

Balancing measures or a balanced accounting of improvement impact: a qualitative analysis of individual and focus group interviews with improvement experts in Scotland

Madalina Toma,¹ Tobias Dreischulte,² Nicola M Gray,¹ Diane Campbell,³ Bruce Guthrie⁴

¹Scottish Improvement Science Collaborating Centre (SISCC), University of Dundee School of Nursing and Health Science, Dundee, UK

²NHS Tayside, Dundee, UK

³Improvement Academy, NHS Tayside, Dundee, UK

⁴Quality, Safety and Informatics Research Group, Population Health Sciences Division, University of Dundee School of Medicine, Dundee, UK

Correspondence to

Dr Bruce Guthrie, Scottish Improvement Science Collaborating Centre (SISCC), University of Dundee School of Nursing and Midwifery, 11 Airlie Place, Dundee DD1 4HJ, UK; B.Guthrie@dundee.ac.uk

Received 1 February 2017

Revised 1 September 2017

Accepted 23 September 2017

Published Online First

21 October 2017

ABSTRACT

Background As quality improvement (QI) programmes have become progressively larger scale, the risks of implementation having unintended consequences are increasingly recognised. More routine use of balancing measures to monitor unintended consequences has been proposed to evaluate overall effectiveness, but in practice published improvement interventions hardly ever report identification or measurement of consequences other than intended goals of improvement.

Methods We conducted 15 semistructured interviews and two focus groups with 24 improvement experts to explore the current understanding of balancing measures in QI and inform a more balanced accounting of the overall impact of improvement interventions. Data were analysed iteratively using the framework approach.

Results Participants described the consequences of improvement in terms of desirability/undesirability and the extent to which they were expected/unexpected when planning improvement. Four types of consequences were defined: expected desirable consequences (*goals*); expected undesirable consequences (*trade-offs*); unexpected undesirable consequences (*unpleasant surprises*); and unexpected desirable consequences (*pleasant surprises*).

Unexpected consequences were considered important but rarely measured in existing programmes, and an improvement pause to take stock after implementation would allow these to be more actively identified and managed. A balanced accounting of all consequences of improvement interventions can facilitate staff engagement and reduce resistance to change, but has to be offset against the cost of additional data collection.

Conclusion Improvement measurement is usually focused on measuring intended *goals*, with minimal use of balancing measures which when used, typically monitor *trade-offs* expected before implementation. This paper proposes that improvers and leaders should seek a balanced accounting of all consequences of improvement across the life of an improvement programme, including deliberately pausing after implementation to identify and quantitatively or qualitatively evaluate any *pleasant* or *unpleasant surprises*.

BACKGROUND

Unintended consequences with negative or positive effects on care processes and outcomes can occur with any change in complex systems like healthcare organisations,¹⁻³ and so are an important potential problem in quality improvement (QI).⁴⁻⁶ More routine use of balancing measures to account for and manage unintended consequences of improvement interventions is recommended by a number of organisations.⁷⁻¹⁰ The Institute of Healthcare Improvement (IHI), for example, describes measurement in improvement programmes in terms of process and outcome measures focused on delivering predefined intended benefits, and balancing measures in terms of negative unintended consequences in other parts of the healthcare system (box 1).⁷⁻⁸ Reflecting this perspective, hospital readmission rates are often used as a balancing measure for interventions aiming to reduce the length of hospital stay, since it is plausible that shortening length of stay could mean discharging patients who are then unable to manage at home.¹¹⁻¹³

Despite calls for a more systematic accounting of all side effects of improvement interventions,¹⁴⁻¹⁵ a number of systematic reviews have shown that balancing measures appear rarely used or reported in practice. A review of the application of Plan Do Study Act (PDSA) methods found that only 6 (6.4%) of 94 included studies reported any ‘*disconfirming observations*’ about the intervention,¹⁶ and only 1 of 100 included studies in a systematic review of perioperative care improvement interventions reported



To cite: Toma M, Dreischulte T, Gray NM, et al. *BMJ Qual Saf* 2018;**27**:547–556.

Box 1 Institute of Healthcare Improvement (IHI) recommended types of measures^{7,8}

'Use a balanced set of measures for all improvement efforts: outcomes measures, process measures, and balancing measures.

1. Outcome Measures: How does the system impact the values of patients, their health and wellbeing? What are impacts on other stakeholders such as payers, employees, or the community?
2. Process Measures: Are the parts/steps in the system performing as planned? Are we on track in our efforts to improve the system?
3. Balancing Measures (looking at a system from different directions/dimensions): Are changes designed to improve one part of the system causing new problems in other parts of the system?'

Adapted from IHI (text is verbatim quote but examples are omitted and text is renumbered).

an 'unfavourable or unintended sign, symptom or event'.¹⁷ These findings are consistent with other reviews, including one of the applications of improvement methodologies in surgery which found that none of 34 included studies reported on unintended consequences,¹⁸ and another where only 1 of 121 studies of interventions to reduce patient falls and catheter-associated infections measured any unintended consequences.¹⁹ Several other studies in the latter review provided anecdotal evidence of 'unexpected occurrences',¹⁹ but robust evaluation of such claims is rare in improvement programmes more generally.²⁰ There is additionally little evidence that improvers routinely consider the potential for unexpected consequences postimplementation,²¹ and the amount of missing data about outcomes other than goals is often significant.^{22,23} The aim of this paper is to explore current understanding of balancing measures in healthcare improvement, including the range of consequences that could, or should be considered to inform a more balanced accounting of the overall impact of improvement interventions.

