

University of Dundee

Balancing measures or a balanced accounting of improvement impact

Toma, Madalina; Dreischulte, Tobias; Gray, Nicola M; Campbell, Diane; Guthrie, Bruce

Published in:
BMJ Quality & Safety

DOI:
[10.1136/bmjqs-2017-006554](https://doi.org/10.1136/bmjqs-2017-006554)

Publication date:
2018

Document Version
Peer reviewed version

[Link to publication in Discovery Research Portal](#)

Citation for published version (APA):

Toma, M., Dreischulte, T., Gray, N. M., Campbell, D., & Guthrie, B. (2018). Balancing measures or a balanced accounting of improvement impact: a qualitative analysis of individual and focus group interviews with improvement experts in Scotland. *BMJ Quality & Safety*, 27(7), 547-556. <https://doi.org/10.1136/bmjqs-2017-006554>

General rights

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

1 **Balancing measures or a balanced accounting of improvement impact. A qualitative**
2 **analysis of individual and focus group interviews with improvement experts in**
3 **Scotland**

4 **ABSTRACT**

5

6 **Background**

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

12

13 **Methods**

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

18

19 **Results**

20 Participants described the consequences of improvement in terms of desirability/undesirability and
21 the extent to which they were expected/unexpected when planning improvement. Four types of
22 consequences were defined: expected desirable consequences (*goals*); expected undesirable
23 consequences (*trade-offs*); unexpected undesirable consequences (*unpleasant surprises*) and
24 unexpected desirable consequences (*pleasant surprises*). Unexpected consequences were
25 considered important but rarely measured in existing programmes, and an improvement pause to
26 take stock after implementation would allow these to be more actively identified and managed. A
27 balanced accounting of all consequences of improvement interventions can facilitate staff
28 engagement and reduce resistance to change, but has to be offset against the cost of additional data
29 collection.

30

31 **Conclusion**

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

38

39 **Keywords**

40 Quality improvement; Measurement of quality; Balancing measures; Unintended consequences;
41 Expected, unexpected, desirable and undesirable consequences

42 **Word count for the abstract: 275**

43 **Word count for the text of the manuscript: 4000**

44 BACKGROUND

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

56

57 Despite calls for a more systematic accounting of all side effects of improvement interventions,^{14,15} a
58 number of systematic reviews have shown that balancing measures appear rarely used or reported
59 in practice. A review of the application of Plan Do Study Act (PDSA) methods found that only six
60 (6.4%) of 94 included studies reported any “*disconfirming observations*” about the intervention,¹⁶
61 and only one of 100 included studies in a systematic review of perioperative care improvement
62 interventions reported an “*unfavourable or unintended sign, symptom or event*”.¹⁷ These findings
63 are consistent with other reviews, including one of the application of improvement methodologies in
64 surgery which found that none of 34 included studies reported on unintended consequences,¹⁸ and
65 another where only one of 121 studies of interventions to reduce patient falls and catheter-
66 associated infections measured any unintended consequences.¹⁹ Several other studies in the latter
67 review provided anecdotal evidence of “*unexpected occurrences*”,¹⁹ but robust evaluation of such
68 claims is rare in improvement programmes more generally.²⁰ There is additionally little evidence that
69 improvers routinely consider the potential for unexpected consequences post-implementation,²¹
70 and the amount of missing data about outcomes other than goals is often significant.^{22,23} The aim of
71 this paper is to explore current understanding of balancing measures in healthcare improvement,
72 including the range of consequences that could, or should be considered to inform a more balanced
73 accounting of the overall impact of improvement interventions.

74 **METHODS**

75 **Design and participants**

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

80 We used purposive sampling to include a broad spectrum of stakeholders with expertise in metrics
81 and measure design in healthcare QI or relevant clinical and/or academic experience in
82 improvement implementation. Participants in both phases of the study included improvement
83 advisors, clinical academics, providers of health and social care services, policy-makers and patient
84 representatives identified from relevant publication records and major conferences on QI, members
85 of QI groups, on-line searches of open-access information and research teams' networks and
86 contacts. Participants were largely based in Scotland, where comprehensive healthcare, which is free
87 at the point of care, is provided to all residents by the taxpayer-funded National Health Service
88 (NHS). Digital maturity of the system varies, with all primary care practices exclusively using
89 electronic medical records (EMR) with widespread electronic data sharing (including for example,
90 primary care sharing of data for hospital use in an emergency care summary, electronic transmission
91 of letters and discharge summaries, and automated laboratory results transmission), but hospitals
92 being at various stages of EMR implementation. NHS Scotland has invested significantly in staff
93 training in improvement and introduced a number of centrally led national safety and quality
94 improvement programmes²⁴, largely (but not exclusively) based on the IHI Model for Improvement.
95 Additional participants with particular expertise or known interest in measurement were purposively
96 recruited from England and the United States. All participants were actively involved in service
97 improvement across various settings including social care, mental health, public health, medicine for
98 the elderly, maternity, neonatal and paediatric care.

