Universiteit
Leiden
The Netherlands

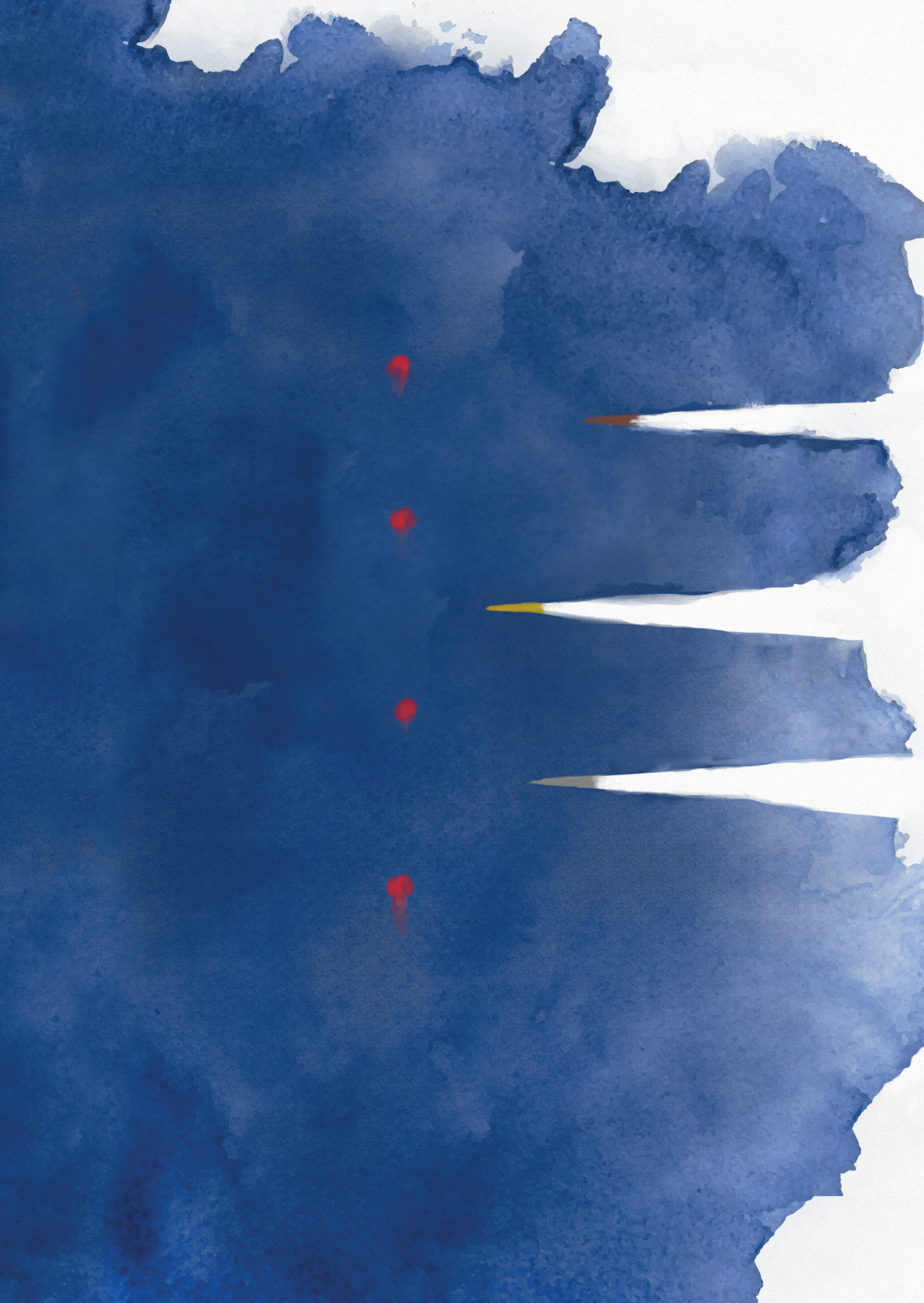
What makes the best performing hospital? the IQ Joint study

Schie, P. van

Citation
Schie, P. van. (2023, November 8). *What makes the best performing hospital?: the IQ Joint study*. Retrieved from https://hdl.handle.net/1887/3656771

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3656771

Note: To cite this publication please use the final published version (if applicable).



Chapter 9

Summary, general discussion, and future perspectives

The objective of this thesis was to study how arthroplasty registries can improve their feedback to orthopaedic surgeons in order to give direction to quality improvement initiatives (QII) that improve care for total hip and knee (THA and TKA). A second aim was to evaluate the effectiveness of such improved feedback on patient outcomes.

Giving feedback on performance indicators is a frequently used approach to improve the quality of care delivered. In this context, feedback is defined as the provision of clinical outcome summaries to healthcare providers or organizations intended to initiate activities to improve the performance of delivered care(1,2). Internationally, feedback from arthroplasty registries is provided in various ways. In the Netherlands, clinical outcomes are shown at the hospital-level in a real-time secured web-based dashboard from the LROI. The extent of between-hospital variation is shown in an anonymized version in annual reports.(3,4) The effect of feedback varies (i.e., from a 9% decrease to a 70% increase), but an optimal design will reasonably improve patient care.(5) This thesis provided an overview of national and international between-hospital variation in clinical outcomes to investigate whether improvement is achievable. In order to optimize the content of the feedback, methodological studies have been performed to investigate whether outlier hospitals can be detected earlier, whether reasons for higher revision rates can be identified, and a composite outcome measure is developed and tested. As awareness of performance by surgeons in combination with motivation to improve is more likely to result in targeted QII improving quality, associations with such awareness were assessed to increase feedback effectiveness.(6-8) The knowledge obtained in combination with up-to-date theory for providing effective feedback was incorporated in a multifaceted quality improvement intervention and tested on its effectiveness in a cluster randomised controlled trial.(5,9-11)

This chapter starts with a summary of the main findings, including the practical implications of the previous chapters. Subsequently, these findings and relevant methodological issues are discussed in the context of available literature, and finally, recommendations for future practice and research are given.

