Streamlining surgical instrument counting: a matrixed multiple case study on the fidelity of weighing systems in the operating room

Kooijmans, Anton M.; van der Elst, Maarten; van den Dobbelsteen, John J.

DOI
10.1186/s43058-025-00808-8

Publication date
2025

Document Version
Final published version

Published in
Implementation Science Communications

Citation (APA)
Kooijmans, A. M., van der Elst, M., & van den Dobbelsteen, J. J. (2025). Streamlining surgical instrument counting: a matrixed multiple case study on the fidelity of weighing systems in the operating room. *Implementation Science Communications*, *6*(1), Article 118. https://doi.org/10.1186/s43058-025-00808-8

Kooijmans *et al.*
Implementation Science Communications (2025) 6:118
https://doi.org/10.1186/s43058-025-00808-8

Implementation Science Communications

RESEARCH Open Access

Streamlining surgical instrument counting: a matrixed multiple case study on the fidelity of weighing systems in the operating room

Anton M. Kooijmans[1,2*], Maarten van der Elst[1,2] and John J. van den Dobbelsteen[1]

Abstract

Background Many technologies have been developed to aid in surgical instrument counting, but wide adoption is rare. A technology that has been widely adopted around 20 years ago is the weighing scale. Lessons can be extracted from its sustainment and fidelity, and applied to the development and implementation of new labour-saving technologies in healthcare.

Methods We conducted semi-structured interviews with experienced staff in four hospitals that use weighing systems in their surgical instrument cycle, which we analysed according to the Matrixed Multiple Case Study (MMCS) methodology. Hospitals were designated a low, medium, or high sustainment and fidelity score, after which influencing factors were identified. These factors were categorised according to the i-PARIHS domains of Innovation, Recipient, Context, and Facilitation. Within-site analysis and cross-site analysis was performed to identify influencing factors associated with a high or low level of sustainment or fidelity.

Results All hospitals showed a high sustainment. Two hospitals showed low fidelity, and two showed high fidelity. Twenty-one total influencing factors were identified, divided among all i-PARIHS domains. All hospitals experienced similar limitations of the technology, and all hospitals showed signs of facilitation efforts during the implementation phase. In low-fidelity hospitals, interdepartmental coordination and trust in technology were limited, in contrast to high-fidelity hospitals. A large and/or complex surgical instrument inventory hindered fidelity of the weighing system.

Conclusions 20 years after implementation, there is varying success concerning the fidelity of weighing systems for surgical instrument counting. All participating hospitals have adapted their workflow to the limitations of the technology in different ways. Given the relative straight-forwardness of weighing scales as a technology, our findings underline the complexity of implementation processes, regardless of the complexity of the innovation.

Keywords Surgical instrument counting, Labour-saving technology, Weighing scale fidelity

Contributions to the literature

- Implementing labour-saving technologies in healthcare is difficult. Lessons can be learnt from already existing technologies, such as the use of weighing systems for surgical instrument counting.
- Even straightforward technologies, such as weighing scales, require unique adaptations to local workflows. Some hospitals could benefit from a designated facili-

*Correspondence:
Anton M. Kooijmans
a.m.kooijmans@tudelft.nl
[1] Delft University of Technology, Mekelweg 2, Delft 2628 CD, the Netherlands
[2] Reinier de Graaf Hospital, Delft, the Netherlands

tator who ensures interdepartmental coordination for optimal use of (more complex) technologies.
- If a technology is used across multiple departments in one organisation, a lack of interdepartmental coordination results in a lack of trust in the technology by the users.

Background

Surgical material counting is a labour-intensive task of operating room (OR) nurses. For many years already, attempts are made to develop a technology that is able to support or fully automate this job [1–3]. However, these technologies rarely get implemented on a wide scale in hospital settings. Many of these developments don't leave the feasibility stage or are implemented in one hospital, after which dissemination is halted [4–6]. Understanding the factors that influence implementation and integration into clinical practice is essential. Engineers, clinicians, and administrators looking to implement a new technology in their hospital organisation could benefit from the lessons learnt from already existing technologies.