METHODS**Design and participants**

The research was carried out in two phases, with semistructured interviews used in the initial phase to formulate a draft conceptual framework for considering all consequences of improvement which was then explored using focus group interviews to refine and elaborate the framework, and to consider its wider applicability.

We used purposive sampling to include a broad spectrum of stakeholders with expertise in metrics and measure design in healthcare QI or relevant clinical and/or academic experience in improvement

implementation. Participants in both phases of the study included improvement advisors, clinical academics, providers of health and social care services, policymakers and patient representatives identified from relevant publication records and major conferences on QI, members of QI groups, online searches of open-access information and research teams' networks and contacts. Participants were largely based in Scotland, where comprehensive healthcare, which is free at the point of care, is provided to all residents by the taxpayer-funded National Health Service (NHS). Digital maturity of the system varies, with all primary care practices exclusively using electronic medical records (EMR) with widespread electronic data sharing (including, for example, primary care sharing of data for hospital use in an emergency care summary, electronic transmission of letters and discharge summaries, and automated laboratory results transmission), but hospitals being at various stages of EMR implementation. NHS Scotland has invested significantly in staff training in improvement and introduced a number of centrally led national safety and QI programmes,²⁴ largely (but not exclusively) based on the IHI Model for Improvement. Additional participants with particular expertise or known interest in measurement were purposively recruited from England and the USA. All participants were actively involved in service improvement across various settings including social care, mental health, public health, medicine for the elderly, maternity, neonatal and paediatric care.

Data collection and analysis

Phase 1: semistructured interviews to formulate the framework

Twelve face-to-face semistructured interviews and three telephone interviews each lasting for approximately 1 hour explored participants' understanding of balancing measures as part of a broader discussion about QI methods in health and social care. Individual interviews followed a topic guide based on the published literature and two pilot interviews. Data were analysed according to the principles of the framework approach²⁵ by developing codes and categories from the transcripts and grouping them into a preliminary coding matrix. The Diffusion of Innovation literature²⁶⁻³⁰ was used to reinterpret the initial matrix and generate a more structured framework reflecting participants' conceptualisation of balancing measures. The researcher who conducted the interviews (MT) coded all transcripts with a selection of transcripts and the emerging framework reviewed by a second experienced researcher (BG) to refine the coding.

Phase 2: focus group interviews to refine the framework

Two focus groups were conducted to explore the current understanding of balancing measures in QI and to elaborate the framework generated in phase 1. The draft framework was shared in a briefing paper prior to the focus group meeting and was used to inform initial

discussions within the groups. Focus groups were facilitated by two experienced moderators (MT and BG), lasted about 75 min each and took place on a single day. Interviews and focus groups were conducted in a non-directive manner, with participants encouraged to talk openly and with relative freedom to steer the discussion. The main researcher kept a journal with field notes reflecting on the research process, including prior assumptions that might have influenced the findings. Data were analysed using an iterative and stepwise process. The framework developed in phase 1 was used as a coding matrix in the analysis. Codes from focus group transcripts were grouped into subthemes, which were then allocated to one of the domains of the initial framework. One researcher (MT) coded all data and the wider team met regularly to reach consensus on the final framework structure, discuss additional categories and resolve any disagreement.

All interviews and focus group data were audio-recorded, transcribed verbatim and analysed using NVivo V.11.

FINDINGS

Participants' characteristics

Semistructured interviews with 15 participants and 2 focus groups with 24 participants (two of whom were also interviewed in phase 1) were completed. Participants had a wide range of roles in improvement and implementation science. Thirty-two participants came

from Scotland, four from England and one from the USA (table 1).

Phase 1: semistructured interviews

Identifying key themes and concepts

When asked about their overall understanding of balancing measures, participants initially emphasised negative consequences of improvement in other parts of the healthcare system, paralleling the IHI definition.

'My understanding is that a balancing measure is essentially something that you put in place because you recognise that often you can go in with the best of intent to improve an issue, you can deliver the improvement but you just end up creating more problems somewhere else.' (Improvement advisor)

Specific examples were again typically framed negatively, often as 'adverse' or 'knock-on' effects. Some of these were described as predictable from the outset, and measured routinely in the local improvement context.

'The mental health safety programme has balancing measures around recovery, about being very clear that one way of improving safety could lead to less positive risk-taking, which would be a very negative unintended consequence. We always use the Scottish recovery indicator, making sure that we promote recovery-oriented practices and we're not clamping down on folk.' (Mental healthcare provider)

Table 1 Characteristics of participants in both phases of the study

Participants' primary roles and responsibilities	Setting	Phase 1: semistructured interviews n=15*	Phase 2: focus group interviews n=24†
<i>Improvement advisors</i> with relevant clinical background and healthcare improvement expertise both locally and nationally, external to the local clinical and managerial teams	Primary care, maternity, neonatal and paediatrics, mental health, healthcare-associated infections, high-risk medicines	5	5
People with a university or similar <i>academic base and perspective</i> , relevant clinical background and healthcare improvement expertise both locally and nationally	Public health, palliative care, primary care, community engagement, health inequalities, capacity and capability building	3	9
<i>Providers of healthcare services</i> including clinicians in leadership positions in quality and safety who retain a significant role within their routine clinical practice, being involved in delivering healthcare improvement both locally and nationally	Primary care, mental health, medicine for the elderly, public health, maternity, neonatal and paediatrics	2	5
<i>Providers of social care services</i> in leadership positions in quality and safety who are involved in facilitating improvements both locally and nationally across the integrated health and social care services	Community health and social care partnerships	1	1
<i>Policy makers and commissioners</i> involved in monitoring performance and setting the general direction of quality improvement	Healthcare, social care, education and early years	3	2
<i>Patient representatives</i> advising health boards on the most efficient ways of accounting for the views and experiences of the people who use the local services	Cardiac care and rehabilitation, dementia care, maternity care	1	2

*All interviews were conducted face-to-face except two academics and one policymaker who were interviewed by telephone.