Summary and practical implications

Arthroplasty registries were initially established to compare implant survival and monitor the safety of different orthopaedic implants, like total hip and knee arthroplasties (THA and TKA). In recent years, however, registries have also been as quality systems across the healthcare system to show the variation between hospitals for numerous clinical outcome measures, thus providing feedback to hospitals on their performance. The latter is usually compared with a reference standard (i.e., the

benchmark).(12-19) Most registries provide feedback through annual reports intended to encourage QII in low-performing hospitals.(12-19) However, this information needs to be viewed by an action of the orthopaedic surgeon (i.e., log into the secured website to view the data within a secured site). Furthermore, also top-performing hospitals may be interested in comparing their performance with hospitals from other countries to stimulate further improvement within specific domains. Fair international hospital comparison is only achievable when consistent outcome definitions are used, as these will determine the frequency of occurrence.(20-23) Consistency in outcome measure definitions makes it also possible to merge international data providing better opportunities to detect rare safety issues earlier (e.g., the metal-on-metal hip arthroplasty disaster, modular femoral neck corrosion etc.), which will prevent thousands of patients from being exposed to poor performing implant designs, thus decreasing unnecessary suffering in future.(24,25) **Chapter 2** showed that among registry reports and arthroplasty cohort studies: revision, readmission, and complications are the most frequently reported clinical outcomes, with considerable differences in their outcome rates between hospitals, indicating a vast improvement potential for at least some hospitals. However, part of the variation may be explained by the significant heterogeneity in the following domains: 1) outcome definitions, including what is a revision, readmission, or complications, 2) duration of follow-up and starting point of follow-up, 3) characteristics included in patient-mix adjustment, and 4) type of patients- and hospital included. This thesis showed that revision of the implant within five years, readmission within 30 days, and complications up to 2 years postoperative were the most commonly used outcome measures in arthroplasty reports. However, none of these definitions had a perfect agreement with the other domains for THA, TKA, and THA&TKA combined. The least consensus was found on whether or not to adjust for patient characteristics, let alone which characteristics should be included in the adjustment. Although the latter as well as the other domains investigated in this study are essential for fair hospital comparison.(23) In the future, partnerships of arthroplasty registries such as the International Society of Arthroplasty Registries can play a leading role, not only in international collaboration but also striving for more uniformity in the definitions and methods used.(26)

Reporting the between-hospital variation in clinical outcomes and identifying positive and negative outlier hospitals is a simple and effective way to get insight into hospital performance, provided that the rankability (i.e., the percentage of total variation that is explained by "true" hospital differences rather than chance variation) is acceptable. (27-30) However, assessing how to pursue improvement for a given clinical outcome can be challenging, particularly for summary outcomes such as all-cause revision, but this can be facilitated by examining specific indications for revision that may be the reason for the identified worse performance on all-cause revision. **Chapter 3**

showed large variation in 1-year all-cause revision rates between Dutch hospitals with moderate rankability (61%) for THA and low rankability (46%) for TKA within a 3-year time frame, indicating huge improvement potential for a considerable part of the Dutch hospitals. Earlier detection of poor performance using a 1-year time frame has the advantage that QII can be introduced earlier; however, this resulted in low rankabilities and is therefore not recommended. Underlying reasons for worse performance on all-cause revision were found for 12 of the 13 negative outlier hospitals for THA and 3 of the 7 for TKA, mainly consisting of infection (both for THA and TKA) and dislocations (only THA). Implant loosening and technical failure (only TKA) were less likely to be the underlying reason for the worse performance. Rankabilities for the specific indications for revision were all low within a 3-year time frame for THA and TKA, except for infection for THA, for which the rankability was moderate (i.e., 61%). As rankabilities within a 1-year timeframe were all low, it is recommended to use a 3-year time frame to identify underlying reasons for worse hospital performance on all-cause revision.

Where **chapter 3** showed that earlier detection of poorer performance could not be done reliably using funnel plots within a 1-year time frame, the monthly monitoring of revision rates using CUSUM charts with 5 control limits shown in **chapter 4** was able to detect worsening performance earlier than the conventional funnel plots. The first signal for negative outliers was generated at a median of 18 months for THA and 21 months for TKA within a 3-year time frame. CUSUM charts thereby enable detection of deteriorating patterns earlier, making it possible to introduce QII earlier than waiting for the results to appear in the funnel plot after 3 years. This thesis adds to the existing literature how much earlier a signal was generated and with what accuracy (i.e., 97% both for THA and TKA) compared with the traditional funnel plot with a 3-year time frame. These results are highly relevant for registries and scientific associations deciding whether to implement CUSUM charts in their organisation to improve quality.(31) The results on accuracy will contribute to professionals' confidence in CUSUM charts. In response to these findings, the Dutch Arthroplasty Register (LROI) has added CUSUM charts to their routine dashboard reporting on clinical outcomes to provide hospitals with a tool for earlier detection and thereby the opportunity to introduce QII earlier to improve patient care.(32)

By increasing the number of events, composite outcomes may also enable that differences between hospitals in their performance to be detected sooner.(33-36) In **chapter 5**, an ordered composite outcome with all combinations of clinical outcomes (i.e., revision, readmission, complications, and upper-quartile LOS) ranked from best to worst according to the patient's perspective was developed and tested on its ability to differentiate between hospitals. The newly developed composite showed higher

rankability than individual clinical outcomes due to the larger variation between hospitals when more information is included. The composite could reliably differentiate between hospitals in their performance using a 1-year time frame, rather than requiring the usual 3-year time frame, allowing the introduction of QII earlier. An additional advantage is that the composite can measure more aspects of delivered quality of care as hospitals may perform well on one outcome while at the same time performing worse on another. The composite overcomes this issue and shows whether a patient had a revision but also whether they were readmitted, experienced complications, or had a prolonged LOS indicating more specifically where improvement is possible (e.g., in patients with a normal LOS, without complications who were readmitted). The new composite is widely applicable as it may help healthcare providers to select for which patient groups' medical records have to be reviewed to investigate whether and how care can be improved. For instance, rather than reviewing all records for patients who were readmitted, it allows to selectively review only those with a normal LOS without any registered complications, to understand e.g., whether information at discharge needs to be improved to avoid readmission or discuss whether the readmission was needed or could have been treated at the outpatient clinic. Another application is that it is visible in one outcome whether a focus to improve, for example, length of stay, does not come at the expense of another outcome (e.g., readmission). Finally, patients can simply check how often a procedure goes as planned (i.e., without any clinical outcome occurring).

Clinical outcomes such as revision, readmission, complications, and LOS measure unintended adverse events and generally occur with low frequency for THA and TKA.(37) However, up to 10% and 20% of patients following THA and TKA, respectively, are dissatisfied with results, mainly related to persistent pain and disability.(38,39) Patient-Reported Outcome Measures (PROMs), on the other hand, measure the intended outcomes such as pain reduction, functionality improvement, and health-related quality of life, and would therefore complement these clinical outcomes by identifying potential additional areas for improvement in these intended outcomes.(40-44) Similar to the need for a high level of data completeness regarding clinical outcomes to ensure there is no selection bias, we also need high response rates of patients completing both pre- and postoperative questionnaires to allow calculating the improvement in PROMs, or if that is not feasible at least gain insight into how those who complete questionnaires are a selection of all patients. Dutch response rates, as in other national and regional arthroplasty registries, were low (i.e., less than 61% of patients completed the preoperative PROM questionnaire, and only about 40% of patients completed preoperative and postoperative PROM questionnaires) compared with the above 98% completeness of revision, surgical procedure, implant, and patient characteristics data in registries for both THA and TKA.(40-45) Previous