99 **Data collection and analysis**

100 ***Phase 1 -Semi-structured interviews to formulate the framework***

101 Twelve face to face semi-structured interviews and three telephone interviews each lasting for
102 approximately one hour explored participants' understanding of balancing measures as part of a
103 broader discussion about QI methods in health and social care. Individual interviews followed a topic
104 guide based on the published literature and two pilot interviews. Data were analysed according to
105 the principles of the Framework approach²⁵ by developing codes and categories from the transcripts

106 and grouping them into a preliminary coding matrix. The Diffusion of Innovation literature²⁶⁻³⁰ was
107 used to reinterpret the initial matrix and generate a more structured framework reflecting
108 participants' conceptualisation of balancing measures. The researcher who conducted the interviews
109 (MT) coded all transcripts with a selection of transcripts and the emerging framework reviewed by a
110 second experienced researcher (BG) to refine the coding.

111 ***Phase 2 – Focus groups interviews to refine the framework***

112 Two focus groups were conducted to explore the current understanding of balancing measures in QI
113 and to elaborate the framework generated in Phase 1. The draft framework was shared in a briefing
114 paper prior to the focus group meeting and was used to inform initial discussions within the groups.
115 Focus groups were facilitated by two experienced moderators (MT and BG), lasted about 75 minutes
116 each, and took place on a single day. Interviews and focus groups were conducted in a non-directive
117 manner, with participants encouraged to talk openly and with relative freedom to steer the
118 discussion. The main researcher kept a journal with field notes reflecting on the research process,
119 including prior assumptions that might have influenced the findings. Data were analysed using an
120 iterative and step-wise process. The framework developed in phase one was used as a coding matrix
121 in the analysis. Codes from focus groups transcripts were grouped into sub-themes, which were then
122 allocated to one of the domains of the initial framework. One researcher (MT) coded all data and the
123 wider team met regularly to reach consensus on the final framework structure, discuss additional
124 categories, and resolve any disagreement.

125 All interviews and focus group data were audio-recorded, transcribed verbatim and analysed using
126 NVivo11.

127 **FINDINGS**

128 **Participants' characteristics**

129 Semi-structured interviews with 15 participants and two focus groups with 24 participants (two of
130 whom were also interviewed in phase 1) were completed. Participants had a wide range of roles in
131 improvement and implementation science. 32 participants came from Scotland, four from England
132 and one from the United States. (Table 1).

133 **Phase 1-Semi-structured interviews**

134 ***Identifying key themes and concepts***

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

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

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

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

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

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

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

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

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

172 *“Most safety programmes haven't paid much attention to balancing measures. From forty-*
173 *nine pages of measures [in a safety improvement programme], there's probably only two or*
174 *three balancing measures like readmission rates, average length of stay or reintubation rates*
175 *when reducing the time patients spend on a ventilator after surgery (...)”* (Policy maker
176 health and social care)

177 **Formulating the framework**

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

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

189 **Phase 2-Focus group interviews**

190 **Mapping key themes and concepts**

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

194 *“A lot of potential consequences are known at the start. ‘Oh, we need to actually count that,*
195 *it will be an interesting balancing measure’. In a recent project focused on improving growth*
196 *by early enteral feeding and maximise use of parenteral nutrition, the rates of necrotising*

197 *enterocolitis and community-acquired bloodstream infections had reasonable potential for a*
198 *balancing measure.” (Provider of neonatology services)*

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

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

208 Participants also mentioned desirable unintended consequences referring to *“serendipitous side*
209 *effects or bonuses which are not planned as original programme outcomes”* (academic and primary
210 care provider), which they said were important to consider in order to obtain a balanced view of the
211 overall impact of improvement interventions.