†Two participants in the interviews (one improvement advisor and one academic) also attended the focus groups.

Other negative consequences were described as only emerging as a potential problem after initial implementation, requiring improvers to be sensitive to the possibility of harm, and to be ready to ask themselves ‘right, what are we going to put in place to measure these adverse effects and see whether the improvement is actually causing any harm?’ (Academic and public health specialist), in order to inform further investigation or action.

‘Work to increase rates of early discharge and reduce length of stay led to patients being discharged into inappropriate conditions which in turn caused an increase in costs and readmission rates (...) That should be a wee bit of a red flag for you to think ‘why is everybody coming back? Are they coming back in because of surgical site infections or because you didn’t get their medicines reconciliation right on discharge?’ (...)’ (Improvement advisor)

Less commonly, participants described unanticipated positive or beneficial consequences. Although they were often uncertain whether these could be considered ‘balancing measures’ since they did not balance the benefits of improvement, they were highly valued by those who had experience of them.

‘A QI initiative aimed at improving writing and reading skills in secondary schools led to a reduction in absence rates as a result of better students’ engagement with different activities across the school (...) It was actually quite surprising and certainly a delightful outcome that we can now flip into a new piece of work to support children to become more engaged across their whole learning journey.’ (Provider of social care services)

However, in practice, the use of balancing measures was perceived to be rare in large-scale healthcare improvement programmes.

‘Most safety programmes haven’t paid much attention to balancing measures. From forty-nine pages of measures [in a safety improvement programme], there’s probably only two or three balancing measures like readmission rates, average length of stay or reintubation rates when reducing the time patients spend on a ventilator after surgery (...)’ (Policymaker health and social care)

Formulating the framework

In summary, when first asked about balancing measures, participants typically started from the position that measures should be implemented to assess undesirable unintended consequences of improvement work. However, their subsequent description of balancing measures also included unanticipated desirable consequences, and considerable discussion of the extent to which all consequences were predictable from the outset. Drawing on the Diffusion of Innovation literature,^{26–30} we developed an initial framework that describes the range of consequences that improvement could have, in terms of their desirability and the

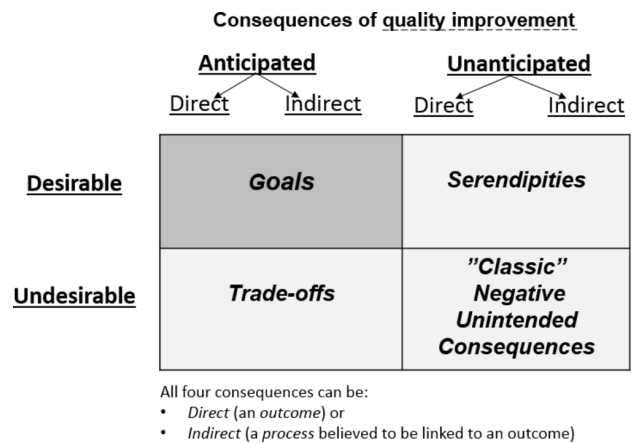


Figure 1 Draft framework of types of consequences of quality improvement projects (derived from phase 1 data and the literature, sent to phase 2 participants before the focus groups).

extent to which they were anticipated when planning improvement.

Four types of consequence were defined at this stage and described as *goals*, *trade-offs*, *classic negative unintended consequences* and *serendipities* (figure 1, sent to phase 2 participants before the focus groups).

Phase 2: focus group interviews

Mapping key themes and concepts

Similar to the individual interviews, focus group participants initially described balancing measures in terms of trade-offs, that is, negative unintended consequences of QI that were expected from the outset.

‘A lot of potential consequences are known at the start. ‘Oh, we need to actually count that, it will be an interesting balancing measure’. In a recent project focused on improving growth by early enteral feeding and maximise use of parenteral nutrition, the rates of necrotising enterocolitis and community-acquired bloodstream infections had reasonable potential for a balancing measure.’ (Provider of neonatology services)

However, as in the individual interviews, participants discussed several examples when undesirable consequences only became apparent after implementation, with examples from the same area of care targeted by improvement, as well as other parts of the wider system.

‘Inducing pregnant women at 40 weeks aimed to decrease the risk of stillbirth and newborn death but led to the use of extra interventions such as continuous fetal monitoring (...) which in turn increased costs and decreased overall patient satisfaction. Also woman who had a serious medical need for an induction could not get on the schedule because all of the hospital beds were occupied by women being electively induced.’ (Provider of maternal and infant healthcare)

Participants also mentioned desirable unintended consequences referring to ‘*serendipitous side effects or bonuses which are not planned as original programme*

outcomes' (academic and primary care provider), which they said were important to consider in order to obtain a balanced view of the overall impact of improvement interventions.