studies have already shown differences in various patient characteristics, such as patients completing questionnaires, in general, being healthier, more likely to be white, with higher literacy rates, and lower rates of cognitive impairment, including dementia.(46-48) To better understand whether missing PROM data for THA and TKA may result in under- or overestimation of PROMS improvement scores, **chapter 6** used clinical outcome rates (i.e., revision, readmission, complications, and upper-quartile LOS) to examine whether these differed between respondents and non-respondent, as well as their association with PROM improvement scores. Chapter 6 showed that respondents to PROM questionnaires less often experienced adverse events. This likely results in an overestimation of the clinically relevant improvement in PROMs as adverse events were associated with a lower likelihood to achieve a clinically relevant PROM improvement. Given the observed patient-level associations in **chapter 6**, it is likely that initiatives to improve the quality of care by reducing readmission, complications, and long LOS for both THA and TKA patients will lead to more patients achieving clinically relevant improvement in HOOS-PS and KOOS-PS scores. Hospital differences in PROM response rates were not associated with differences in adverse event rates, suggesting that estimated between-hospital differences in PROM improvement are likely unaffected.

Feedback may be methodologically sound, but it is only effective if it is viewed and interpreted by practitioners. **Chapter 7** showed that only half (i.e., 55%) of Dutch orthopaedic surgeons performing THA and TKA were aware of their outlier performance status regarding revision rates. Awareness was higher among surgeons that more often logged in on the LROI dashboard, more often interpreted funnel plots correctly, and more often could recall the 1-year revision rates of their surgeon group. Thirty-eight percent of THA and 26% of TKA surgeons met all three conditions necessary to act upon the feedback information, i.e., logging in, correct interpretation of funnel plot, and could recall their 1-year revision rates. Forty-five percent of surgeons in a hospital identified as a negative outlier reported not seeing their worsening performance coming, meaning they continued to provide care without modifications. Thus, a focus on making feedback more effective is very important, as 85% of surgeons indicated that they did start QII once being identified as having worse performance. Logging in on the LROI dashboard should be made more attractive and encouraged, for example, by emphasizing the importance of already reported clinical outcomes (e.g., revision rates) and adding new outcomes such as prosthesis survival, complications, readmissions, and length-of-hospital-stay as these are considered relevant by a large part of surgeons. Second, teaching material should become available to improve interpretation skills of statistical presentation of data like funnel plots, CUSUM charts etc., or provide explanatory text with these funnel plots on the LROI dashboard to help surgeons understand what the data in the figure

represent. Third, feedback should be sent to become easily accessible, readily read (e.g., infographics), and tailored to single hospitals rather than expecting surgeons to make any selections and investigate themselves.

The knowledge obtained from the previous chapters with contemporary theory for providing effective feedback was incorporated into a multifaceted quality improvement intervention and tested on its effectiveness in **chapter 8**.(5,9-11) The intervention was applied over eight months and included monthly feedback, education on the interpretation of the feedback, and an action implementation toolbox including evidence-based QII. Hospitals that received the intervention improved 4.3% more on the Textbook Outcome (i.e., the absence of revision, readmission, complications, and long LOS, the best possible outcome in the composite developed in **chapter 5**) compared with control hospitals. It was found that the effect size was larger for intervention hospitals that introduced QII, suggesting that these QII were likely the reason for the better outcomes. The median number of TKA surgeries performed was considerably smaller than for THA in intervention hospitals, which could explain why the effect for TKA was not significant when outcomes for THA and TKA were analysed separately, even though the effect size pointed in the same direction of improvement. In addition, the difference could be explained by the lower baseline risks for revision and complications for TKA, associated with smaller absolute risk reduction. These findings in **chapter** 8 support that frequent feedback to surgical teams should be supplemented by interactive education and facilitated by evidence-based QII tailored to specific outcomes to improve the quality of care regarding THA and TKA effectively.

General discussion

Components of the quality intervention

Quality improvement interventions are a common strategy to improve patient outcomes, but with highly variable effects across studies.(5,49,50) Two meta-analyses show that quality interventions using only one single intervention component are less effective, with little to no improvement.(51,52) Thus, quality interventions should be designed in a multifaceted way, including components addressing, for example, the gap in knowledge and (surgical) skills, but may also address other components such as audit and feedback that will allow to evaluate the impact of changes made and spur further improvement actions. Even though education on quality parameters is needed in most quality improvement efforts, serious limitations are present when they are used without proper context and a predetermined goal.(9) Education can only be effective if it solves a knowledge problem which is a barrier for quality

improvement. Nevertheless, designing proper multifaceted interventions is less than easy and straightforward. The quality intervention reported in **chapter 8** serves as a model for future intervention efforts by combining theory and previous evidence, thus increasing the likelihood of effectiveness in daily clinical practice.

The first intervention component consisted of monthly updated feedback sent by email to all individual orthopaedic surgeons. The monthly time interval was chosen based on previous evidence that feedback on performance is more effective when given repeatedly and not only once. The reason might be that recipients are more likely to perceive such repeated feedback as more relevant and accurate than only once a year since monthly data feedback is closer to current performance and can thus be easier related to current clinical practice.(5,53) Furthermore, these monthly feedbacks also allow for timely evaluation of the introduced QII, and thus any subsequent improvement actions without delay can be done. Arthroplasty registries can support this as data are routinely collected that could easily be used for near real-time monitoring of clinical outcomes. Sending the feedback by email was based on orthopaedic surgeons' preference of receiving feedback by email and our finding that 33% of orthopaedic surgeons never logged in on the LROI dashboard (**chapter 7**). This adds to the literature regarding the most effective mode of delivery (e.g., electronic, paper, face-to-face) and frequency (e.g., monthly, quarterly, or yearly) which have not been well assessed to date. Moreover, our intervention tapped in on previous knowledge that feedback is more effective when it is also discussed orally in a group, preferably by a senior colleague, rather than just presenting written data to individuals.(5) As for the former, the researcher visited all intervention sites to explain the feedback orally, although this intervention might have been improved if a senior colleague would visit the hospitals.(54,55) Another important aspect we considered was the sender of the feedback. It is known that when feedback comes from a regulatory body, recipients may be more likely to activate affective processes (e.g., distress), distracting attention from the specific task requiring change. In contrast, the current intervention was nested in the registry from which they "trusted" data and procedures for how data were collected and processed.(1)

