection methods are rarely used in practice, as they are time-consuming, labour-intensive and expensive [14, 15]. Hence, they are not sustainable or practical in the long term for busy healthcare environments.

Previous literature has suggested that it is time to review and update data collection methods with 'fresh eyes' [10]. Therefore, NHS England commissioned Haelo (an independent innovation and improvement science centre hosted by Salford Royal Hospital NHS Foundation Trust) [16] to explore whether the NHS Safety Thermometer approach could be applied to collect medication safety data, which could be used for learning and measurement, and to support organizations in decreasing the risk of harm from medication error over time.

Choice of solution

The NHS Safety Thermometer, developed in 2010 as part of a national safety improvement programme in England, is a tool that has enabled organizations to collect data on common harms on 1 day each month and to track improvement over time [17]. The original NHS Safety Thermometer measures harm from pressure ulcers, falls, venous thromboembolism and urine infections in catheterized patients. It also provides a composite measure of 'harm free' care, defined as the absence of the measured harms [18].

Following the national rollout of the Safety Thermometer specialist groups and frontline teams identified that this methodology could be used for additional patient safety issues. Four 'next generation' Safety Thermometers were developed for maternity, mental health, children and young people and, the subject of this paper, the Medication Safety Thermometer (MedsST).

A national multidisciplinary steering group was commissioned by NHS England and facilitated by Haelo. This group initiated the development of the MedsST, an instrument that aimed to support local measurement of harm from medication, and related improvement. The MedsST also needed to allow for data to be aggregated and assessed at regional and national levels, in line with the NHS Outcomes Framework, which requires a focus on the 'incidence of medication errors causing serious harm' [19].

The steering group adhered to the Safety Thermometer design principles, that the tool would: have clinically valid definitions, be efficient, be used wherever the patient is treated, provide immediate access to data over time, measure all harm experienced by the patient regardless of preventability, measure harm at the patient level enabling a composite measure of 'harm free' care and be easy to aggregate [18, 20].

Approach to implementation

A plan for developing the MedsST was constructed using a driver diagram framework (Fig. 1). Alpha-testing (from January 2013 to
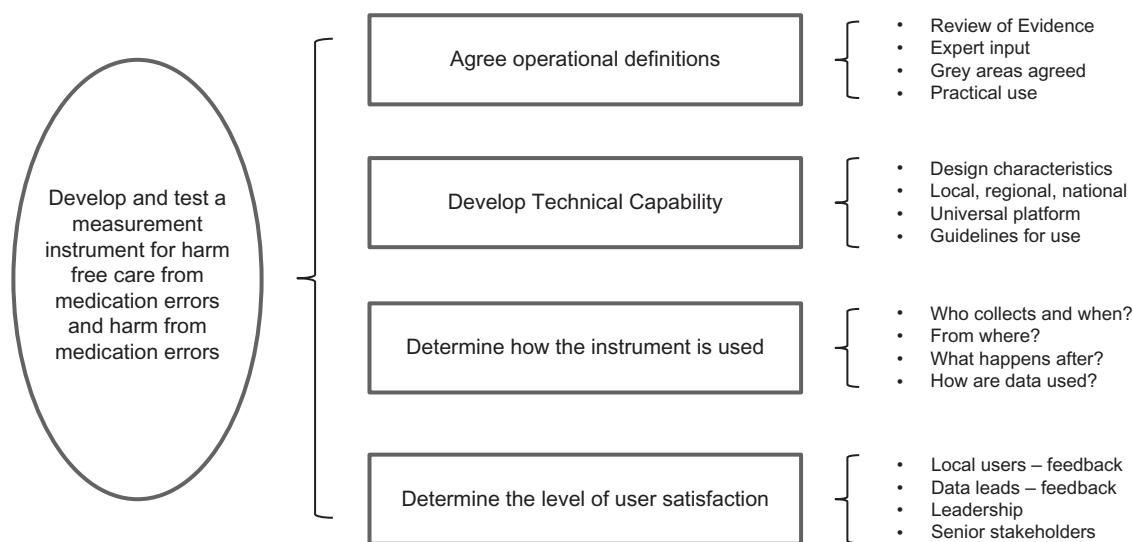

Figure 1 Project Plan Framework—adapted from Power *et al.* [18].

March 2013) involved very early tests with eight alpha-sites in Greater Manchester and one alpha-site in London. Beta-testing ('the pilot phase') ran from April 2013 to September 2013. In addition, a 6-month regional Commissioning for Quality and Innovation (CQUIN) payment target was introduced from April 2013 to March 2014 to incentivize the Greater Manchester organizations to continue testing the tool. CQUIN targets are used as financial levers in addition to baseline funding for organizations in the NHS [21]. Participation in the beta testing phase was open to all organizations and led to 43 sites joining the pilot phase. The national rollout of the MedsST occurred in October 2013 and collection of feedback for improving the MedsST has continued.

Agreeing on operational definitions

It was decided to focus on harm due to high-risk medicines and develop measures of harm related to errors involving these (Tables 1–3).

Technical development

Initially, a paper-based prototype instrument was tested in alpha-sites; data were entered into a spreadsheet and e-mailed to Haelo. Monthly feedback was used to design the next iteration of the form.