'The Book Bug sessions were established to strengthen attachment between parents and children by encouraging them to share and enjoy books together. One of the measures, which wasn't a balancing measure in the first instance but turned into one, was an increased interest from parents to improve their own literacy, bearing in mind that they had a young child that would need supported through school.' (Public health specialist)

However, even when unintended consequences were clearly identified, concerns were raised about the difficulty of creating or implementing a fully balanced set of measures, since data were not usually available from the outset unless routinely accessible from an existing source.

'I think we struggle with balancing measures. We always know we should think about them beforehand, but don't know how to deal with what comes up during the project (...) I think in safety we probably talk more about negative expected consequences, and the unexpected ones are the tip of the iceberg stuff (...) I don't think we become aware of them very often and we tend to then think 'oh it would have been nice to have data on that at the beginning'. (...) they almost feel like a missed opportunity.' (Academic capacity building)

Barriers and facilitators to using balancing measures

In terms of measure design, the majority of interviewees found the distinction between 'process', 'outcome' and 'balancing' measures in some of the improvement literature confusing, since balancing measures could relate to processes and outcomes depending on the context.

'We tend to be quite prescriptive about the family of measures and putting things into baskets of process and outcome and balancing measures is not always helpful. I don't think we pay enough attention to balancing measures and I'm not sure whether they're the right ones either (...) Readmission rates and average length of stay are balancing measures, but they could also be outcomes or processes that we might measure.' (Academic and palliative care provider)

Participants broadly perceived balancing measures to be important and relatively underused but reflected on the increasing burden of data collection in already resource-constrained systems.

'The time that we spend collecting or looking for data is time we don't spend delivering patient care, so there's a cost to this. Having balancing measures could be disproportionately expensive (...) just one of those things when measures are added on and on and nothing's changing. You're just collecting for the sake of collecting. You need to consider these measures very

carefully or it's a waste of peoples' time.' (Provider of geriatric healthcare)

However, there was a general agreement that engaging those involved in delivering care in the choice and design of measures from the outset would likely lead to better understanding of the rationale for measuring and could help minimise the burden of data collection.

'If the work is owned by the frontline staff, if it's their piece of improvement and if they've developed their own balancing measures then they're not going to think that measurement is too onerous in the same way as other would if they don't understand why they're measuring.' (Policymaker education and early years)

More importantly, the overall process of considering unintended consequences and implementing balancing measures was perceived to have value in its own right in terms of improving staff engagement with improvement and overcoming resistance to change.

You find a lot of latent resistance because people are genuinely worried about an unintended consequence and they don't engage in the work. You can introduce your checklist and it is fantastic, but it really annoys the staff because 'this is just going to take up a huge amount of time' (...) Using a balancing measure can convince your communities that improvement is needed and could be a goodwill builder if people know that you're monitoring and taking their concerns seriously. (Academic community engagement)

Refining the framework

Figure 2 shows a revised version of the framework that takes account of focus group findings, including the language used (eg, 'expected' rather than 'anticipated'). Desirability was described as a clear dichotomy, but expectations were perceived as more of a spectrum. While an initial measurement plan can define consequences expected from the outset (*goals* and *trade-offs*), participants thought that improvement programmes might need to plan for a 'pause' after implementation to account for unexpected consequences, both desirable and undesirable. The language of 'serendipities' and 'classic negative unintended consequences' was disliked, and renamed. The four types of consequences in the revised framework (figure 2) were therefore: *improvement goals*: the expected and desirable consequences of the improvement programme, defined by the initial measurement plan; *improvement trade-offs*: the expected but undesirable consequences of the improvement programme, and implicitly believed to be smaller in magnitude than the goals (and so an acceptable compromise); *pleasant surprises*: unexpected and desirable consequences emerging after implementation; *unpleasant surprises*: unexpected and undesirable consequences emerging after implementation.

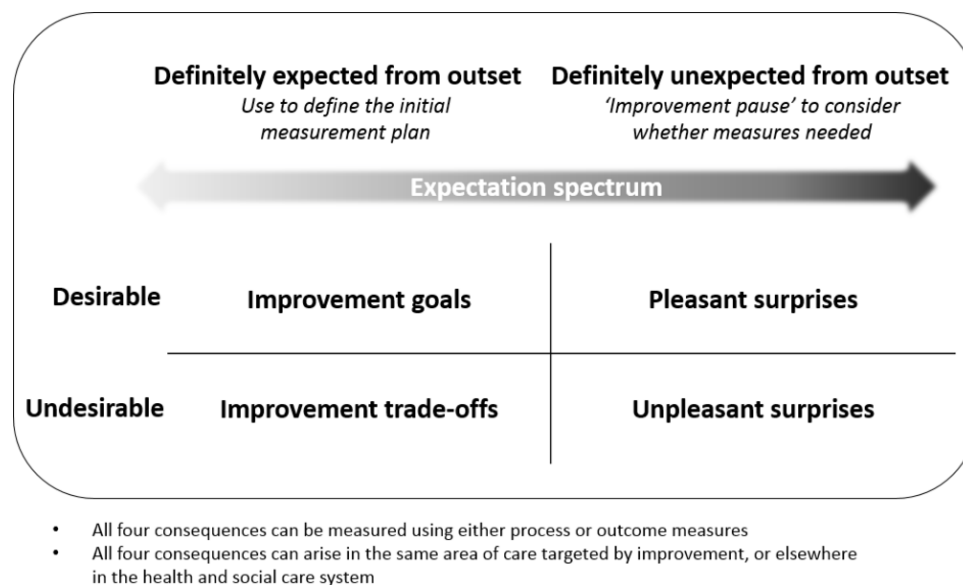


Figure 2 Refined framework of types of consequences of quality improvement projects (derived from phase 1 data and the literature, and refined after phase 2 focus groups).