As for the feedback, the comparator hospitals are essential in helping clinicians assess their performance and identify differences between current and desirable performance.(56) However, the choice of comparators may have critical implications for what message is conveyed by the feedback and how recipients will react.(57) For this reason, we performed stratified randomisation to ensure that, e.g., academic hospitals treating complex cases would not only be compared with private hospitals treating only the healthiest patients (which can never be captured entirely by case-mix adjustment). We included funnel plots in the feedback, as this may already be

a familiar visual presentation because it was available in the LROI dashboard and annual reports, although we realized this required some education based on our findings that 39% of the orthopaedic surgeons did not interpret the funnel plot correctly (**chapter 7**). The funnel plot allowed participants to compare performance with other hospitals. However, the funnel plot will only give an average estimate of performance across a period e.g., 3 years, so we added a CUSUM chart to more clearly show trends over time e.g., indicating whether hospitals were moving in the right direction after implementing a new QII. This trend of performance change (as shown in CUSUM charts) is more motivating to introduce new QII than the distance between performance and best performers (as shown in funnel plots).(58,59) In addition, trends increase the credibility of feedback and enable the introduction of QII according to the *Plan-Do-Study-Act (PDSA)* cycle, in which recipients continuously self-assess their performance and the effect of QII when deciding whether or not to take action. Trends, therefore, add meaningful information and should be added by default to performance feedback. In addition, **chapter 4** showed that changes in hospital performance were detected 18 months and 15 months earlier for THA and TKA, respectively, than the conventional funnel plots using a 3-year time frame.

With regard to comparing hospital performances, a recently published review suggests that comparison with high-performing peers is preferred over benchmarking hospitals to a national average, as it shows that top performance could be achieved.(60) In addition, psychological theories suggest that clinicians are less likely to accept an ”externally imposed” performance goal (e.g., 1-year revision rate of less than 1% imposed by an outside party) and that recipients in such cases are more likely to reject feedback recommendations and pursue self-conceived performance levels.(56,58,61). The funnel plot fits well with this understanding, as the performance compared to other hospitals and performance outcomes can be adjusted for differences in case-mix. Due to case-mix adjustment, hospitals that mainly treat patients without comorbidities and therefore expected to have lower frequencies of adverse events could be fairly compared with hospitals that mainly treat patients with multiple comorbidities.(20-22) However, when large differences in hospital performances consist, low performers may experience the feedback as unfeasible and reject the feedback. Tailoring the feedback to individual hospitals could avoid feedback rejection, for example, by comparing low performers with the top 50% and average performers with the top 10%.

The second intervention component consisted of education to interpret the feedback on performance, which was provided during an on-site visit in the first month of the intervention. The latter was attended by a majority of the orthopaedic surgeons within a hospital (**chapter 8**). An online educational video and pocket card containing a summary of the educational meetings were available as a reference. Education was

needed as **Chapter 7** showed that 39% of orthopaedic surgeons could not interpret funnel plots correctly. Even more, they often overestimated their performance if unaware of their performance data.(56) Even more important, "unawareness" will limit the undertaking of necessary QII, because it is assumed that performance is good enough even though there may be room for improvement. It seems evident that correctly interpreted feedback will improve the quality of care, as 17 out of 20 orthopaedic surgeons indicated that they would conduct QII when becoming aware of worse performance compared to the national benchmark. At the end of the educational meeting, explicit goals and specific actions for improvement were discussed, as this will improve implementation and intervention effectiveness.(5,49) This also aligns with theories that goals aimed for can make feedback more tangible for clinicians and thus help to facilitate better-focused action plans, which facilitate steps needed to achieve predefined goals.(62,63) The improvement process is more effective when goals are considered specific, measurable, achievable, relevant, and time-bound, allowing the recipients' attention to be more productive on the task. (53,58,64) We encouraged surgeon groups to set their own goals and create their own action plans as goals may otherwise not be acceptable for a subset of clinicians, even if a credible orthopaedic authority set them (e.g., a national association).(56,58,61) Not embracing the set goals will improve the chance of feedback rejection and impede intentions to improve care, thus likely diluting the effects of the quality interventions. (61) In this way, we also aimed for an established engagement with goals and action plans. To further stimulate engagement, bimonthly questionnaires were sent to all orthopaedic surgeons allocated to the intervention group to verify compliance with self-set goals and action plans. Furthermore, monitoring of progress in achievement of these plans was done.

The final intervention component was an action implementation toolbox including evidence-based QII for each clinical outcome reported in the feedback, which was added to overcome the barrier of translating feedback into what needs to be improved in clinical practice. Clinicians have often been shown to lack the skills or knowledge to interpret statistical feedback and formulate what QII is necessary to improve. (5,7,65,66) The toolbox bridges this gap and lowers the barrier to implementing evidence-based quality improvement initiatives. Adding a toolbox to a quality intervention has shown to be an improvement compared to feedback alone, but only in process indicators and not clinical outcomes.(11) As shown in **chapter 8**, our intervention showed an improvement in clinical outcomes, which could be due to the fact that we included evidence-based QII in the toolbox that targeted the outcomes. In contrast, the toolbox in the study of Roos-Blom et al. mostly targeted process measures, such as the availability of a protocol rather than the outcomes.(11)

In this thesis, we worked on a model for the design of a quality improvement intervention, where insight was gained into potential barriers of the target group in **chapter 7,** and feedback was methodologically improved in **chapters 3, 4, and 5**. Gained knowledge on these topics was supplemented with theory to provide effective feedback (i.e., education, discussing goals and actions, and a toolbox). However, there were also indications for intervention improvement as four of the ten intervention hospitals indicated that they needed additional information on the interpretation of funnel plots and CUSUM chats despite the offered educational session. Two hospitals indicated that they would appreciate more QIIs in the toolbox, and seven hospitals indicated that they would like to be matched with hospitals to exchange information on best practices and identify areas for further improvement. Unfortunately, the latter was initially planned but was not executed due to government restrictions related to the COVID-19 pandemic.(67) Finally, the intervention period could have lasted longer than eight months, as the bimonthly surveys showed that some intervention hospitals started implementing QII after several months. Therefore, it is possible that the end effect of a QII has not yet been achieved at the time the intervention was evaluated after eight months. The intervention, as reported in this thesis, will probably not meet the target group's needs in the future as it is possible that new components would fit better with barriers at that time. This makes designing an appropriate quality intervention, like quality improvement, a continuous improvement process.