Around 20 years ago, in the Netherlands, weighing scales were introduced as a solution to aid in instrument counting, aiming to limit human error and increase patient safety [7]. A recent survey among all Dutch hospitals shows that around a quarter of them have implemented a weighing system in their instrument lifecycle (Kooijmans et al.: Surgical materials counting in Dutch hospitals: a comparison of protocols and procedures over time, submitted). Weighing scales are a straightforward technology, but, according to the survey they are not always used in the way they were originally intended; to count surgical instruments. This raises questions about the innovation's implementation success and long-term use.

Implementation science is a growing field in healthcare. While medical interventions or treatments are mostly focused on evidence-based effectiveness, they will not be effective if they are not implemented well. As such, implementation outcomes serve as necessary preconditions for making desired changes in clinical or health service outcomes [8]. A multitude of such outcomes have been described in the literature. Proctor et al.'s taxonomy of implementation outcomes provides a clear overview of concepts: acceptability, adoption, appropriateness, cost, feasibility, fidelity, penetration, sustainability [8]. Considering that implementation of the weighing system took place around twenty years ago, early-stage implementation outcomes are not relevant for this study. Penetration is not relevant as all investigated hospitals use the innovation. Relevant outcomes for this late stage of implementation are fidelity and sustainment. Both terms will be further explained below. See Fig. 1 for a schematic overview of these terms and the subsequent analysis.

The fidelity of an implemented practice describes the level to which the practice is used as it was intended during implementation [8]. Many definitions of fidelity exist in the literature, but the framework proposed by Carrol et al. will be used in this study [9]. Carroll divides fidelity into several subcategories: content (the 'active ingredients' of the intervention; or the knowledge the intervention seeks to deliver to its recipients); coverage (whether the intervention is used for all procedures it was initially intended for); frequency (whether the intervention is used as often as intended); and duration (whether the intervention is used as long as intended). Sustainability or sustainment is an implementation outcome that refers to the level of sustained use of the innovation, and the integration of the innovation into regular workflows [8]. Rabin et al. describes three stages of sustainment [10]: 1) Passage (a single event such as transition from temporary to permanent funding); 2) Cycle or routine (i.e. repetitive reinforcement of the importance of the intervention; organisational procedures and behaviours); 3) Niche saturation (the extent to which an intervention is integrated into all subsystems of an organisation). Passage indicates a low sustainment, cycle or routine indicates a medium sustainment, and niche saturation indicates a high sustainment.

In general, all hospitals use a similar instrument life cycle [11]: Starting at the end of a surgical procedure, instruments are brought to a processing room, from which larger batches of instrument trays are transported to the sterile processing department (SPD). There, instruments are first grossly cleaned to remove blood, tissue, and other material, after which they are manually scrubbed to further decontaminate them. Then, instruments are placed in large automatic washers, where they are further cleaned according to a strictly audited regimen. Cleaned instruments are then taken out of the washers into the 'clean zone', where they are assembled into their respective instrument trays. Fully assembled trays are then wrapped in a polypropylene wrap and subsequently sterilised. Sterile instrument trays are kept in storage until they are needed again for a surgical procedure. Generally, instruments are weighed at the end of surgery and during assembly (after cleaning). The latter is then used as reference for the next postoperative weighing procedure. Weighing stations can also be added before start of surgery and when instrument trays arrive at the SPD (before cleaning).

The instrument lifecycle can be complicated, as large hospitals can manage over 100,000 instrument trays and 2.6 million instruments annually. Considering weighing