Guidance for instrument use and data collection

Safety Thermometers have been designed to be used as part of routine healthcare, in acute and community settings to encourage continuity of care [22].

The NHS Safety Thermometer data collection is made at the point of care by a healthcare professional who reviews the patient's documentation and performs a physical examination where necessary. For example, the presence of a pressure ulcer, when the skin is inspected, is classed as a 'harm' in the original Safety thermometer. Early discussions between the steering group and the first tests of change revealed difficulties with this methodology when measuring harm from medicines. In particular, harm from medication may not be apparent at the time of review. This 'uncoupling' of the error from the harm required a stepped approach to measuring error and harm. This characteristic is unique to the MedsST and differentiates it from the original NHS Safety Thermometer.

Guidance documents were developed to support teams in testing the tool [20]. It was recommended that Step 1 data (process errors) were collected by nurses, and Step 2 data (triggers of harm) by pharmacists and nurses together. The third step involved a multidisciplinary 'huddle' to discuss if harm had actually occurred. In hospital settings, this would involve at least the nurse, pharmacist and junior doctor looking after the patient on the ward, and in the community this may involve a phone call from a nurse or pharmacist to the GP overseeing the patient's care.

Feedback and satisfaction with the instrument

The main methods of feedback to the steering group included: monthly meetings via a virtual conferencing platform, monthly surveys and regular phone calls and e-mails with volunteers who had tested the tool. The data collected using the tool, and the feedback and satisfaction data were discussed regularly within the steering group. Once changes were agreed, a new version of the tool was circulated. The development team hypothesized that, with increased satisfaction and ease of use, the number of patients surveyed and the number of organizations using the tool would increase.

Ethics

Data were collected for NHS service improvement rather than research; therefore, research ethics committee approval was not required. No patient identifiable data were collected. The data were collected monthly as part of routine care, therefore causing no burden to patients and the burden on the staff was evaluated using surveys and identified as minimal.

PDSA testing and instrument refining

Safety Thermometers have been developed using improvement science, in particular Plan, Do, Study, Act (PDSA) cycles, which provide a structure for iterative testing of changes to improve quality systems. Each measure and definition included was developed using numerous cycles.

To date (May 2016), there have been 16 versions of the MedsST with multiple small changes per version, with each version tested for 2–3 months. Version 16 has now been used for over a year, with no current plans for Version 17. Version 16 includes subversions for acute and community settings. The most recent version of the MedsST is available from www.safetythermometer.nhs.uk [23].

Agreeing on operational definitions

In order to measure outcomes of harm from medication, proxy measures were identified but early tests revealed that this approach alone would not provide clinically valid definitions of harm. Attributing harm to medication error was complex due to several factors. For example, there may be some time between an error occurring and the harm being apparent (such as omission of an anticoagulant) or it may be difficult to establish if the error alone had caused the harm (such as confusion due to opiate overdose, which could also be due to a competing cause, for instance, a severe infection). To ensure only a manageable proportion of the most high-risk patients were triggering Step 2, each operational definition was refined several times (Table 1). In addition, process measures that may indicate potential harm were also focused on including medication omissions, allergy status and medicines reconciliation completion.

Technical development

As the number of users increased, an online version using SurveyMonkey® replaced the spreadsheet method. Once feedback indicated that the form was suitable, online platforms were developed, including a dedicated web tool and an application that could be used on phones or tablets, which also allowed offline data collection. This reduced the data collection time and anecdotal feedback suggests most organizations take <2 minutes per patient (excluding interruptions and when Step 3 is triggered).

Recommendations for use and observations of use

Through testing, the steering group agreed a recommended sample for data collection: all patients on five surgical wards and five medical wards per hospital, on the same day each month and all patients (up to 200) in community settings. However, organizations could choose to scale up their collection sample over time. Suggested dates for data collection were published in the MedsST guidance [20] and were used by the majority of organizations.

Feedback from surveys and observations revealed that data have been collected by a variety of professionals (Tables 1 and 2). Anecdotal feedback suggested in some, but not all organizations, Steps 1 and 2 data were regularly analysed at ward and senior

Table 1 Changes in operational definitions over time (using Version 1, 8 and 16 for illustration). The most recent version on the MedsST is available from www.safetythermometer.nhs.uk