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

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

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

226 **Barriers and facilitators to using balancing measures**

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

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

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

238 *“The time that we spend collecting or looking for data is time we don’t spend delivering*
239 *patient care, so there’s a cost to this. Having balancing measures could be disproportionately*
240 *expensive (...) just one of those things when measures are added on and on and nothing’s*
241 *changing. You’re just collecting for the sake of collecting. You need to consider these*
242 *measures very carefully or it’s a waste of peoples’ time.” (Provider of geriatric healthcare)*

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

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

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

253 *“You find a lot of latent resistance because people are genuinely worried about an*
254 *unintended consequence and they don’t engage in the work. You can introduce your checklist*
255 *and it is fantastic, but it really annoys the staff because “this is just going to take up a huge*
256 *amount of time’ (...) Using a balancing measure can convince your communities that*

257 *improvement is needed and could be a goodwill builder if people know that you're*
258 *monitoring and taking their concerns seriously.” (Academic community engagement)*

259 **Refining the framework**

260 Figure 2 shows a revised version of the framework that takes account of focus groups findings,
261 including the language used (eg ‘expected’ rather than ‘anticipated’). Desirability was described as a
262 clear dichotomy, but expectations were perceived as more of a spectrum. While an initial
263 measurement plan can define consequences expected from the outset (*goals* and *trade-offs*),
264 participants thought that improvement programmes might need to plan for a ‘pause’ after
265 implementation to account for unexpected consequences, both desirable and undesirable. The
266 language of ‘serendipities’ and ‘classic negative unintended consequences’ was disliked, and
267 renamed. The four type of consequences in the revised framework (Figure 2) were therefore:

268 **Improvement goals:** the expected and desirable consequences of the improvement programme,
269 defined by the initial measurement plan; **Improvement trade-offs:** the expected but undesirable
270 consequences of the improvement programme, and implicitly believed to be smaller in magnitude
271 than the goals (and so an acceptable compromise); **Pleasant surprises:** unexpected and desirable
272 consequences emerging after implementation; **Unpleasant surprises:** unexpected and undesirable
273 consequences emerging after implementation.

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

276 **DISCUSSION**

277 **Summary of findings**

278 Participants started by discussing balancing measures in terms of undesirable consequences which
279 were expected before or early in implementation (*trade-offs*) and which could offset some of the
280 intended benefits of improvements (*goals*). Although a range of examples were discussed, most
281 participants agreed that such measures were relatively rarely used. Participants additionally
282 emphasised that many consequences only became apparent after implementation, and these
283 unexpected consequences could be either desirable or undesirable (*pleasant* or *unpleasant*
284 *surprises*) and could accrue in the same part of the system as the improvement work, or other parts.
285 There was frequent confusion as to what a balancing measure should measure, since the implication
286 of many existing framings^{7,8} is that balancing measures are distinct in some way from process and
287 outcome measures, rather than any type of consequence being measurable in terms of processes
288 and outcomes. Involving front-line staff in identifying unintended consequences and balancing
289 measure design was perceived to increase engagement with improvement and reduce resistance to

290 change. Balancing measures were seen as a necessary and integral part of evaluating the impact of
291 an improvement programme, as well as a pragmatic way of engaging sceptics constructively by
292 understanding their legitimate concerns around implementation. However, the value of designing
293 and implementing balancing measures has to be offset against their cost in the context of overall
294 measurement burden.

295 **Strengths and limitations of the study**

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

309 **Comparison with existing literature**

310 The existing improvement literature on measurement design emphasises the importance of
311 developing a balanced set of measures during the planning of an improvement programme,^{7-10 31-33}
312 often distinguishing between process and outcome measures for goals, and balancing measures for
313 expected undesirable consequences (*trade-offs*) which are easily predictable from the outset (Box
314 1). However, participants in this study found this framing too narrow because they were concerned
315 about unexpected undesirable consequences (*unpleasant surprises*) and valued unexpected
316 desirable consequences (*pleasant surprises*), neither of which could be defined prior to intervention
317 implementation.