All four consequences can be measured using either process or outcome measures and can arise in the same area of care targeted by improvement, or elsewhere in the health and social care system.

DISCUSSION

Summary of findings

Participants started by discussing balancing measures in terms of undesirable consequences which were expected before or early in implementation (*trade-offs*) and which could offset some of the intended benefits of improvements (*goals*). Although a range of examples were discussed, most participants agreed that such measures were relatively rarely used. Participants additionally emphasised that many consequences only became apparent after implementation, and these unexpected consequences could be either desirable or undesirable (*pleasant or unpleasant surprises*) and could accrue in the same part of the system as the improvement work, or other parts. There was frequent confusion as to what a balancing measure should measure, since the implication of many existing framings^{7 8} is that balancing measures are distinct in some way from process and outcome measures, rather than any type of consequence being measurable in terms of processes and outcomes. Involving front-line staff in identifying unintended consequences and balancing measure design was perceived to increase engagement with improvement and reduce resistance to change. Balancing measures were seen as a necessary and integral part of evaluating the impact of an improvement programme, as well as a pragmatic way of engaging sceptics constructively by understanding their legitimate concerns around implementation. However, the value of designing and implementing

balancing measures has to be offset against their cost in the context of overall measurement burden.

Strengths and limitations of the study

A strength of the study is that it drew on both empirical data from a purposively wide range of stakeholders and existing literature on unintended consequences. A limitation is that the sample was largely recruited from Scotland which may limit generalisability. However, NHS Scotland has a history of centrally led, and broadly successful efforts to introduce system-wide improvement interventions, most commonly based on the IHI Model for Improvement including training and implementation of national safety programmes in acute hospitals, mental healthcare and primary care.²⁴ Participants therefore had experience of a number of improvement programmes to draw on, although limited implementation of electronic medical records in hospitals means that perceptions of the burden of data collection will at least partly reflect that data used in national improvement programmes currently almost entirely consist of bespoke data collected by clinical staff. Findings were consistent across the diverse range of stakeholders (including those outside in Scotland), and we believe that the measurement issues faced by improvement programmes in Scotland are likely to be relevant in other countries and systems worldwide.

Comparison with existing literature

The existing improvement literature on measurement design emphasises the importance of developing a balanced set of measures during the planning of an improvement programme,^{7-10 31-33} often distinguishing between process and outcome measures for goals, and balancing measures for expected undesirable

Table 2 Published examples of trade-offs and pleasant and unpleasant surprises in the improvement literature

Study	Improvement goals	Other consequences	Examples of balancing measures prior, during and postimplementation
Kavanagh <i>et al</i> ³⁵	To improve the timeliness of management of vaso-occlusive pain events in children with sickle cell disease in paediatric emergency departments	Expected undesirable consequences (<i>trade-offs</i>)	<i>Mean time from triage to the second intravenous opioid dose</i> was introduced as a balancing measure because of concern that the use of intranasal fentanyl as the first-line intervention might delay subsequent intravenous dosing. Other <i>trade-offs</i> measured included <i>readmission rates</i> within 24 hours of discharge, episodes of <i>respiratory depression</i> and inpatient <i>length of stay</i> .
Dewan <i>et al</i> ³⁶	To decrease unnecessary routine complete blood count testing in a low-risk cohort of postoperative patients in the paediatric intensive care units	Expected undesirable consequences (<i>trade-offs</i>)	Balancing measures were implemented for <i>haemoglobin level below 8g/dL</i> in patients for whom complete blood counts were actually sent and <i>blood transfusions up to 7 days postoperatively</i> for any patients in the cohort.
Duvoisin <i>et al</i> ³⁷	To reduce the number of unnecessary diagnostic tests such as complete blood count and C-reactive protein in infants with risk factors for early-onset neonatal sepsis	Unexpected desirable consequences (<i>pleasant surprises</i>)	There was preintervention concern that reduction in the use of diagnostic tests would <i>delay the initiation of antibiotic treatment</i> , but unexpectedly the intervention resulted in <i>earlier treatment of infection</i> on average.
Bell <i>et al</i> ³⁸	To reduce the preoperative use of antimicrobials associated with <i>Clostridium difficile</i> infection	Unexpected undesirable consequences (<i>unpleasant surprises</i>)	The new surgical prophylaxis regimen of four doses of flucloxacillin 1 g plus single-dose gentamicin 4 mg/kg unexpectedly led to <i>increased rates of postoperative acute kidney injury</i> in orthopaedic patients, large enough to lead to the termination of the intervention through a change in the national antibiotic policy recommendation for orthopaedic surgical prophylaxis.
Strom <i>et al</i> ³⁹	To evaluate the effectiveness of a customised nearly hard stop alert in reducing concomitant orders for warfarin and trimethoprim-sulfamethoxazole compared with the standard practice of a pharmacist intervention programme	Unexpected undesirable consequences (<i>unpleasant surprises</i>)	Unexpected <i>delays in indicated anticoagulant and/or antimicrobial treatment initiation</i> were deemed sufficiently serious to warrant discontinuation of the improvement intervention.

consequences (*trade-offs*) which are easily predictable from the outset (box 1). However, participants in this study found this framing too narrow because they were concerned about unexpected undesirable consequences (*unpleasant surprises*) and valued unexpected desirable consequences (*pleasant surprises*), neither of which could be defined prior to intervention implementation.