Measure/ step	Step	Version 1	Version 8	Version 16[a]
Allergy status documented	1	Was the medicine allergy status documented in the clinical record in this care setting (including no known allergies)?	Was the medicine allergy status documented in the patient's clinical record in this care setting (including no known drug allergies) e.g. on prescription or Medication Administration Review and Request (MARR) chart?	Same as version 8.
Medicines reconciliation initiated		Were all medications documented as reconciled within 24 hours of admission to this care setting?	Was medicines reconciliation for all medicines undertaken (started) within 24 hours of admission to this care setting?	Same as version 8.
Omission of medication		Had the patient had an omitted dose of any medication in the last 24 hours?	Had the patient had an omitted dose of any medication in the last 24 hours (excluding food supplements)?	Was the patient on any of the following medications: anticoagulants, opioids, insulin or anti-infectives (excluding food supplements & oxygen). If so, had any of these (or 'any other prescribed medicines') been omitted and for what reason? Reasons: Patient refused, outstanding reconciliation, medicine not available, route not available, patient absent at medication round, not documented or other
Omission of high-risk medication		Not included in Version 1	Were omitted doses (see above) any of the following: anticoagulant, insulin, opiate, anti-infective (antibiotics, antifungals, antivirals and antimalarials)?	
Inclusion criteria and triggers for harm from anticoagulants	2	All anticoagulants were included. Triggers: If the patient had a bleed, vitamin K administered or INR outside the following limits—<2, higher than 6	Heparin, LMWH, Warfarin and NOACs (excluding VTE prophylaxis) were included. Triggers: A bleed of any kind or VTE, administration of vitamin K, protamine or clotting factors e.g. octaplex, or an INR greater than 6 or APTT ratio greater than 4	Heparin, LMWH, Warfarin and NOACs (excluding VTE prophylaxis) were included. Triggers: A bleed of any kind or VTE, or administration of vitamin K, protamine or clotting factors e.g. octaplex
Inclusion criteria/ trigger for harm from opiates		All opiates were included. Triggers: Was the prescribed dose more than 50% higher than the previous dose? Was the prescribed starting dose usual for the route to be used? Was the patient showing any symptoms of an overdose or common side-effects?	All opiates included. Triggers: Common complications (including sedation, respiratory depression, confusion), administration of naloxone, increased early warning score or respiratory rate below 12 breaths per minute	Opioids excluding oral codeine, dihydrocodeine and tramadol. Triggers: Administration of Naloxone, respiratory rate is <8 breaths per minute
Inclusion criteria/ trigger for harm from sedatives		All sedatives were included. Triggers: If the patient had any history of dementia or delirium, had administration of Flumazenil or had had a fall	The following injectable sedatives were included: midazolam, lorazepam, diazepam, clonazepam. Triggers: Common complications of over sedation (hypotension, delirium, respiratory depression, reduced Glasgow Coma Score), administration of Flumazenil or increased early warning score	IV or SC sedatives: Midazolam, Lorazepam, diazepam, clonazepam were included Triggers: Common complications (see version 8) or administration of Flumazenil
Inclusion criteria and Triggers for harm from insulin		All insulin included. Triggers: If an intravenous syringe or a non-insulin syringe used for insulin preparation or administration? Was the patient's insulin unit dose and frequency clearly documented? Had the patient had any omitted doses of insulin in the last 24 hours?	All insulin was included. Triggers: Common complications (capillary blood sugar <4 mmol/L, symptoms: anxiety confusion, extreme hunger, fatigue, irritability, sweating or clammy skin, trembling hands), administration of IV dextrose or glucagon, or diabetic ketoacidosis or hyperosmolar hyperglycaemic state	All insulin included. Triggers: Common complications: capillary blood sugar <4 mmol/L or symptoms of hypoglycaemia, administration of IV dextrose or glucagon or diabetic ketoacidosis or hyperosmolar hyperglycaemic state

Table continued

Table 1 Continued

Measure/ step	Step	Version 1	Version 8	Version 16[a]
	3	If any of the above (harms) were identified, the team was to refer to Step 3, which involved a MDT root cause analysis to determine whether there was harm from medication error. *The form for Step 3 was to be confirmed*	If triggered, organizations were recommended to perform an MDT huddle. This would involve a discussion with the doctor, nurse and pharmacist taking care of the patient to ascertain whether harm had occurred. *The form for Step 3 was to be confirmed*	If triggered, organizations were recommended to perform an MDT huddle using *a supplementary page for facilitation*. The form recorded: who was involved with the MDT huddle, their roles and involvement with the patient's care. If harm had occurred, it also recording of the level of harm based on NPSA harm scale [8] and learning and outcomes of Step 3

[a]Version 16 consists of two subversions; acute and community. The acute subversion has been used for illustration purposes in this table, as it is used more predominantly.

management levels. For example, at some sites, MedsST was analysed to see which wards were showing most improvement. Additionally, not all organizations have used Step 3 and, when it has been used, there have been challenges with completing it at the point of care. In hospitals, for example, the patient surveyed may have left the ward by the time the huddle could be arranged. In those organizations that have used Step 3, it has encouraged voluntary incident reporting of harm to allow local investigation and identification, in turn promoting a culture of safety [24].

Feedback and satisfaction with instrument

Virtual conference meetings allowed users and developers to discuss and suggest improvements based on testing and learning. It was often highlighted that organizations were experiencing similar problems, for example, problems with high numbers of referrals from Step 1 to Step 2, due to codeine-based medication post-surgery (Table 2). There has been a steep increase in the number of hospitals using the Web tool and, more recently, the mobile application. Some hospitals have stopped using the MedsST. Anecdotal feedback suggests some hospitals have stopped using the MedsST due to lack of time and resources.

Setting

The MedsST has predominantly been used in secondary care hospitals; however, has also been used in community settings, including community hospitals, domiciliary care and nursing homes.