318 Although there are some studies of *trade-offs*,³⁴⁻³⁶ and *pleasant*³⁷ and *unpleasant*^{38 39} *surprises* (Table
319 2), published improvement interventions rarely report data relating to unintended consequences.¹⁵⁻
320 ^{19 40} This may partly reflect publication bias, since authors are known to emphasise positive results
321 and “*tuck away*”⁴¹ negative contextual features and failures.²³ However, it also likely reflects more
322 general lack of consideration or measurement of unintended consequences, consistent with an

323 observed preoccupation with measuring pre-specified local processes and outcomes (*goals*).⁴²⁻⁴⁴ The
324 implementation of PDSA cycles in healthcare for example has been criticised for often involving an
325 over-simplified “*Do, Do, Do*” approach¹⁵ focused on little and often measurement and delivery of
326 *goals* at the expense of thinking ahead and looking to the future (for *trade-offs*) and reflecting on
327 potential hazards during implementation (for *surprises*).^{45 46}

328 **Implications for improvement programme design**

329 Balancing measures are an integral and core element of commonly used improvement models like
330 the IHI Model for Improvement,^{7 8} but they are sometimes poorly specified and do not appear to be
331 commonly implemented in practice.^{15-19 40} Based on the literature and the findings of this study, we
332 believe that rather than focusing on balancing measures to implement at the start of improvement,
333 improvers and leaders at all levels of management should consider how best to achieve a balanced
334 accounting of the overall impact of improvement across the life of a programme. This requires
335 consideration of all four types of consequence, any of which can be measured in terms of process
336 and outcome. (Figure 2) Such a balanced accounting of impact can be achieved by articulating clear
337 assumptions and formulating explicit predictions for both *goals* and *trade-offs* before
338 implementation,^{14 40 47} and having a planned improvement pause after implementation to
339 deliberately step back from *goal* delivery to take stock and reflect on potential *surprises*.^{46 48} In an
340 ideal world, improvers would consult the available evidence base and seek external input from key
341 stakeholders in order to identify potential *trade-offs*, speculate on and investigate potential
342 *surprises*, and if necessary, to design relevant process and outcome measures to account for them.
343 However, improvement takes place in resource-constrained environments, which will confine what
344 is possible, including for example, the feasibility of measurement in other areas of a complex system.
345 Focusing on a balanced accounting rather than balancing measures also emphasises that qualitative
346 methods have much to offer both for the identification of *trade-offs* before implementation, and for
347 understanding *surprises* after implementation where retrospective measurement may be difficult.⁴⁹

348 ⁵⁰

349 **Implications for reporting quality improvement projects**

350 Few improvement reports mention unintended consequences, despite the Standards for Quality
351 Improvement Reporting Excellence (SQUIRE) guidance¹⁴ including a requirement that reporting
352 should include “*unintended consequences, such as unexpected benefits, problems, failures or costs*
353 *associated with the intervention*” (standard 13e). Of note though is that the SQUIRE explanation and
354 elaboration for this standard⁵¹ focuses more on exploring variation in implementation effectiveness
355 and does not provide any examples of significant elaboration of unintended consequences. As the
356 volume of publications in QI is growing, modification of SQUIRE to clarify that improvement reports

357 *should* report any measured or qualitatively assessed unintended consequences, or report that these
358 were not assessed, would be helpful to contextualise any evidence presented about the
359 achievement of improvement goals.

360 **CONCLUSION**

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

372 **List of abbreviations**

373

374 **QI:** Quality improvement

375 **IHI:** Institute of Healthcare Improvement

376 **NHS:** The National Health Service

377 **EMR:** Electronic Medical Record

378 **PDSA:** Plan Do Study Act

379 **SQUIRE:** Standards for Quality Improvement Reporting Excellence

380

381 **DECLARATIONS**

382

383 **Acknowledgements**

384 This work was undertaken by and on behalf of The Scottish Improvement Science Collaborating
385 Centre (SISCC). We are grateful to our improvement experts for their time and willingness to co-
386 create the *“balanced accounting” framework*. We also acknowledge Julie Anderson and Anita Lee,
387 who provided invaluable feedback on the revised versions of this paper.

388 **Ethics approval and consent to participate**

389 The University of Dundee Research Ethics Committee waived ethical approval for this study (UREC
390 15069). Informed consent to participate in the study was obtained from all participants.