However, similar to other multifaceted quality interventions that were tested, it is unclear to what extent each single component of the intervention (i.e., feedback, education, and an action implementation toolbox) contributed to the 4.3% (95% confidence interval 4.30% to 4.34%) absolute improvement in the intervention compared to control hospitals.(5,49,66) **Chapter 8** showed that the intervention effect was most likely achieved through the introduction of targeted QII, making this likely the causal link to the improved patient outcomes, demonstrating that if surgeons are sufficiently engaged to introduce QII, it will improve patient care.(68)

Sustainability

Even if a quality improvement intervention positively affects the quality of care delivered, maintaining access to resources available during the intervention is likely needed to sustain the improvement or even continue to achieve further gains. Little is known about why a quality intervention is sustainable, as most empirical data demonstrate a lack of sustainability, and only a few studies report on sustainability and adoption in everyday practice after the initial improvement initiative ended.(69-72) Implementing a package of common quality interventions (e.g., feedback, education, alerts) as a quick fix to resolve poor hospital performance may then provide a temporary solution but is generally unsustainable.(69,70,73) In the end, it is not

the number of implemented quality interventions by a hospital that is a measure of success but rather the ability to sustain the interventions in the long term.(74-76) To achieve that goal, an effective and sustainable intervention must offer a solution for the underlying problem as the first step, but it also needs to be adapted to the environment and use resources that will continue to be available after the intervention ends to be sustainable and become part of everyday practice.(73,77)

During intervention design, the application of the intervention components by the Dutch Arthroplasty Registry (LROI) was considered, so that intervention components and resources would remain available as much as possible if the quality intervention proved effective. For example, the CUSUM charts shown in **chapter 4** to enable earlier detection of worsening performance, have already been implemented and are currently part of the routine LROI dashboard as well as promoted through communication by the Dutch Orthopaedic association to reach other hospitals that did not take part in the IQ Joint study.(78) In addition, following the IQ Joint study, it is now also possible to register complications in the LROI database in order to further improve the quality of care and safety. The education video to explain the statistical feedback information and pocket card to be used as a reminder in clinical practice remained available, as well as the action implementation toolbox used during the study. The latter must be kept up-to-date and potentially further expanded if new evidence-based QII effectively improves the targeted clinical outcomes. Furthermore, to promote continued engagement with improving the quality of care as an integrated part of orthopaedic patient care, an annual educational session or workshop may act as a stimulus and platform for exchanging best practices to motivate the hospitals to improve their care continually. Finally, the IQ joint study group provided a positive and safe improvement climate, where the created collaboration of hospitals may continue to improve the quality of care in future projects.

Future perspectives

Relevant future directions of research to increase the effectiveness of quality improvement initiatives and prevent adverse effects are mentioned in the following section. The focus is on further developing quality interventions for a specific problem and determining the effect of the prosthesis on the quality of care

Tailoring quality interventions to a specific problem in practice

Even though there is evidence that some bundled interventions are more effective in improving care, the complicated matter of how quality interventions exercise their effect needs to be further unraveled, as engagement with and impact of interventions

is variable.(5,79) The lack of systematic, coordinated research in this field and gaps in studies reporting on the likely mechanism of the intervention effect perpetuates this problem. Only reporting outcome effects without concurrent publication of the process evaluation leaves the reader to guess at the reason for the main findings because they cannot learn from the underlying processes or barriers that may or may not have affected the outcome. Process evaluation is therefore needed to contextualize and understand the effect of the intervention.(80) Process evaluations can focus on the uptake of intervention components and are often based on a mix of interviews, focus groups, and field notes, which can substantiate the intervention's fidelity. For example, as was reported in **chapter 8**, the number of study participants interpreting the feedback, attending educational sessions, and introducing QII were reported. More knowledge on factors hindering or facilitating effective interventions is needed.

Just as care is tailored to the individual patient, quality interventions should also be tailored to the daily barriers and problems clinicians face. Initiatives that truly help clinicians achieve their goals are likely to be well-received, unlike many quality interventions leading to unenthusiastic engagement and unsuccessful outcomes due to little clinical relevance.(49,50) For that matter, some studies state that more than a third of participants who sign up for a quality intervention are not actively participating in the intervention.(81) These numbers are unlikely to increase as long as a one-size-fits-all approach is maintained that focuses on the outcome from a theoretical perspective rather than a clinically relevant perspective, understanding the resources and skills required to achieve better performances. To address this problem, the participating hospitals defined which clinical outcome they wanted to improve (i.e., revision, readmission, complications, and long LOS). Nevertheless, we could still assess the intervention's effect by focusing on a composite outcome as the primary outcome, including all clinical outcomes. In addition, prior to the implementation of the intervention, insight was obtained into the desired resources and skill level of clinicians. The latter requires investments in improving methods to provide information clinicians need to improve on but also ensures a better understanding of the clinical physicians' attention we are trying to capture and the behaviour we are trying to change. Who are these clinicians, what matters to them, and how do their goals align with intervention goals? Answering these questions will facilitate more tailored interventions clinicians want to participate in. Systemically categorising clinicians' preferences, knowledge, skills, and goals is the first step toward achieving a targeted understanding of what needs to be addressed to ensure the quality intervention is designed to drive practice changes. Other topics relevant for future research include the culture in which quality interventions take place, which is part of the contextual factors that can influence the effect enormously. The culture within a clinic strongly influences how feedback is being responded to, even if it is highly credible (i.e., trusting), from a trustworthy source, and constructive (i.e.,

usefulness and actionability), which will generally strengthen the effectiveness of quality interventions.(82) In addition, the culture significantly influences how clinicians work together, engagement and motivation to work on quality improvement.(68) Finally, a hospital culture with room to implement QII is needed in which clinicians feel supported even when the effect is disappointing. Truly exercising PDSA cycles to improve care also means abandoning interventions if there are no or negative effects.(83) Future research should focus on quality intervention trials incorporating up-to-date evidence- and theory-based best practices and address knowledge gaps. In addition, there should be a shift towards tailored audit and feedback studies rather than one-size-fits-all two-arm studies where multiple quality interventions are tested simultaneously.(84) Then the most effective intervention is implemented in clinical practice or taken as a starting point for future research.(1)

Implant choice

Rather than focusing on the quality of care delivered, another possibility to improve care is to focus on choosing the most optimal implant for a particular patient. **Chapter 3** showed that most negative outlier hospitals for overall revision were also a negative outlier for a specific indication for revision (e.g., infection, dislocation, or implant loosening), allowing hospitals to introduce targeted QII to improve the quality of care delivered, thereby lowering their revision rates. However, the specific implant was not included in the analysis, while the implant type can have significant effects on the overall revision rate as well as on specific indications for revision.(85-88) For example, metal-on-metal, large-head, uncemented, and resurfacing implants increased the risk of revision surgery after primary THA compared with metal-on-polyethylene, 32mm diameter heads, and cemented implants, respectively.(85,88-90) Also, introducing a new implant can be a reason for an increase in revisions since the instrumentation is slightly different as well as details on implant placement. The latter affects the surgical team and the surgeon. The Orthopaedic Data Evaluation Panel (ODEP) aims to guide surgeons' implant choices by classifying implants based on whether they have revision rates at an acceptable predefined level (i.e., an established external benchmark based on revision rates). Data are based on both observational single cohort studies as well as registry data with sufficient follow-up and sample size.(91) However, specific patients may benefit more from one type of implant while others may benefit more from a different type depending on patient and implant characteristics. Therefore, a prediction model to guide the most optimal implant choice for a specific patient (group) by estimating the lowest revision risk tailored to the individual patient (group) based on items like sex, age, femoral head size, comorbidity score, and activity level could, in theory, improve performance of arthroplasty surgery for patients.(92-94) Such a prediction model could represent the next step forward in improving the quality of care for patients after THA and TKA.