Lessons learned

Repeated PDSA cycles confirmed that attributing harm to medication error at a single time point is highly complex [4, 9], and it is necessary to use different steps to observe errors, triggers of harm and actual harm. The original plan was for the MedsST to involve a simple bedside point of care audit, similar to the NHS Safety Thermometer, which focused on harm as an outcome of medication error. However, the resulting instrument extends this and focuses on both potential *and* actual harm due to medication [10].

Adverse events are often multifactorial and it can be challenging to attribute harm to a medication [9]. By using a number of steps, this complexity was partially addressed, as only those patients that triggered potential harm indicators were investigated for actual harm. Previous tools, such as the IHI global trigger tool, have demonstrated the need for using numerous steps [4]. Although various steps are required, trigger tools must be as time- and resource-efficient as possible [25, 26]. A previous study, using the IHI global trigger tool for Adverse Drug Events (ADEs), reported that 20 minutes was required to screen a single patient's record, and the study required a doctor and pharmacist to spend one-half to one day per site retrospectively reviewing a random sample of charts that contained triggers [26]. The study used a 39-item ADE trigger tool and only nine of the 39 triggers used accounted for 94.4% of ADEs detected [26]. Focusing review on triggers more predictive of an adverse event, as the MedsST does, is a better use of resources and may be more likely to improve patient safety [25, 26].

PDSA methodology can be effectively used to develop a measurement system

As previous research has suggested, it is occasionally necessary to simply 'get on with it' to assess the outcomes and the methods by which we can learn and improve a system [27]. However, this should not be a 'quick and dirty process' and requires an efficient plan, which may be constantly revised [27]. As indicated in Table 1, some definitions were expanded and then retracted to the original definition over several versions because, until changes are tested, it is difficult to know their impact.

The overarching aim to develop a tool to allow measurement and aid improvement of medication safety over time was achieved. Feedback from organizations using Step 3 suggests the MedsST triggers have been useful to identify actual harm from high-risk medications, and may have contributed to increased incident reporting and encouraged multidisciplinary teamwork. However, the focus on actual harm was expanded to also include potential harm (using process measures) and some organizations have focused on potential harm only. Although the focus of the MedsST may differ to what was originally planned, the PDSA cycle approach is quality driven and learning from 'failed' tests is equally as important as learning from success, and often the most valuable lessons are learnt from failure, which enables course correction [28].

Measuring medication error and harm at the point of care is beneficial and a multidisciplinary approach is vital. The data collected and analysed provide a baseline to establish whether further improvement work impacts medication safety and if it is maintained [12]. The simple act of collecting data should not be underestimated and, as data are mainly collected at the point of care by the multidisciplinary team (MDT), this process alone may help to improve safety culture and awareness at a local level [29].

Table 2   Summary of PDSA cycles involved in developing Step 1

Plan	Step 1 would focus on error potential and be completed for all patients. It involves collecting demographic data regarding the patient, their medications, omissions and drug allergy documentation, and identify patients taking any of the four classes of medicines reported to the UK's NRLS as most likely to cause death and severe harm between 2005 and 2010 [8] if not prescribed, dispensed or administered appropriately: anticoagulants, injectable sedatives, insulin and opiates. Step 1 was the first stage of a proxy harm measurement system, and if a patient was on any of the aforementioned medications, Step 2 would be triggered (Table 3). Prediction: Step 1 would be collected by nurses and they would be comfortable using Step 1, if they were not, this would be highlighted in user feedback. Step 1 would refer a small proportion of patients who were on high-risk drugs to allow a manageable 'snapshot' of the level of harm from medication errors. Testing would confirm if the high-risk drug class definitions were appropriate for this, or whether they were under-sensitive or oversensitive
Do	Testing was gradually scaled up as the tool improved, based on feedback from each test. First, very small tests on one patient were undertaken, then one ward, multiple wards, alpha sites (nine hospitals), beta sites (43 hospitals), and finally all sites continued to feedback after the official testing phase ended. Frontline teams collected data and fed back their experience of using the form, for example, how easy data collection was and how long it took. Feedback was collected at regular intervals and assessed at biweekly steering group meetings, facilitated by the development team, to ascertain the most efficient method of collecting data. Feedback platforms included online forums and surveys, verbal reports and meetings. Observations were also undertaken to better understand the impact of problems, such as the order of questions in regards to ease of data collection
Study	The prediction was not entirely correct as some definitions were not appropriate, for various reasons highlighted below. The main learning points from testing were: • In addition to nurses, Step 1 data were also collected by pharmacists, preregistration pharmacists, clinical auditors and healthcare assistants • The wording of some questions in Step 1 was not relevant or appropriate for community care settings • The conceptual order of the questions did not enable the easiest and quickest collection of data, and was not necessarily taking <10 minutes per patient. The order, although seemingly logical, actually meant that most teams were looking for data in one place for the first question, moving somewhere else on the record to get the data for the next questions, and then going back to their original source for data for the third question • A large number of patients who were at a very low risk of harm were triggering Step 2 due to being on opioids. Qualitative feedback from testers indicated that they felt that patients on low doses or low risk opioids were going through to Step 2 unnecessarily, as there was very little risk of harm occurring and that this was very time-consuming and disengaging. This was mostly due to low dose codeine, usually compounded with paracetamol as co-codamol. This had often been prescribed as 'when required' and not always necessarily used by the patient • There was a need identified for an appropriate denominator to understand the proportion of omissions of high-risk medication. In early versions, data about the number of patients who had had omissions of high-risk medications was collected, however, data about the number of patients who were on the high-risk medications initially were not collected. This meant that users were using the whole population of patients surveyed as a denominator, as opposed to the population of patients on a high-risk medication, leading to sampling bias
Act	Actions taken in response to study of tests included: • Development of a community subversion, in which the wording was amended to make Step 1 more relevant to practice in community • Individual definitions were revised to make the tool more practical. For example, it was decided to exclude oral codeine, dihydrocodeine and tramadol, as the problems they were causing in data collection outweighed the benefit of keeping them. The concept of the MedsST is to give a *snapshot* of harm and it is not possible to include all medications, even though they all have the potential to cause harm • The form was reordered so that questions were grouped together around the likely source of information. Multiple PDSAs were conducted to redesign all of the questions, thus *increasing ease* of data collection and reducing the time required • A new question was introduced about the number of patients on critical medication
Unresolved issues	Feedback from users has highlighted that the wording remains unsuitable for community settings; further refining is required. Some organizations are still taking longer than 10 minutes to survey each patient; further investigation is required to explore the potential reasons for this