391 **Consent for publication**

392 Not applicable

393 **Availability of data and material**

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

396 **Competing interests**

397 The authors declare that they have no competing interests.

398

399 **Funding**

400 This paper presents core work from the Scottish Improvement Science Collaborating Centre, funded
401 by the Scottish Funding Council, Health Foundation, Chief Scientist Office, and NHS Education
402 Scotland with in-kind contributions from participating partner universities and health boards. The
403 funding bodies had no role in the design of the study and collection, analysis, and interpretation of
404 data and in writing the manuscript
405

406 **Authors' contributions**

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

411 **REFERENCES**

412

- 413 1. Merton RK. The Unanticipated Consequences of Purposive Social Action. *Am Sociol Rev*
414 1936;**1**:894-904.
- 415 2. Wachter RM. Expected and unanticipated consequences of the quality and information
416 technology revolutions. *JAMA* 2006;**295**:2780-3. DOI:10.1001/jama.295.23.2780
- 417 3. Han YY, Carcillo JA, Venkataraman ST, Clark RSB, *et al.* Unexpected increased mortality after
418 implementation of a commercially sold computerized physician order entry system. *Pediatrics*
419 2005;**116**:1506-12. DOI:10.1542/peds.2005-1287
- 420 4. Auerbach AD, Landefeld CS, Shojania KG. The tension between needing to improve care and
421 knowing how to do it. *N Engl J Med* 2007;**357**:608-13. DOI:10.1056/NEJMs070738
- 422 5. Bardach N, Cabana M. The unintended consequences of quality improvement. *Curr Opin Pediatr*
423 2009;**21**:777-82. DOI: 10.1097%2FMOP.0b013e3283329937
- 424 6. Dixon-Woods M, McNicol S, Martin G. Ten challenges in improving quality in healthcare: lessons
425 from the Health Foundation's programme evaluations and relevant literature. *BMJ Qual Saf*
426 2012;**21**:876-84. DOI: 10.1136/bmjqs-2011-000760
- 427 7. Institute of Healthcare Improvement: *Science of Improvement: Establishing Measures*.
428 [http://www.ihp.org/resources/pages/howtoimprove/scienceofimprovementestablishingmeasures.as](http://www.ihp.org/resources/pages/howtoimprove/scienceofimprovementestablishingmeasures.aspx)
429 [px](http://www.ihp.org/resources/pages/howtoimprove/scienceofimprovementestablishingmeasures.aspx). Accessed 21 Mar 2016.
- 430 8. Langley GJ, Nolan KM, Nolan TW, *et al.* The Improvement Guide: A Practical Approach to
431 Enhancing Organizational Performance. San Francisco,CA:Jossey-Bass;1996.
- 432 9. Clarke J, Davidge M, James L. How-to Guide for Measurement for Improvement, *Patient Safety*
433 *First* 2009.
434 [https://eoeleadership.hee.nhs.uk/sites/default/files/Patient%20Safety%20First%20How%20To%20G](https://eoeleadership.hee.nhs.uk/sites/default/files/Patient%20Safety%20First%20How%20To%20Guide%20measurement%20for%20improvement.pdf)
435 [uide%20measurement%20for%20improvement.pdf](https://eoeleadership.hee.nhs.uk/sites/default/files/Patient%20Safety%20First%20How%20To%20Guide%20measurement%20for%20improvement.pdf). Accessed 03 Feb 2016.
- 436 10. NHS Quality Improvement Scotland: *Introduction to Measurement for Improvement*.
437 [http://www.qihub.scot.nhs.uk/knowledge-centre/qualityimprovementtopics/measurement-for-](http://www.qihub.scot.nhs.uk/knowledge-centre/qualityimprovementtopics/measurement-for-improvement.aspx)
438 [improvement.aspx](http://www.qihub.scot.nhs.uk/knowledge-centre/qualityimprovementtopics/measurement-for-improvement.aspx). Accessed 21 Jan 2016.
- 439 11. Asti L, Magers JS, Keels E, *et al.* A quality improvement project to reduce length of stay for
440 neonatal abstinence syndrome, *Pediatrics* 2015;**135**:1494-e500. DOI: 10.1542/peds.2014-1269