References

1. Ivers NM, Sales A, Colquhoun H, Michie S, Foy R, Francis JJ, et al. No more 'business as usual' with audit and feedback interventions: towards an agenda for a reinvigorated intervention. Implement Sci. 2014;9:14.
2. Brehaut JC, Eva KW. Building theories of knowledge translation interventions: use the entire menu of constructs. Implement Sci. 2012;7:114.
3. Website orf the Dutch Arthroplasty register (LROI). https://www.lroi.nl/. Checked on June 14, 2021.
4. Dutch Arthroplasty Register (LROI). Annual report 2019. www.lroi-report.nl.
5. Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, et al. Audit and feedback: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev. 2012(6):Cd000259.
6. van der Veer SN, de Keizer NF, Ravelli AC, Tenkink S, Jager KJ. Improving quality of care. A systematic review on how medical registries provide information feedback to health care providers. Int J Med Inform. 2010;79(5):305-23.
7. de Vos Maartje LG, van der Veer SN, Graafmans WC, de Keizer NF, Jager KJ, Westert GP, et al. Process evaluation of a tailored multifaceted feedback program to improve the quality of intensive care by using quality indicators. BMJ quality & safety. 2013;22(3):233-41.
8. Davis DA, Mazmanian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. JAMA. 2006;296(9):1094-102.
9. Soong C, Shojania KG. Education as a low-value improvement intervention: often necessary but rarely sufficient. BMJ quality & safety. 2020;29(5):353-7.
10. Gude WT, Roos-Blom MJ, van der Veer SN, de Jonge E, Peek N, Dongelmans DA, et al. Electronic audit and feedback intervention with action implementation toolbox to improve pain management in intensive care: protocol for a laboratory experiment and cluster randomised trial. Implement Sci. 2017;12(1):68.
11. Roos-Blom MJ, Gude WT, de Jonge E, Spijkstra JJ, van der Veer SN, Peek N, et al. Impact of audit and feedback with action implementation toolbox on improving ICU pain management: cluster-randomised controlled trial. BMJ quality & safety. 2019;28(12):1007-15.
12. Norwegian Arthroplasty Register (NAR). Annual Report 2020. http://nrlweb.ihelse.net/eng/Rapporter/Report2020_english.pdf.
13. Dutch Arthroplasty Register (LROI). Online Annual Report. https://www.lroi-report.nl/.
14. Swedish Knee Arthroplasty Register (SKAR). Annual Report 2020. http://myknee.se/pdf/SVK_2020_Eng_1.0.pdf.
15. Danish Hip Arthroplasty Register (DHAR). Annual report 2020. http://danskhoftealloplastikregister.dk/wp-content/uploads/2020/11/dhr-aarsrapport-2020_til_offentliggoerelse-1.pdf.
16. Swedish Hip Arthroplasty Register (SHAR). Annual Report 2018. https://registercentrum.blob.core.windows.net/shpr/r/Arsrapport_2018_Hoftprotes_ENG_26mars_Final-rJepCXNsLI.pdf.
17. Danish Knee Arthroplasty Register (DKAR). Annual Report 2020. https://www.sundhed.dk/content/cms/99/4699_dkr-arsrapport-2020_offentliggorelse.pdf.
18. Finnish Arthroplasty Register (FAR). Online Annual Report. https://www.thl.fi/far/#index.
19. Swiss Arthroplasty Register (SAR). Annual Report 2020. https://www.siris-implant.ch/de/Downloads&category=16.

20. Lenguerrand E, Whitehouse MR, Beswick AD, Kunutsor SK, Burston B, Porter M, et al. Risk factors associated with revision for prosthetic joint infection after hip replacement: a prospective observational cohort study. Lancet Infect Dis. 2018;18(9):1004-14.
21. Prokopetz JJ, Losina E, Bliss RL, Wright J, Baron JA, Katz JN. Risk factors for revision of primary total hip arthroplasty: a systematic review. BMC Musculoskelet Disord. 2012;13:251.
22. Jasper LL, Jones, C. A., Mollins, J., Pohar, S. L., Beaupre, L. A.. Risk factors for revision of total knee arthroplasty: a scoping review. BMC Musculoskelet Disord. 2016;17:182.
23. Signorini DF, Weir NU. Any variability in outcome comparisons adjusted for case mix must be accounted for. BMJ. 1999;318(7176):128.
24. Drummond J, Tran P, Fary C. Metal-on-Metal Hip Arthroplasty: A Review of Adverse Reactions and Patient Management. Journal of functional biomaterials. 2015;6(3):486-99.
25. Pfuntner A, Wier LM, Stocks C. Most Frequent Procedures Performed in U.S. Hospitals, 2011: Statistical Brief #165. Healthcare Cost and Utilization Project (HCUP) Statistical Briefs. Rockville (MD): Agency for Healthcare Research and Quality (US); 2006.
26. International Society of Arthroplasty Registries (ISAR). https://www.isarhome.org/home.
27. van Dishoeck AM, Koek MB, Steyerberg EW, van Benthem BH, Vos MC, Lingsma HF. Use of surgical-site infection rates to rank hospital performance across several types of surgery. Br J Surg. 2013;100(5):628-36; discussion 37.
28. van Dishoeck AM, Lingsma HF, Mackenbach JP, Steyerberg EW. Random variation and rankability of hospitals using outcome indicators. BMJ quality & safety. 2011;20(10):869-74.
29. Adab P, Rouse AM, Mohammed MA, Marshall T. Performance league tables: the NHS deserves better. BMJ. 2002;324(7329):95-8.
30. Cram P, Cai X, Lu X, Vaughan-Sarrazin MS, Miller BJ. Total knee arthroplasty outcomes in top-ranked and non-top-ranked orthopedic hospitals: an analysis of Medicare administrative data. Mayo Clin Proc. 2012;87(4):341-8.
31. Neuburger J, Walker, K., Sherlaw-Johnson, C., van der Meulen, J., Cromwell, D. A.. Comparison of control charts for monitoring clinical performance using binary data. BMJ quality & safety. 2017;26(11):919-28.
32. Landelijke Registratie Orthopedische Interventies (LROI); CUSUM chart implementation. https://www.lroi.nl/nieuws/dashboard-uitgebreid-survival-cusum-en-indicatoren.
33. Hofstede SN, Ceyisakar IE, Lingsma HF, Kringos DS, Marang-van de Mheen PJ. Ranking hospitals: do we gain reliability by using composite rather than individual indicators? BMJ quality & safety. 2019;28(2):94-102.
34. Karthaus EG, Lijftogt N, Busweiler LAD, Elsman BHP, Wouters M, Vahl AC, et al. Textbook Outcome: A Composite Measure for Quality of Elective Aneurysm Surgery. Ann Surg. 2017;266(5):898-904.
35. Dimick JB, Birkmeyer NJ, Finks JF, Share DA, English WJ, Carlin AM, et al. Composite measures for profiling hospitals on bariatric surgery performance. JAMA Surg. 2014;149(1):10-6.
36. Lingsma HF, Bottle A, Middleton S, Kievit J, Steyerberg EW, Marang-van de Mheen PJ. Evaluation of hospital outcomes: the relation between length-of-stay, readmission, and mortality in a large international administrative database. BMC Health Serv Res. 2018;18(1):116.
37. Van Schie P, Van Bodegom-Vos L, Van Steenbergen LN, Nelissen R, Marang-van de Mheen PJ. A more comprehensive evaluation of quality of care after total hip and knee arthroplasty: combining 4 indicators in an ordered composite outcome. Acta Orthop. 2022;93:138-45.
38. Rolfson O, Kärrholm J, Dahlberg LE, Garellick G. Patient-reported outcomes in the Swedish Hip Arthroplasty Register: results of a nationwide prospective observational study. J Bone Joint Surg Br. 2011;93(7):867-75.