Although more complex than anticipated, it was possible to collect similar medication safety data in different settings. Testing revealed that the MedsST needed to be different in community and acute settings, as the resources in each setting are considerably different.

Lessons learned from the data

The focus on medicines reconciliation helps to improve continuity of care between healthcare settings [22]. Some of the medicines reconciliation rates observed from the national MedsST data are similar to rates from previous research. For example, national MedsST data show that ~73% of patients are having medicines reconciliation within 24 hours (Fig. 2a). This figure is similar to findings from a previous study evaluating medicines reconciliation rates in one UK

hospital ($n = 70\%$) [30]. The aforementioned data, however, indicate that the standard of 95%, previously suggested by the National Institute for Health and Care Excellence for medicines reconciliation within 24 hours [30, 31], is not being met. Organizations should be encouraged to use the MedsST when assessing further improvement work to increase medicine reconciliation completion rates.

Other MedsST data have varied from data collected in previous research. For example, MedsST data suggest 22% of patients have at least one dose omission per day (Fig. 2b) and 5.7% of patients experience an omission of a critical medication (Fig. 2c). These omission data are lower than omission rates from previous research studies, which estimate that 80% of patients have an omitted dose [32]. This variance is may be due to a number of factors such as whether studies measure the rate of omissions of doses,

Table 3 Brief summary of PDSA cycles involved in developing Step 2

Plan	The plan was for Step 2 to be completed for all patients triggered in Step 1 due to receiving one or more of the high-risk medications. Step 2 was the second stage of a proxy harm measurement system mentioned in Table 2. For example, if a patient identified in Step 1 as being on an anticoagulant, and then in Step 2 it was established that they had had a bleed, these two factors together would be classed as a harm. Similarly to Step 1, PDSA methodology was used to develop the measures so that data would be simple to collect, the burden of data collection is minimal and measures are easily understood and clinically valid. Prediction: Step 2 would be collected by nurses and pharmacists together, who would be comfortable with identifying the harms listed. The definitions used for the triggers of harm from medication error would be appropriate for identifying potential harms. Feedback from users would identify if the definitions used were appropriate or not
Do	Data were collected on all patients identified in Step 1 who were on any of the drugs from the four high-risk classes. These patients would go through to Step 2 where a nurse and pharmacist would collect data on whether the triggers of harm had occurred. First, very small tests on one patient were undertaken, then one ward, multiple wards, alpha sites (nine hospitals), beta sites (43 hospitals), and finally all sites continued to feedback after the official testing phase ended. Frontline teams collected data and fed back on their experience of using the form, for example, how easy data collection was and how long it took. Feedback was collected at regular intervals and assessed at biweekly steering group meetings, facilitated by the development team to ascertain the most efficient method of collecting data. Feedback platforms included *online* forums and surveys, verbal reports and meetings. Observations were also undertaken to explore the feedback and better understand the impact of problems, such as the order of questions regarding ease of data collection
Study	The prediction that teams would be comfortable with identifying harms was not entirely correct, and a need for revisions was confirmed as each definition went through multiple PDSA cycles. Qualitative feedback from several PDSAs indicated that the attempt to define harm related to medication errors was extremely complex and that the measures were not representative of actual medication harm. Some of the key individual issues identified were:
	• Instead of Step 2 being collected by a nurse and pharmacist as recommended, it was mainly collected solely by a pharmacist or in some cases solely by a nurse. In addition, other professionals, such as pharmacy technicians were collecting data for Step 2
	• Certain terminology was not understood by all data collectors depending on their professional background. For example, one of the triggers of harm from injectable sedatives included assessing the patient's 'early warning score'. However, feedback indicated that most of the data for Step 2 was being collected by the pharmacy team who, as opposed to nurses, were not familiar with this term. In addition, different organizations had different definitions of 'early warning scores' and not all organizations used them
	• Attributing a harm to a medication error using a trigger was difficult. It is absolutely vital to have multidisciplinary discussions to ascertain the likelihood of whether harm has occurred due to a medication error. In many cases, it was not possible to be certain that a harm was only related to medication. There could be other factors to consider, making it difficult to decide if a harm could be classed as a medication harm
Act	Definitions of each individual measure were refined and tested through PDSA cycles numerous times, resulting actions included:
	• Refinement of Step 2 to exclude certain triggers. For example, the use of an 'early warning score' as a trigger of harm was removed in version 8
	• There was strong consensus from the steering group and the testers that, in order to understand if a harm was caused by a medication, there needed to be a multidisciplinary discussion involving nurses, doctors and pharmacists when collecting data on medication harms. This lead to official testing of Step 3, in volunteering organizations, after the launch date (October 2014) when Step 2 was more refined and stable
Unresolved issues	The argument for continuing to include 'when required' opioids. Some harm may be missed, as harm may occur from low dose opioids. Many organizations have not been using Step 3 and referring harms from Step 2 for MDT discussion. Further qualitative exploration is required to find out why organizations are not using Step 3