- 441 12. Mittal V, Darnell C, Walsh B, *et al.* Inpatient bronchiolitis guideline implementation and resource
442 utilization. *Pediatrics* 2014;**133**:e730-e7. DOI: 10.1542/peds.2013-2881
- 443 13. Turkington P, Power M, Hunt C, *et al.* There is another way: empowering frontline staff caring for
444 acutely unwell adults. *Int J Qual Health Care* 2013;**26**:71-8. DOI: 10.1093/intqhc/mzt084.
- 445 14. Ogrinc G, Davies L, Goodman D, *et al.* SQUIRE 2.0 (Standards for Quality Improvement Reporting
446 Excellence): revised publication guidelines from a detailed consensus process. *BMJ Qual Saf*
447 2016;**25**:986-92. DOI: 10.1136/bmjqs-2015-004411
- 448 15. Reed JE, Card AJ. The problem with Plan-Do-Study-Act cycles. *BMJ Qual Saf* 2016;**25**:147-52. DOI:
449 10.1136/bmjqs-2015-005076
- 450 16. Curnock E, Ferguson J, McKay J, *et al.* Healthcare Improvement and Rapid PDSA Cycles of
451 Change: A Realist Synthesis of the Literature. *NHS Education for Scotland* 2012.
452 http://www.nes.scot.nhs.uk/media/1389875/pdsa_realist_synthesis.pdf. Accessed 03 June 2017.
- 453 17. Jones EL, Lees N, Martin G, *et al.* How Well Is Quality Improvement Described in the
454 Perioperative Care Literature? A Systematic Review. *Jt Comm J Qual Patient Saf* 2016;**42**:196-206.
455 DOI: 10.1016/S1553-7250(16)42025-8
- 456 18. Nicolay CR, Purkayastha S, Greenhalgh A, *et al.* Systematic review of the application of quality
457 improvement methodologies from the manufacturing industry to surgical healthcare. *Br J Surg*
458 2012;**99**:324-35. DOI: 10.1002/bjs.7803
- 459 19. Manojlovich M, Lee S, Lauseng D. A systematic review of the unintended consequences of clinical
460 interventions to reduce adverse outcomes. *J Patient Saf* 2016;**12**:173-9. DOI:
461 10.1097/PTS.0000000000000093
- 462 20. McDonald KM, Schultz EM, Chang C. Evaluating the State of Quality-Improvement Science
463 through Evidence Synthesis: Insights from the Closing the Quality Gap Series. *Perm J* 2013;**17**:52-61.
464 DOI: 10.7812/TPP/13-010
- 465 21. Pronovost P, Wachter R. Proposed standards for quality improvement research and publication:
466 one step forward and two steps back. *Qual Saf Health Care* 2006;**15**:152-3. DOI:
467 10.1136/qshc.2006.018432
- 468 22. Brainard J, Hunter PR. Do complexity-informed health interventions work? A scoping review.
469 *Implement Sci* 2016;**11**:127. DOI: 10.1186/s13012-016-0492-5

- 470 23. Jones, E. SIQINS: Strengthening Reporting of Quality Improvement Interventions and Methods in
471 Surgery. <https://lra.le.ac.uk/bitstream/2381/39747/1/2017-JONES-EL-PHD.pdf>. Accessed 29 May
472 2017.
- 473 24. The Scottish Patient Safety Programme (SPSP).
474 <http://www.scottishpatientsafetyprogramme.scot.nhs.uk/>. Accessed 11 May 2017.
- 475 25. Ritchie J, Lewis J, McNaughton Nicholls C, *et al.* Qualitative Research Practice. A guide for social
476 science students and researchers. 2nd ed. London: Sage Publications; 2003.
- 477 26. Ash JS, Sittig DF, Dykstra RH, Guappone K, *et al.* Categorizing the unintended sociotechnical
478 consequences of computerized provider order entry. *Int J Med Inform* 2007;**76**:21-7. DOI:
479 10.1016/j.ijmedinf.2006.05.017
- 480 27. Ash JS, Sittig DF, Poon EG, *et al.* The extent and importance of unintended consequences related
481 to computerized provider order entry. *J Am Med Inform Assoc* 2007;**14**:415-23. DOI:
482 10.1197/jamia.M2373
- 483 28. Bloomrosen M, Starren J, Lorenzi NM, *et al.* Anticipating and addressing the unintended
484 consequences of health IT and policy: a report from the AMIA 2009 Health Policy Meeting. *J Am Med*
485 *Inform Assoc* 2011;**18**:82-90. DOI: 10.1136/jamia.2010.007567
- 486 29. Rogers, EM. Diffusion of Innovations. 4th ed. New York: Free Press; 1995.
- 487 30. Campbell EM, Sittig DF, Ash JS, *et al.* Types of Unintended Consequences Related to
488 Computerized Provider Order Entry. *J Am Med Inform Assoc* 2006;**13**:547-56. DOI:
489 10.1197/jamia.M2042
- 490 31. Scottish Government: *Quality Improvement and Data Measurement-What Non-Executive*
491 *Directors need to know*. 2016. <http://www.gov.scot/Resource/0049/00492311.pdf>. Accessed 01
492 May 2016.
- 493 32. U.S. Department of Health and Human Services Health Resources and Services Administration:
494 *Performance Management and Measurement* 2011.
495 <http://www.hrsa.gov/quality/toolbox/508pdfs/performanceandmeasurement.pdf>
496 Accessed 09 Mar 2016.
- 497 33. Health Quality Ontario: *Measurement for Quality Improvement* 2013.
498 <http://www.hqontario.ca/Portals/0/documents/qi/qi-measurement-primer-en.pdf>. Accessed 13 Apr
499 2016.