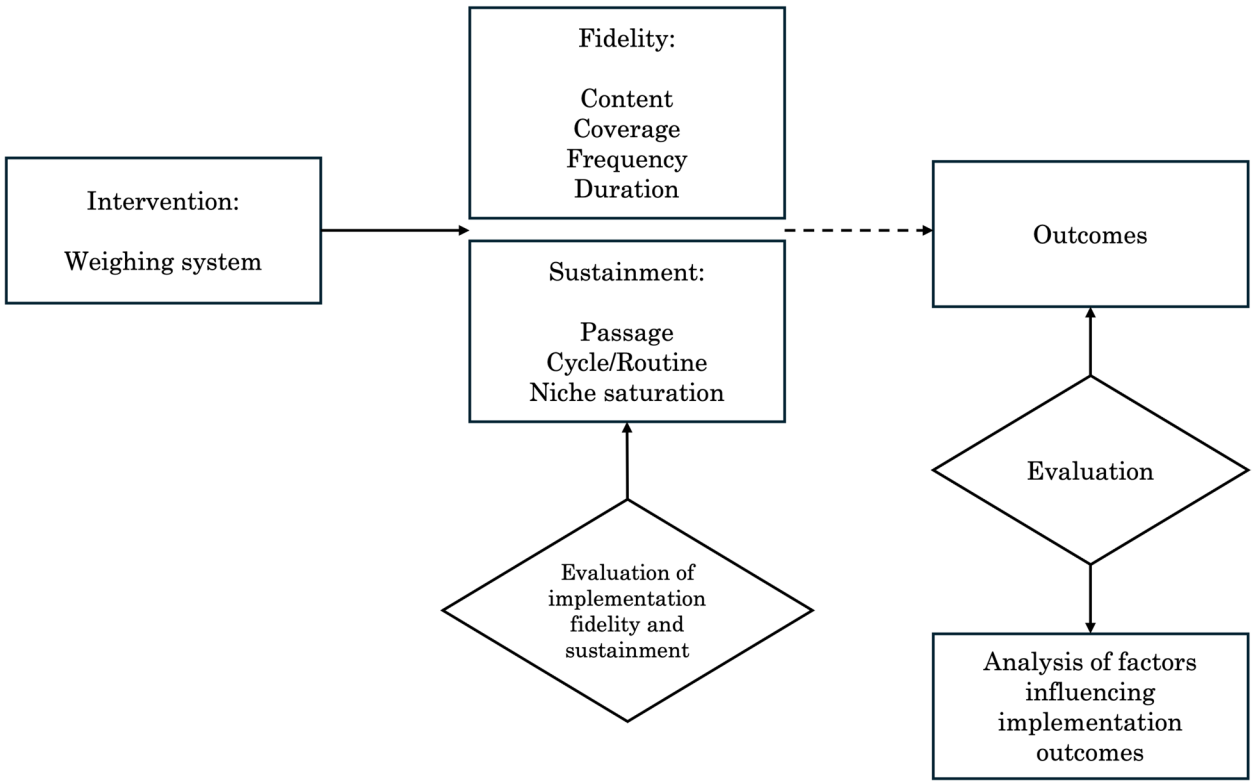

Fig. 1 Schematic overview of implementation fidelity and sustainment and the subsequent analysis of factors influencing these outcomes

specifically, it is estimated that 5.9% of trays contain broken instruments, potentially impacting their weight [12]. Furthermore, maintenance and repairs can alter an instrument's properties, exchanging its handle, or sharpening its cutting blade, for example [13]. Similarly, instruments can remain in use for many years before being discarded, given that the estimated lifespan of an instrument is 300–900 sterilisation cycles [12]. This means that several iterations of the same instrument type can be in circulation at the same time [14]. Lastly, variation of a few grams in a single instrument is amplified in a tray that could contain a hundred instruments.

In this qualitative study we investigated several hospitals that use the weighing system in their surgical instrument lifecycle, determining factors that influence the relevant implementation outcomes of fidelity and sustainment. To structure the analysis and to compare multiple hospital organisations, we used the matrixed multiple case study (MMCS) methodology, as described by Kim et al. [15]. This methodology encompasses nine steps of data collection, processing, and analysis, making use of research team consensus meetings to limit subjectivity in the study's findings.

Methods

Data collection

This qualitative study was performed according to the COREQ checklist for reporting of qualitative studies [16]. We conducted semi-structured interviews with OR and SPD senior staff members with at least five years of experience in the current department, in hospitals that use weighing scales in their surgical instrument lifecycle. Informed consent for participation in the study was obtained from all participants. The study was approved by the TU Delft Human Research Ethics Committee. Interview questions can be seen in Appendix A. If possible, we recorded audio of the interviews, as well as meeting notes when necessary.

Data analysis

We used the matrixed multiple case study (MMCS) methodology's nine analytical steps to analyse the interview data [15]. Consensus was reached on all steps by the research project's three team members (the authors), as well as four additional research group members (see Acknowledgments). We determined the evaluation goal (step 1) to be to identify the way in which fidelity and sustainment differed across the hospitals and which factors are of influence on this. Next, we defined fidelity

and sustainment (step 2) as respectively: how the current usage of the weighing system compares to the intended use at the time of implementation (i.e. instrument counting as a replacement for manual counting), and how the weighing system is integrated in the instrument lifecycle and related processes.