or the rate of patients with omitted doses [33]. Other reasons include: studies examining different drug classes or whether data are collected from electronic prescribing and administration systems, which have the potential to impact omissions and identifying the rate of omissions [34]. Therefore, standardization of how omissions are measured is required and in the context of the MedsST, local improvement has been encouraged, rather than comparison between organizations.

Lessons learned from the data provide many opportunities for further improvement work, which can be presented in a variety of ways. The Pareto Chart in Fig. 2d shows that 80% of critical omissions were with only two of the four critical risk medications (anti-infectives and opioids). Therefore, the most parsimonious approach of reducing omissions may be to focus improvement efforts on reducing omissions of anti-infectives and opioids in the first instance.

Data collected by the MedsST are presented in run charts on the website [23]. This allows users to study variation in data over time and understand the impact of changes with minimal mathematical complexity [35]. The run charts make data accessible and understandable to a range of different healthcare professionals. Special cause variation occurred in August 2014 (Fig. 2b and c), when there was a decrease in the number of omissions coincident with the introduction of Version 16. This was due to a change in the way omissions data were collected and the operational definitions that were first implemented in Version 16 (Table 1). To address this, further guidance and support was provided to organizations. This was done by producing additional guidance and providing support via group WebExes, and one-to-one phone support to certain organizations. The data stabilized from September 2014 onwards, suggesting that challenges with data collection had been somewhat resolved.

Over 230 000 patients have been surveyed using the MedsST in over 100 organizations (June 2016). As the number of patients surveyed using the MedsST has increased, the denominator for each of the medication safety measures is larger, which has reduced variation. A decrease in variation occurred in early 2015, as illustrated in Fig. 2a–c; in January 2015 the number of patient surveyed was 7425 compared to 5271 patients in December 2015. Furthermore, the
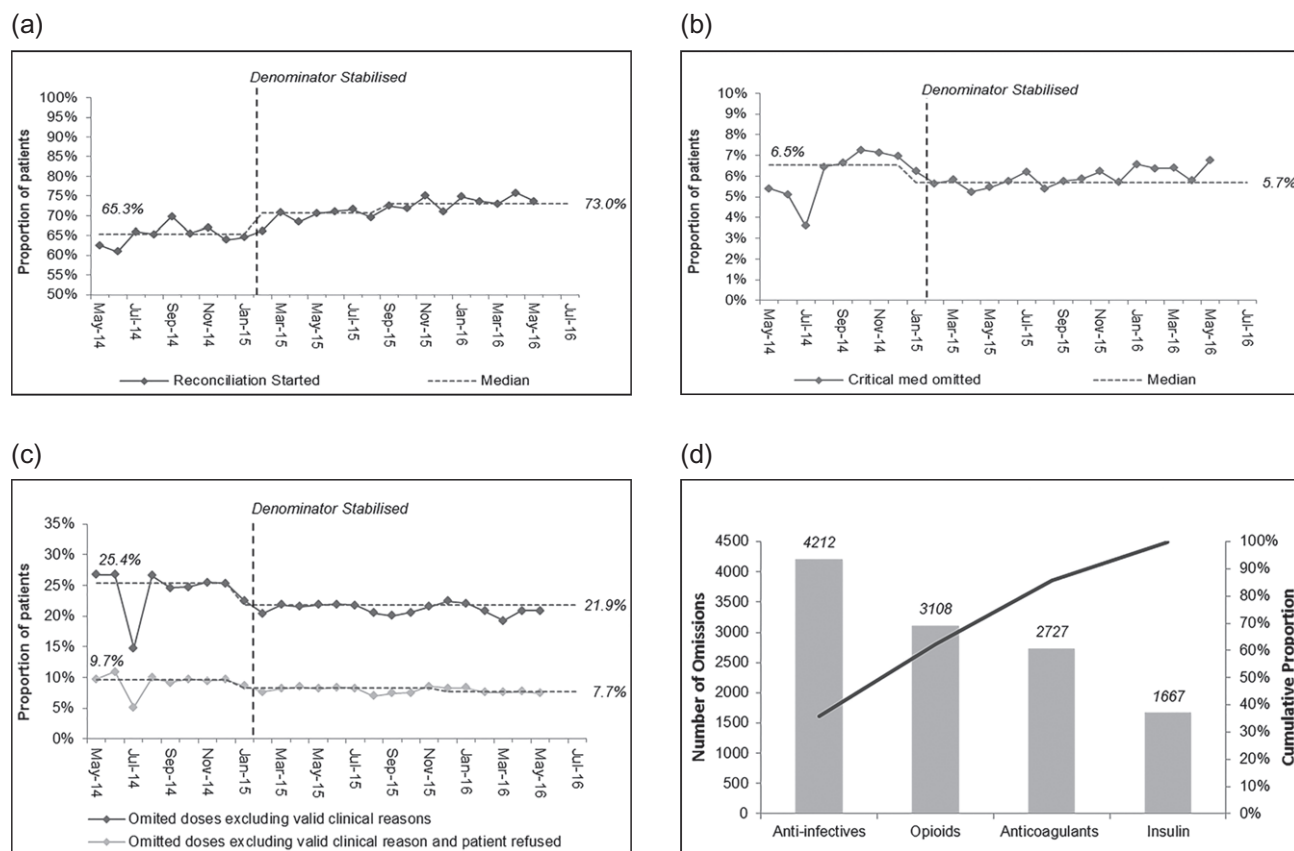
Figure 2 Medicines reconciliation and omissions data over 24 months. (**a**) Proportion of patients with a medicines reconciliation started in the last 24 hours of admission to setting. (**b**) Proportion of patients with omissions of critical medicine(s) in the last 24 hours (The last 24 hours from the point of data collection). (**c**) Proportion of patients who have had an omitted dose in the last 24 hours (Anti-infectives include: antibiotics, antifungals, antivirals and antimalarials). (**d**) Number of critical omissions by medication class (between October 2013 and April 2016). The line denotes the cumulative frequency of omissions.