- 500 34. Shojania KG, Ranji SR, McDonald KM, *et al.* Effects of quality improvement strategies for type 2
501 diabetes on glycemic control: A meta-regression analysis. *JAMA* 2006;**296**:427-40. doi:
502 10.1001/jama.296.4.427
- 503 35. Kavanagh PL, Sprinz PG, Wolfgang TL, *et al.* Improving the Management of Vaso-Occlusive
504 Episodes in the Pediatric Emergency Department. *Pediatrics* 2015;**136**:e1016-25. DOI:
505 10.1542/peds.2014-3470
- 506 36. Dewan M, Galvez J, Polsky T, *et al.* Reducing Unnecessary Postoperative Complete Blood Count
507 Testing in the Pediatric Intensive Care Unit. *Perm J* 2017;**21**:16-051. DOI: 10.7812/tpp/16-051
- 508 37. Duvoisin G, Fischer C, Maucort-Boulch D, *et al.* Reduction in the use of diagnostic tests in infants
509 with risk factors for early-onset neonatal sepsis does not delay antibiotic treatment. *Swiss Med Wkly*
510 2014;**144**:w13981. DOI: 10.4414/smw.2014.13981
- 511 38. Bell S, Davey P, Nathwani D, *et al.* Risk of AKI with Gentamicin as Surgical Prophylaxis. *J Am Soc*
512 *Nephrol* 2014;**25**:2625-32. DOI: 10.1681%2FASN.2014010035
- 513 39. Strom BL, Schinnar R, Aberra F, *et al.* Unintended effects of a computerized physician order entry
514 nearly hard-stop alert to prevent a drug interaction: a randomized controlled trial. *Arch Intern Med*
515 2010;**170**:1578-83. DOI: 10.1001/archinternmed.2010.324
- 516 40. Taylor MJ, McNicholas C, Nicolay C, *et al.* Systematic review of the application of the plan-do-
517 study-act method to improve quality in healthcare. *BMJ Qual Saf* 2014;**23**:290-8. DOI:
518 10.1136/bmjqs-2013-001862
- 519 41. Matosin N, Frank E, Engel M, *et al.* Negativity towards negative results: a discussion of the
520 disconnect between scientific worth and scientific culture. *Dis Model Mech* 2014;**7**:171-3. DOI:
521 10.1242/dmm.015123
- 522 42. Raleigh VS, Foot C. *Getting the measure of quality. Opportunities and challenges* 2010.
523 [https://www.kingsfund.org.uk/sites/files/kf/Getting-the-measure-of-quality-Veena-Raleigh-](https://www.kingsfund.org.uk/sites/files/kf/Getting-the-measure-of-quality-Veena-Raleigh-Catherine-Foot-The-Kings-Fund-January-2010.pdf)
524 [Catherine-Foot-The-Kings-Fund-January-2010.pdf](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.
- 525 43. Dixon-Woods M, Amalberti R, Goodman S, *et al.* Problems and promises of innovation: why
526 healthcare needs to rethink its love/hate relationship with the new. *BMJ Qual Saf* 2011;**20**: i47-i51.
527 DOI: 10.1136/bmjqs.2010.046227
- 528 44. Dixon-Woods M, Martin G. Does quality improvement improve quality? *Future Hospital Journal*
529 2016;**3**:191-4. DOI: 10.7861/futurehosp.3-3-191

530 45. Vincent C, Burnett S, Carthey J. The measurement and monitoring of safety. Drawing together
531 academic evidence and practical experience to produce a framework for safety measurement and
532 monitoring. *Health Foundation Spotlight report* 2013.
533 [http://www.health.org.uk/sites/health/files/TheMeasurementAndMonitoringOfSafety_fullversion.p](http://www.health.org.uk/sites/health/files/TheMeasurementAndMonitoringOfSafety_fullversion.pdf)
534 [df](http://www.health.org.uk/sites/health/files/TheMeasurementAndMonitoringOfSafety_fullversion.pdf). Accessed 29 July 2016.