39. Dunbar MJ, Richardson G, Robertsson O. I can't get no satisfaction after my total knee replacement: rhymes and reasons. Bone Joint J. 2013;95-b(11 Suppl A):148-52.
40. Johnston B, Patrick D, Devji T, al e. Chapter 18: Patient_reported outcomes. Cochrane Handbook for Systematic Reviews of Interventions. 2019.
41. Makhni EC. Meaningful Clinical Applications of Patient-Reported Outcome Measures in Orthopaedics. J Bone Joint Surg Am. 2021;103(1):84-91.
42. FDA. Guidance for Indusctry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. 2009.
43. Black N. Patient reported outcome measures could help transform healthcare. BMJ. 2013;346:f167.
44. Rolfson O, Eresian Chenok K, Bohm E, Lübbeke A, Denissen G, Dunn J, et al. Patient-reported outcome measures in arthroplasty registries. Acta Orthop. 2016;87 Suppl 1(Suppl 1):3-8.
45. Bohm ER, Kirby S, Trepman E, Hallstrom BR, Rolfson O, Wilkinson JM, et al. Collection and Reporting of Patient-reported Outcome Measures in Arthroplasty Registries: Multinational Survey and Recommendations. Clin Orthop Relat Res. 2021;479(10):2151-66.
46. Hutchings A, Grosse Frie K, Neuburger J, van der Meulen J, Black N. Late response to patient-reported outcome questionnaires after surgery was associated with worse outcome. J Clin Epidemiol. 2013;66(2):218-25.
47. Jahagirdar D, Kroll T, Ritchie K, Wyke S. Patient-reported outcome measures for chronic obstructive pulmonary disease : the exclusion of people with low literacy skills and learning disabilities. Patient. 2013;6(1):11-21.
48. Gibbons E. Patient-reported outcome measures and the evaluation of services. Health Serv Deliv Res. 2016;4:55-68.
49. Tuti T, Nzinga J, Njoroge M, Brown B, Peek N, English M, et al. A systematic review of electronic audit and feedback: intervention effectiveness and use of behaviour change theory. Implement Sci. 2017;12(1):61.
50. Weiss D, Dunn SI, Sprague AE, Fell DB, Grimshaw JM, Darling E, et al. Effect of a population-level performance dashboard intervention on maternal-newborn outcomes: an interrupted time series study. BMJ quality & safety. 2018;27(6):425-36.
51. Shojania KG, Ranji SR, McDonald KM, Grimshaw JM, Sundaram V, Rushakoff RJ, et al. Effects of quality improvement strategies for type 2 diabetes on glycemic control: a meta-regression analysis. JAMA. 2006;296(4):427-40.
52. Tricco AC, Ivers NM, Grimshaw JM, Moher D, Turner L, Galipeau J, et al. Effectiveness of quality improvement strategies on the management of diabetes: a systematic review and meta-analysis. Lancet. 2012;379(9833):2252-61.
53. DR I, CD F, MS T. Consequences of individual feedback on behaviour in organisations. J Appl Psychol. 1979:349-71.
54. Hysong SJ, Best RG, Pugh JA. Audit and feedback and clinical practice guideline adherence: making feedback actionable. Implement Sci. 2006;1:9.
55. Kluger AN, Van Dijk D. Feedback, the various tasks of the doctor, and the feedforward alternative. Med Educ. 2010;44(12):1166-74.
56. Gude WT, Roos-Blom MJ, van der Veer SN, Dongelmans DA, de Jonge E, Francis JJ, et al. Health professionals' perceptions about their clinical performance and the influence of audit and feedback on their intentions to improve practice: a theory-based study in Dutch intensive care units. Implement Sci. 2018;13(1):33.