hypothesis, that the number of organizations and patients surveyed would increase as the satisfaction and ease of use increased, was correct. This is also suggested by the fact that the majority of Greater Manchester organizations chose to continue using the MedsST, despite no longer receiving CQUIN payments after April 2014.

However, some organizations have stopped using the MedsST. Detailed analysis of such cases is warranted for further learning. Individual organizational data published online [29] demonstrates that, despite the constraints of using a tool that is relatively new, some organizations have improved [29]. This suggests that solutions to common problems may exist in the user community. Certain MedsST users, who are positive deviants, may have knowledge that can be generalized and, if the solutions have been generated within the MedsST user community, they may be more readily adopted in other organizations [36–38].

Suggestions for future work

Further research is required to explore how the MedsST is used in practice and to evaluate its utility. A mixed-method approach may be suitable for this. Investigation of variance in the use of the MedsST is warranted, for example, to explore the barriers preventing some organizations from using Step 3. Investigation of variance of the actual MedsST data is also warranted. Lessons can be learnt from organizations who have shown improvement in their MedsST

data. The positive deviance approach may be useful to explore how the MedsST can successfully be used for improvement.

Conclusion

The MedsST provides a refined methodology for measuring medication safety and its improvement over time. The PDSA approach has been particularly helpful in developing the tool. The increased engagement may be due to the refinement of the tool relying on regular feedback from frontline users; however, further research is required to ascertain this. The MedsST is inherently practical and easy to use, and has been used by over 100 healthcare organizations across the UK. To the best of our knowledge, it is the only tool measuring medication safety on a monthly basis. Data collection has led to demonstrable improvement in some organizations, but not all, indicating the need for further development and evaluation.

Acknowledgements

References

1. Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. *BMJ* 2001;**322**:517–9.

2. Department of Health. *An Organisation with a Memory: A Report from an Expert Working Group on Learning from Adverse Events in the NHS.* London: Department of Health, 2000.

3. de Vries EN, Ramrattan MA, Smorenburg SM *et al.* The incidence and nature of in-hospital adverse events: a systematic review. *Qual Saf Health Care* 2008;**17**:216–23.

4. Rozich JD, Haraden CR, Resar RK. Adverse drug event trigger tool: a practical methodology for measuring medication related harm. *Qual Saf Health Care* 2003;**12**:194–200.

5. Tache SV, Sonnichsen A, Ashcroft DM. Prevalence of adverse drug events in ambulatory care: a systematic review. *Ann Pharmacother* 2011;**45**:977–89.

6. Morimoto T, Gandhi TK, Seger AC *et al.* Adverse drug events and medication errors: detection and classification methods. *Qual Saf Health Care* 2004;**13**:306–14.

7. Lewis PJ, Dornan T, Taylor D *et al.* Prevalence, incidence and nature of prescribing errors in hospital inpatients: a systematic review. *Drug Saf* 2009;**32**:379–89.

8. Cousins DH, Gerrett D, Warner B. A review of medication incidents reported to the National Reporting and Learning System in England and Wales over six years (2005–10). *Br J Clin Pharmacol* 2012;**74**:597–604.

9. McLeod M Measuring medication errors. In: Tully MP, Dean Franklin B. *Safety in Medication Use*. London: CRC Press, 2015;61–72.