535 46. Vincent C, Burnett S, Carthey J. Safety measurement and monitoring in healthcare: a framework
536 to guide clinical teams and healthcare organisations in maintaining safety. *BMJ Qual Saf*
537 2014;**23**:670-7. DOI: 10.1136/bmjqs-2013-002757

538 47. McQuillan RF, Silver SA, Harel Z, *et al*. How to Measure and Interpret Quality Improvement Data.
539 *Clin J Am Soc Nephrol* 2016;**11**:908-14. DOI: 10.2215%2FCJN.11511015

540 48. Lee SL. The Extended Surgical Time-Out: Does It Improve Quality and Prevent Wrong-Site
541 Surgery? *Perm J* 2010;**14**:19-23.

542 49. Powell AE, Rushmer RK, Davies HTO. A systematic narrative review of quality improvement
543 models in health care. *Healthcare Improvement Scotland* 2009.
544 [http://www.healthcareimprovementscotland.org/previous_resources/hta_report/a_systematic_nar](http://www.healthcareimprovementscotland.org/previous_resources/hta_report/a_systematic_narrative_review.aspx)
545 [rative_review.aspx](http://www.healthcareimprovementscotland.org/previous_resources/hta_report/a_systematic_narrative_review.aspx). Accessed 20 March 2017.

546 50. Portela MC, Pronovost PJ, Woodcock T, *et al*. How to study improvement interventions: a brief
547 overview of possible study types. *BMJ Qual Saf* 2015;**24**:325-36. DOI: 10.1136/bmjqs-2014-003620

548 51. Goodman D, Ogrinc G, Davies L, *et al*. Explanation and elaboration of the SQUIRE (Standards for
549 Quality Improvement Reporting Excellence) Guidelines, V.2.0: examples of SQUIRE elements in the
550 healthcare improvement literature. *BMJ Qual Saf* 2016;**25**:e7. DOI: 10.1136/bmjqs-2015-004480

551 52. Dayan M, Edwards N. Learning from Scotland's NHS. *Nuffield Trust Research report* 2017.
552 <https://www.nuffieldtrust.org.uk/research/learning-from-scotland-s-nhs>. Accessed 27 August 2017.

553

554

555

556

557

558

559

560 **Box 1: Institute of Healthcare Improvement recommended types of measures**⁷⁸

“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?”

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

562 **Table 1: Characteristics of participants in both phases of the study**

Participants' primary roles and responsibilities	Setting	Phase 1: Semi-structured interviews N=15*	Phase 2: Focus group interviews N=24**
Improvement advisors 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 academic base and perspective , 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
Providers of healthcare services 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
Providers of social care services 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
Policy-makers and commissioners involved in monitoring performance and setting the general direction of quality improvement	Health care, social care, Education and early years	3	2
Patient representatives 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

563 **All interviews were conducted face to face except two academics and one policy maker which were interviewed by telephone;*

564 ***Two participants in the interviews (one improvement advisor and one academic) also attended the focus groups;*

565
566

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

568

Study	Improvement goals	Other consequences	Examples of balancing measures prior, during and post implementation
Kavanagh 2015 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 (trade-offs)	<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, <i>episodes of respiratory depression</i> and <i>inpatient length of stay</i> .
Dewan 2017 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 (trade-offs)	Balancing measures were implemented for <i>haemoglobin level below 8 g/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 2014 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 (pleasant surprises)	There was pre-intervention 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 2014 ³⁸	To reduce the pre-operative use of antimicrobials associated with Clostridium difficile infection	Unexpected undesirable consequences (unpleasant surprises)	The new surgical prophylaxis regimen of four doses of flucloxacillin 1g plus single dose gentamicin 4mg/kg unexpectedly led to <i>increased rates of post-operative 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 2010 ³⁹	To evaluate the effectiveness of a customized nearly hard stop alert in reducing concomitant orders for warfarin and trimethoprim- sulfamethoxazole compared with the standard practice of a pharmacist intervention program.	Unexpected undesirable consequences (unpleasant surprises)	Unexpected <i>delays in indicated anticoagulant and/or antimicrobial treatment initiation</i> were deemed sufficiently serious to warrant discontinuation of the improvement intervention.

569 **Figures/illustrations**

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

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