57. Landis-Lewis Z, Douglas GP, Hochheiser H, Kam M, Gadabu O, Bwanali M, et al. Computer-Supported Feedback Message Tailoring for Healthcare Providers in Malawi: Proof-of-Concept. AMIA Annu Symp Proc. 2015;2015:814-23.
58. AN K, A D. The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory. Psychol Bull. 1996(119):254-84.
59. RE J, M H, C-H C. The importance of velocity, or why speed may matter more than distance. Organ Psychol Rev. 2013(3):62-85.
60. Gude WT, Brown B, van der Veer SN, Colquhoun HL, Ivers NM, Brehaut JC, et al. Clinical performance comparators in audit and feedback: a review of theory and evidence. Implement Sci. 2019;14(1):39.
61. Gude WT, van Engen-Verheul MM, van der Veer SN, de Keizer NF, Peek N. How does audit and feedback influence intentions of health professionals to improve practice? A laboratory experiment and field study in cardiac rehabilitation. BMJ quality & safety. 2017;26(4):279-87.
62. Locke EA, Latham GP. Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. Am Psychol. 2002;57(9):705-17.
63. Carver CS, Scheier MF. Control theory: a useful conceptual framework for personality-social, clinical, and health psychology. Psychol Bull. 1982;92(1):111-35.
64. GT D. There's a S.M.A.R.T. way to write management's goals and objectives. Manage Rev. 1981:35-6.
65. de Vos ML, van der Veer SN, Wouterse B, Graafmans WC, Peek N, de Keizer NF, et al. A multifaceted feedback strategy alone does not improve the adherence to organizational guideline-based standards: a cluster randomized trial in intensive care. Implement Sci. 2015;10:95.
66. Ivers NM, Grimshaw JM, Jamtvedt G, Flottorp S, O'Brien MA, French SD, et al. Growing literature, stagnant science? Systematic review, meta-regression and cumulative analysis of audit and feedback interventions in health care. J Gen Intern Med. 2014;29(11):1534-41.
67. ClinicalTrial.gov. https://clinicaltrials.gov/ct2/show/NCT04055103?term=Arthroplasty&cntry=NL&city=Leiden&draw=2&rank=2.
68. Wolfstadt JI, Cohen-Rosenblum A. 'You can't do quality between surgical cases and tea time': barriers to surgeon engagement in quality improvement. BMJ quality & safety. 2022:bmjqs-2022-015083.
69. Glasgow JM, Davies ML, Kaboli PJ. Findings from a national improvement collaborative: are improvements sustained? BMJ quality & safety. 2012;21(8):663-9.
70. Bridges J, May C, Fuller A, Griffiths P, Wigley W, Gould L, et al. Optimising impact and sustainability: a qualitative process evaluation of a complex intervention targeted at compassionate care. BMJ quality & safety. 2017;26(12):970-7.
71. Lipitz-Snyderman A, Needham DM, Colantuoni E, Goeschel CA, Marsteller JA, Thompson DA, et al. The ability of intensive care units to maintain zero central line-associated bloodstream infections. Arch Intern Med. 2011;171(9):856-8.
72. Pronovost PJ, Watson SR, Goeschel CA, Hyzy RC, Berenholtz SM. Sustaining Reductions in Central Line-Associated Bloodstream Infections in Michigan Intensive Care Units: A 10-Year Analysis. Am J Med Qual. 2016;31(3):197-202.
73. Burke RE, Marang-van de Mheen PJ. Sustaining quality improvement efforts: emerging principles and practice. BMJ quality & safety. 2021.
74. Greene SM, Reid RJ, Larson EB. Implementing the learning health system: from concept to action. Ann Intern Med. 2012;157(3):207-10.
75. Kilbourne AM, Glasgow RE, Chambers DA. What Can Implementation Science Do for You? Key Success Stories from the Field. J Gen Intern Med. 2020;35(Suppl 2):783-7.

76. Jackson GL, Cutrona SL, White BS, Reardon CM, Orvek E, Nevedal AL, et al. Merging Implementation Practice and Science to Scale Up Promising Practices: The Veterans Health Administration (VHA) Diffusion of Excellence (DoE) Program. Jt Comm J Qual Patient Saf. 2021;47(4):217-27.
77. Trbovich P, Shojania KG. Root-cause analysis: swatting at mosquitoes versus draining the swamp. BMJ quality & safety. 2017;26(5):350-3.
78. van Schie P, van Bodegom-Vos L, van Steenbergen LN, Nelissen R, Marang-van de Mheen PJ. Monitoring Hospital Performance with Statistical Process Control After Total Hip and Knee Arthroplasty: A Study to Determine How Much Earlier Worsening Performance Can Be Detected. J Bone Joint Surg Am. 2020;102(23):2087-94.
79. Sales A. Reporting on implementation trials with null findings: the need for concurrent process evaluation reporting. BMJ quality & safety. 2022.
80. Eccles MP, Foy R, Sales A, Wensing M, Mittman B. Implementation Science six years on--our evolving scope and common reasons for rejection without review. Implement Sci. 2012;7:71.
81. Desveaux L, Ivers NM, Devotta K, Ramji N, Weyman K, Kiran T. Unpacking the intention to action gap: a qualitative study understanding how physicians engage with audit and feedback. Implement Sci. 2021;16(1):19.
82. Watling C, Driessen E, van der Vleuten CP, Vanstone M, Lingard L. Beyond individualism: professional culture and its influence on feedback. Med Educ. 2013;47(6):585-94.
83. Leis JA, Shojania KG. A primer on PDSA: executing plan-do-study-act cycles in practice, not just in name. BMJ quality & safety. 2017;26(7):572-7.
84. Boggan J. Tailoring audit and feedback - a precision intervention but with added complexity? BMJ quality & safety. 2022.
85. López-López JA, Humphriss RL, Beswick AD, Thom HHZ, Hunt LP, Burston A, et al. Choice of implant combinations in total hip replacement: systematic review and network meta-analysis. BMJ. 2017;359:j4651.
86. Deere KC, Whitehouse MR, Porter M, Blom AW, Sayers A. Assessing the non-inferiority of prosthesis constructs used in hip replacement using data from the National Joint Registry of England, Wales, Northern Ireland and the Isle of Man: a benchmarking study. BMJ Open. 2019;9(4):e026685.
87. Wyatt M, Frampton C, Whitehouse M, Deere K, Sayers A, Kieser D. Benchmarking total hip replacement constructs using noninferiority analysis: the New Zealand joint registry study. BMC Musculoskelet Disord. 2021;22(1):719.
88. Smith AJ, Dieppe P, Vernon K, Porter M, Blom AW. Failure rates of stemmed metal-on-metal hip replacements: analysis of data from the National Joint Registry of England and Wales. Lancet. 2012;379(9822):1199-204.
89. Clarke A, Pulikottil-Jacob R, Grove A, Freeman K, Mistry H, Tsertsvadze A, et al. Total hip replacement and surface replacement for the treatment of pain and disability resulting from end-stage arthritis of the hip (review of technology appraisal guidance 2 and 44): systematic review and economic evaluation. Health Technol Assess. 2015;19(10):1-668, vii-viii.
90. Smith AJ, Dieppe P, Howard PW, Blom AW. Failure rates of metal-on-metal hip resurfacings: analysis of data from the National Joint Registry for England and Wales. Lancet. 2012;380(9855):1759-66.
91. Orthopaedic Data Evaluation Panel (ODEP). ODEP Benchmarks. Available at: http://www.odep.org.uk/Benchmarks/ODEPBenchmarks.aspx.
92. Tsikandylakis G, Mohaddes M, Cnudde P, Eskelinen A, Karrholm J, Rolfson O. Head size in primary total hip arthroplasty. EFORT Open Rev. 2018;3(5):225-31.

93. Paxton E, Cafri G, Havelin L, Stea S, Pallisó F, Graves S, et al. Risk of revision following total hip arthroplasty: metal-on-conventional polyethylene compared with metal-on-highly cross-linked polyethylene bearing surfaces: international results from six registries. J Bone Joint Surg Am. 2014;96 Suppl 1(Suppl 1):19-24.
94. Berry DJ, Harmsen WS, Cabanela ME, Morrey BF. Twenty-five-year survivorship of two thousand consecutive primary Charnley total hip replacements: factors affecting survivorship of acetabular and femoral components. J Bone Joint Surg Am. 2002;84(2):171-7.