10. Grissinger M. Measuring up to medication safety in hospitals. *P T* 2009;**34**:10–50.

11. Bates DW, Spell N, Cullen DJ *et al.* The costs of adverse drug events in hospitalized patients. Adverse Drug Events Prevention Study Group. *JAMA* 1997;**277**:307–11.

12. Williams SD, Phipps DL, Ashcroft D. Examining the attitudes of hospital pharmacists to reporting medication safety incidents using the theory of planned behaviour. *Int J Qual Health Care* 2015;**27**:297–304.

13. Hutchinson A, Young T, Cooper K *et al.* Trends in healthcare incident reporting and relationship to safety and quality data in acute hospitals: results from the NRLS. *Qual Saf Health Care* 2009;**18**:5–10.

14. Ferner RE. The epidemiology of medication errors: the methodological difficulties. *Br J Clin Pharmacol* 2009;**67**:614–20.

15. Montesi G, Lechi A. Prevention of medication errors: detection and audit. *Br J Clin Pharmacol* 2009;**67**:651–5.

16. Haelo. What We Do: Haelo. Available from: http://www.haelo.org.uk/what-we-do/ (14 October 2016, date last accessed).

17. Buckley C, Cooney K, Sills E *et al.* Implementing the safety thermometer tool in one NHS trust. *Br J Nurs* 2014;**23**:268–72.

18. Power M, Fogarty M, Madsen J *et al.* Learning from the design and development of the NHS Safety Thermometer. *Int J Qual Health Care* 2014;**26**:287–97.

19. Department of Health. *The NHS Outcomes Framework 2012–2013.* London: Department of Health. 2011. https://www.gov.uk/government/publications/nhs-outcomes-framework-2012-to-2013 (15 July 2016, date last accessed).

20. Haelo. *The Medication Safety Thermometer.* Salford: Haelo Available at: https://www.safetythermometer.nhs.uk/index.php?option=com_wrapper&view=wrapper&Itemid=521 (15 December 2016, date last accessed).

21. Department of Health. *Using the Commissioning for Quality and Innovation (CQUIN) Payment Framework—Guidance on New National Goals for 2012-13.* London: Department of Health, 2012.

22. Leotsakos A, Zheng H, Croteau R *et al.* Standardization in patient safety: the WHO High 5s project. *Int J Qual Health Care* 2014;**26**:109–16.

23. Haelo. The NHS Safety Thermometers: Haelo. Available from: www.safetythermometer.nhs.uk (07 March 2016, date last accessed).

24. Parry G, Cline A, Goldmann D. Deciphering harm measurement. *JAMA* 2012;**307**:2155–6.

25. Naessens JM, O'Byrne TJ, Johnson MG *et al.* Measuring hospital adverse events: assessing inter-rater reliability and trigger performance of the Global Trigger Tool. *Int J Qual Health Care* 2010;**22**:266–74.

26. Singh R, McLean-Plunckett EA, Kee R *et al.* Experience with a trigger tool for identifying adverse drug events among older adults in ambulatory primary care. *Qual Saf Health Care* 2009;**18**:199–204.

27. Reed JE, Card AJ. The problem with Plan-Do-Study-Act cycles. *BMJ Qual Saf* 2015;**25**:147–52.

28. Berwick DM. Improvement, trust, and the healthcare workforce. *Qual Saf Health Care* 2003;**12**:i2–6.

29. Tully MP, Franklin BD Conclusion. In: Tully MP, Franklin BD. *Safety in Medication Use*. London: CRC Press, 2015;263–4.

30. Iddles E, Williamson A, Bradley A *et al.* Are we meeting current standards in medicines reconciliation? A study in a District General Hospital. *BMJ Qual Improv Rep* 2015;**4**:1.

31. Smith J. Improving and Maintaining Medicines Reconciliation on Admission at North Bristol NHS Trust National Institute for Health and Care Excellence, 2015. Available from: https://www.nice.org.uk/sharedlearning/improving-and-maintaining-medicines-reconciliation-on-admission-at-north-bristol-nhs-trust-nbt (05 July 2016, date last accessed).

32. Warne S, Endacott R, Ryan H *et al.* Non-therapeutic omission of medications in acutely ill patients. *Nurs Crit Care* 2010;**15**:112–7.

33. Baqir W, Jones K, Horsley W *et al.* Reducing unacceptable missed doses: pharmacy assistant-supported medicine administration. *Int J Pharm Pract* 2015;**23**:327–32.

34. Coleman JJ, McDowell SE, Ferner RE. Dose omissions in hospitalized patients in a UK hospital: an analysis of the relative contribution of adverse drug reactions. *Drug Saf* 2012;**35**:677–83.

35. Provost LP, Murray S. *The Health Care Data Guide: Learning from Data for Improvement*. San Francisco: John Wiley & Sons, 2011.

36. Lawton R, Taylor N, Clay-Williams R *et al.* Positive deviance: a different approach to achieving patient safety. *BMJ Qual Saf* 2014;**23**:880–3.

37. Marsh DR, Schroeder DG, Dearden KA *et al.* The power of positive deviance. *BMJ* 2004;**329**:1177–9.

38. Baxter R, Taylor N, Kellar I *et al.* Learning from positively deviant wards to improve patient safety: an observational study protocol. *BMJ Open* 2015;**5**:e009650.