Although there are some studies of *trade-offs*,^{34–36} and *pleasant*³⁷ and *unpleasant*^{38, 39} *surprises* (table 2), published improvement interventions rarely report data relating to unintended consequences.^{15–19, 40} This may partly reflect publication bias, since authors are known to emphasise positive results and ‘*tuck away*’⁴¹ negative contextual features and failures.²³ However, it also likely reflects more general lack of consideration or measurement of unintended consequences, consistent with an observed preoccupation with measuring prespecified local processes and outcomes (*goals*).^{42–44} The implementation of PDSA cycles in healthcare, for example, has been criticised for often involving an oversimplified ‘*Do, Do, Do*’ approach¹⁵ focused on little and often measurement and delivery of *goals* at the expense of thinking ahead and looking to the future (for *trade-offs*) and reflecting on potential hazards during implementation (for *surprises*).^{45, 46}

Implications for improvement programme design

Balancing measures are an integral and core element of commonly used improvement models like the IHI Model for Improvement,^{7, 8} but they are sometimes poorly specified and do not appear to be commonly implemented in practice.^{15–19, 40} Based on the literature and the findings of this study, we believe that rather than focusing on balancing measures to implement at the start of improvement, improvers and leaders at all levels of management should consider how best to achieve a balanced accounting of the overall impact of improvement across the life of a programme. This requires consideration of all four types of consequence, any of which can be measured in terms of process and outcome (figure 2). Such a balanced accounting of impact can be achieved by articulating clear assumptions and formulating explicit predictions for both *goals* and *trade-offs* before implementation,^{14, 40, 47} and having a planned improvement pause after implementation to deliberately step back from *goal* delivery to take stock and reflect on potential *surprises*.^{46, 48} In an ideal world, improvers would consult the available evidence base and seek external input from key stakeholders in order to identify potential *trade-offs*, speculate on and investigate potential *surprises*, and if necessary, to design relevant process and outcome measures

to account for them. However, improvement takes place in resource-constrained environments, which will confine what is possible, including, for example, the feasibility of measurement in other areas of a complex system. Focusing on a balanced accounting rather than balancing measures also emphasises that qualitative methods have much to offer both for the identification of *trade-offs* before implementation, and for understanding *surprises* after implementation where retrospective measurement may be difficult.^{49 50}

Implications for reporting QI projects

Few improvement reports mention unintended consequences, despite the Standards for Quality Improvement Reporting Excellence (SQUIRE) guidance¹⁴ including a requirement that reporting should include '*unintended consequences, such as unexpected benefits, problems, failures or costs associated with the intervention*' (standard 13e). Of note though is that the SQUIRE explanation and elaboration for this standard⁵¹ focuses more on exploring variation in implementation effectiveness and does not provide any examples of significant elaboration of unintended consequences. As the volume of publications in QI is growing, modification of SQUIRE to clarify that improvement reports *should* report any measured or qualitatively assessed unintended consequences, or report that these were not assessed, would be helpful to contextualise any evidence presented about the achievement of improvement goals.

CONCLUSION

This study is largely based on analysis of data from interviews carried out in Scotland which has an integrated single-payer healthcare system and relatively well-developed QI infrastructure.⁵² However, improvement interventions in complex systems will often result in unintended consequences irrespective of context, so we believe that the conclusions apply more widely, although the ability of improvers to evaluate or measure unintended consequences will vary, being lower in more fragmented healthcare systems. Overall, the evidence is that improvement programme measurement is usually focused on evaluating intended *goals*, with minimal use of balancing measures which are typically monitoring *trade-offs* expected before implementation. We conclude that a more balanced accounting of the effects of improvement should consider *goals* and predictable *trade-offs* early in the design of an improvement programme, and also pause to take stock of *pleasant and unpleasant surprises* after a period of implementation.

Acknowledgements This work was undertaken by and on behalf of the Scottish Improvement Science Collaborating Centre (SISCC). We are grateful to our improvement experts for their time and willingness to cocreate the 'balanced accounting' framework. We also acknowledge Julie Anderson and Anita Lee who provided invaluable feedback on the revised versions of this paper.

Contributors MT and BG were responsible for planning the study and led the data collection and analysis. TD, NMG and DC contributed to data analysis. MT drafted and led the writing of the manuscript. BG, TD, NMG and DC participated in critically appraising and revising the intellectual content of the manuscript. All authors read and approved the final manuscript.

Funding This paper presents core work from the Scottish Improvement Science Collaborating Centre, funded by the Scottish Funding Council, Health Foundation, Chief Scientist Office, and NHS Education Scotland with in-kind contributions from participating partner universities and health boards. The funding bodies had no role in the design of the study and collection, analysis and interpretation of data, and in writing the manuscript.

Competing interests None declared.

Patient consent Detail has been removed from this case description/these case descriptions to ensure anonymity. The editors and reviewers have seen the detailed information available and are satisfied that the information backs up the case the authors are making.

Ethics approval The University of Dundee Research Ethics Committee granted ethical approval for this study (UREC 15069).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The data used and/or analysed during the current study are available from the corresponding author on reasonable request.

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

- 1 Merton RK. The unanticipated consequences of purposive social action. *Am Sociol Rev* 1936;1:894–904.
- 2 Wachter RM. Expected and unanticipated consequences of the quality and information technology revolutions. *JAMA* 2006;295:2780–3.
- 3 Han YY, Carcillo JA, Venkataraman ST, *et al*. Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system. *Pediatrics* 2005;116:1506–12.
- 4 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.
- 5 Bardach NS, Cabana MD. The unintended consequences of quality improvement. *Curr Opin Pediatr* 2009;21:777–82.
- 6 Dixon-Woods M, McNicol S, Martin G. Ten challenges in improving quality in healthcare: lessons from the Health Foundation's programme evaluations and relevant literature. *BMJ Qual Saf* 2012;21:876–84.
- 7 Institute of Healthcare Improvement. Science of improvement: establishing measures. <http://www.ihl.org/resources/pages/howtoimprove/scienceofimprovementestablishingmeasures.aspx> (accessed 21 Mar 2016).
- 8 Langley GJ, Nolan KM, Nolan TW, *et al*. *The Improvement Guide: A Practical Approach to Enhancing Organizational Performance*. San Francisco, CA: Jossey-Bass, 1996.
- 9 Clarke J, Davidge M, James L. How-to guide for measurement for improvement. *Patient Safety First* 2009 <https://eoleadership.hee.nhs.uk/sites/default/files/Patient%20Safety%20First%20How%20To%20Guide%20measurement%20for%20improvement.pdf> (Accessed 03 Feb 2016).

- 10 NHS Quality Improvement Scotland. *Introduction to measurement for improvement*. <http://www.qihub.scot.nhs.uk/knowledge-centre/qualityimprovementtopics/measurement-for-improvement.aspx> (Accessed 21 Jan 2016).
- 11 Asti L, Magers JS, Keels E, *et al*. A quality improvement project to reduce length of stay for neonatal abstinence syndrome. *Pediatrics* 2015;135:e1494–e1500.
- 12 Mittal V, Darnell C, Walsh B, *et al*. Inpatient bronchiolitis guideline implementation and resource utilization. *Pediatrics* 2014;133:e730–e737.
- 13 Turkington P, Power M, Hunt C, *et al*. There is another way: empowering frontline staff caring for acutely unwell adults. *Int J Qual Health Care* 2014;26:71–8.
- 14 Ogrinc G, Davies L, Goodman D, *et al*. SQUIRE 2.0 (Standards for Quality Improvement Reporting Excellence): revised publication guidelines from a detailed consensus process. *BMJ Qual Saf* 2016;25:986–92.
- 15 Reed JE, Card AJ. The problem with Plan-Do-Study-Act cycles. *BMJ Qual Saf* 2016;25:147–52.
- 16 Curnock E, Ferguson J, McKay J, *et al*. Healthcare improvement and rapid PDSA cycles of change: a realist synthesis of the literature. *NHS Education for Scotland* 2012 http://www.nes.scot.nhs.uk/media/1389875/pdsa_realist_synthesis.pdf (accessed 03 June 2017).
- 17 Jones EL, Lees N, Martin G, *et al*. How well is quality improvement described in the perioperative care literature? A systematic review. *Jt Comm J Qual Patient Saf* 2016;42:196–AP10.
- 18 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.
- 19 Manojlovich M, Lee S, Lauseng D. A systematic review of the unintended consequences of clinical interventions to reduce adverse outcomes. *J Patient Saf* 2016;12:173–9.
- 20 McDonald KM, Schultz EM, Chang C. Evaluating the state of quality-improvement science through evidence synthesis: insights from the closing the quality gap series. *Perm J* 2013;17:52–61.
- 21 Pronovost P, Wachter R. Proposed standards for quality improvement research and publication: one step forward and two steps back. *Qual Saf Health Care* 2006;15:152–3.
- 22 Brainard J, Hunter PR. Do complexity-informed health interventions work? A scoping review. *Implement Sci* 2016;11:127.
- 23 Jones E. SIQINS: Strengthening Reporting of Quality Improvement Interventions and Methods in Surgery. <https://ira.le.ac.uk/bitstream/2381/39747/1/2017-JONES-EL-PHD.pdf> (accessed 29 May 2017).
- 24 The Scottish Patient Safety Programme (SPSP). *Supporting Health and Social Care*. Scotland: NHS, 2017. <http://www.scot.nhs.uk/scotspatientsafetyprogramme.scot.nhs.uk/>. (accessed 11 May 2017).
- 25 Ritchie J, Lewis J, McNaughton Nicholls C, *et al*; Qualitative Research Practice. *A guide for social science students and researchers*. 2nd ed. London: Sage Publications, 2003.
- 26 Ash JS, Sittig DF, Dykstra RH, *et al*. Categorizing the unintended sociotechnical consequences of computerized provider order entry. *Int J Med Inform* 2007;76 (Suppl 1):S21–S27.
- 27 Ash JS, Sittig DF, Poon EG, *et al*. The extent and importance of unintended consequences related to computerized provider order entry. *J Am Med Inform Assoc* 2007;14:415–23.
- 28 Bloomrosen M, Starren J, Lorenzi NM, *et al*. Anticipating and addressing the unintended consequences of health IT and policy: a report from the AMIA 2009 Health Policy Meeting. *J Am Med Inform Assoc* 2011;18:82–90.
- 29 Rogers EM. *Diffusion of Innovations*. 4th ed. New York: Free Press, 1995.
- 30 Campbell EM, Sittig DF, Ash JS, *et al*. Types of unintended consequences related to computerized provider order entry. *J Am Med Inform Assoc* 2006;13:547–56.
- 31 Scottish Government. Quality Improvement and Data Measurement—What Non-Executive Directors need to know. 2016 <http://www.gov.scot/Resource/0049/00492311.pdf> (accessed 01 May 2016).
- 32 HRSA. U.S. Department of health and human services health resources and services administration: *performance management and measurement*. 2011 <http://www.hrsa.gov/quality/toolbox/508pdfs/performanceandmeasurement.pdf> (accessed 09 Mar 2016).
- 33 Health Quality Ontario. Measurement for quality improvement. 2013 <http://www.hqontario.ca/Portals/0/documents/qi/qi-measurement-primer-en.pdf> (accessed 13 Apr 2016).
- 34 Shojania KG, Ranji SR, McDonald KM, *et al*. Effects of quality improvement strategies for type 2 diabetes on glycemic control: a meta-regression analysis. *JAMA* 2006;296:427–40.
- 35 Kavanagh PL, Sprinz PG, Wolfgang TL, *et al*. Improving the management of vaso-occlusive episodes in the pediatric emergency department. *Pediatrics* 2015;136:e1016–e1025.
- 36 Dewan M, Galvez J, Polsky T, *et al*. Reducing unnecessary postoperative complete blood count testing in the pediatric intensive care unit. *Perm J* 2017;21:16–51.
- 37 Duvoisin G, Fischer C, Maucort-Boulch D, *et al*. Reduction in the use of diagnostic tests in infants with risk factors for early-onset neonatal sepsis does not delay antibiotic treatment. *Swiss Med Wkly* 2014;144:w13981.
- 38 Bell S, Davey P, Nathwani D, *et al*. Risk of AKI with gentamicin as surgical prophylaxis. *J Am Soc Nephrol* 2014;25:2625–32.
- 39 Strom BL, Schinnar R, Abera F, *et al*. Unintended effects of a computerized physician order entry nearly hard-stop alert to prevent a drug interaction: a randomized controlled trial. *Arch Intern Med* 2010;170:1578–83.
- 40 Taylor MJ, McNicholas C, Nicolay C, *et al*. Systematic review of the application of the plan-do-study-act method to improve quality in healthcare. *BMJ Qual Saf* 2014;23:290–8.
- 41 Matosin N, Frank E, Engel M, *et al*. Negativity towards negative results: a discussion of the disconnect between scientific worth and scientific culture. *Dis Model Mech* 2014;7:171–3.
- 42 Raleigh VS, Foot C. Getting the measure of quality. *Opportunities and challenges* 2010 <https://www.kingsfund.org.uk/sites/files/kf/Getting-the-measure-of-quality-Veena-Raleigh-Catherine-Foot-The-Kings-Fund-January-2010.pdf> (accessed 19 Jan 2016).
- 43 Dixon-Woods M, Amalberti R, Goodman S, *et al*. Problems and promises of innovation: why healthcare needs to rethink its love/hate relationship with the new. *BMJ Qual Saf* 2011;20(Suppl 1):i47–i51.
- 44 Dixon-Woods M, Martin G. Does quality improvement improve quality? *Future Hosp J* 2016;3:191–4.
- 45 Vincent C, Burnett S, Carthey J. The measurement and monitoring of safety. Drawing together academic evidence and practical experience to produce a framework for safety

- measurement and monitoring. *Health Foundation Spotlight report 2013* http://www.health.org.uk/sites/health/files/TheMeasurementAndMonitoringOfSafety_fullversion.pdf (accessed 29 Jul 2016).
- 46 Vincent C, Burnett S, Carthey J. Safety measurement and monitoring in healthcare: a framework to guide clinical teams and healthcare organisations in maintaining safety. *BMJ Qual Saf* 2014;23:670–7.
 - 47 McQuillan RF, Silver SA, Harel Z, *et al.* How to measure and interpret quality improvement data. *Clin J Am Soc Nephrol* 2016;11:908–14.
 - 48 Lee SL. The extended surgical time-out: does it improve quality and prevent wrong-site surgery? *Perm J* 2010;14:19–23.
 - 49 Powell AE, Rushmer RK, Davies HTO. A systematic narrative review of quality improvement models in health care. *Healthcare Improvement Scotland 2009* http://www.healthcareimprovementscotland.org/previous_resources/hta_report/a_systematic_narrative_review.aspx (accessed 20 Mar 2017).
 - 50 Portela MC, Pronovost PJ, Woodcock T, *et al.* How to study improvement interventions: a brief overview of possible study types. *BMJ Qual Saf* 2015;24:325–36.
 - 51 Goodman D, Ogrinc G, Davies L, *et al.* Explanation and elaboration of the SQUIRE (Standards for Quality Improvement Reporting Excellence) Guidelines, V2.0: examples of SQUIRE elements in the healthcare improvement literature. *BMJ Qual Saf* 2016;25:e7.
 - 52 Dayan M, Edwards N. Learning from Scotland's NHS. *Nuffield Trust Research report 2017* <https://www.nuffieldtrust.org.uk/research/learning-from-scotland-s-nhs> (accessed 27 Aug 2017).