We selected i-PARIHS' domains [17]- innovation, recipients, context, and facilitation – as relevant domains to categorise the identified influencing factors (step 3). We examined the data on potential factors influencing fidelity and sustainment, as well as categorising these factors according to the abovementioned i-PARIHS domains (step 4) by conducting a hybrid inductive-deductive content analysis approach [18] over multiple coding rounds using ATLAS.ti (version 24.0) software.

We assessed fidelity and sustainment per site (step 5) by examining each site's code summary and exemplar quotes. To assess fidelity, we divided this concept into subcategories as described by Carroll et al. [9]: Content, Coverage, and Frequency. While a fourth subcategory is described (Duration), the researchers agreed that this category is irrelevant for this specific study, as the intervention (weighing) is intended to be used ongoingly at the participating sites. We assigned a higher fidelity level to sites that exhibited more characteristics of using the weighing scale for instrument counting, and a lower level if the weighing system was implemented in the instrument workflow, but did not contribute to the staff's knowledge whether all instruments are complete after surgery. Code summaries, based on the aforementioned subcategories, including exemplar quotes for each site were drafted by the main researcher (AK), after which they were reviewed and refined by all project team members. Review and discussion of the summaries of all four sites took approximately four 45-min meetings after which a final discussion took place to synthesise the subcategory scores. Consensus was reached and each site was designated a final fidelity score.

To assess the level of sustainment, we used Rabin et al.'s three stages of sustainment as a guide. A higher sustainment was assigned to sites that showed a deeper integration of the weighing system into standard workflows, and where the weighing process is considered "just the way things are done". As in determining the fidelity scores, reviews and discussions of the four sites' code summaries were held, taking around 45 min each. Consensus was reached and each site was designated a sustainment score.

We then identified influencing factors per site (step 6), categorising them under the i-PARIHS domains. Potential influencing factor descriptions were drafted based on the above-mentioned code summaries. The summaries, factor descriptions, and exemplar quotes were then reviewed by the whole project team and similar factor descriptions were consolidated into one factor.

We organised the data, being the assessed fidelity and sustainment and identified factors per site into sortable matrices (step 7) using Microsoft Excel. Influencing factors were laid out on the horizontal axis, and fidelity and sustainment on vertical axes. We conducted a within-site analysis (step 8), determining whether a factor was 'present', 'somewhat present', 'minimally present', or 'not mentioned', based on the number of times the concept surrounding a specific factor was discussed in aggregate reports from the hospitals' participants. To illustrate, one participant from a certain hospital mentioning a concept two times was counted the same as two participants from the same hospital mentioning the concept once. Aggregate reports were used in the consensus meetings to limit complexity of the discussions and to allow for efficient analysis and decision-making. Then, each factor was determined to be of enabling, hindering, neutral, or unclear influence on implementation. Neutral factors were those we found to have both an enabling and hindering influence. Like earlier in this study, the frequency and type of the identified factors was initially proposed by the main researcher, after which consensus was achieved through discussion sessions with the whole project team.

Finally, we conducted cross-site analysis (step 9), assessing whether certain factors or combinations thereof were present across multiple sites and whether these factors were consistently present in higher or lower levels of fidelity and sustainment. As not every factor was present at all sites or did not come up in the conversation, the interview was finalised with an open question on whether the interviewee had anything to add that was not discussed yet.

As a team, we decided that factors were associated with high/medium/low fidelity and/or sustainment if they were at least somewhat present in all hospitals designated as high/medium/low fidelity and/or sustainment.

Results

8 interviews across 4 hospitals (two academic and two regional centres) were conducted with key OR and SPD staff members. Both academic centres are equipped with more than thirty operating rooms, comprising regular, paediatric, and hybrid ORs, and perform mostly highly complex surgeries. One regional hospital has twenty operating rooms, both regular and hybrid, and perform both regular and highly complex procedures. The other regional hospital is equipped with five ORs, and performs routine surgical procedures as well as paediatric cases. All hospitals implemented the weighing system around the same period, approximately 20 years ago, to aid in

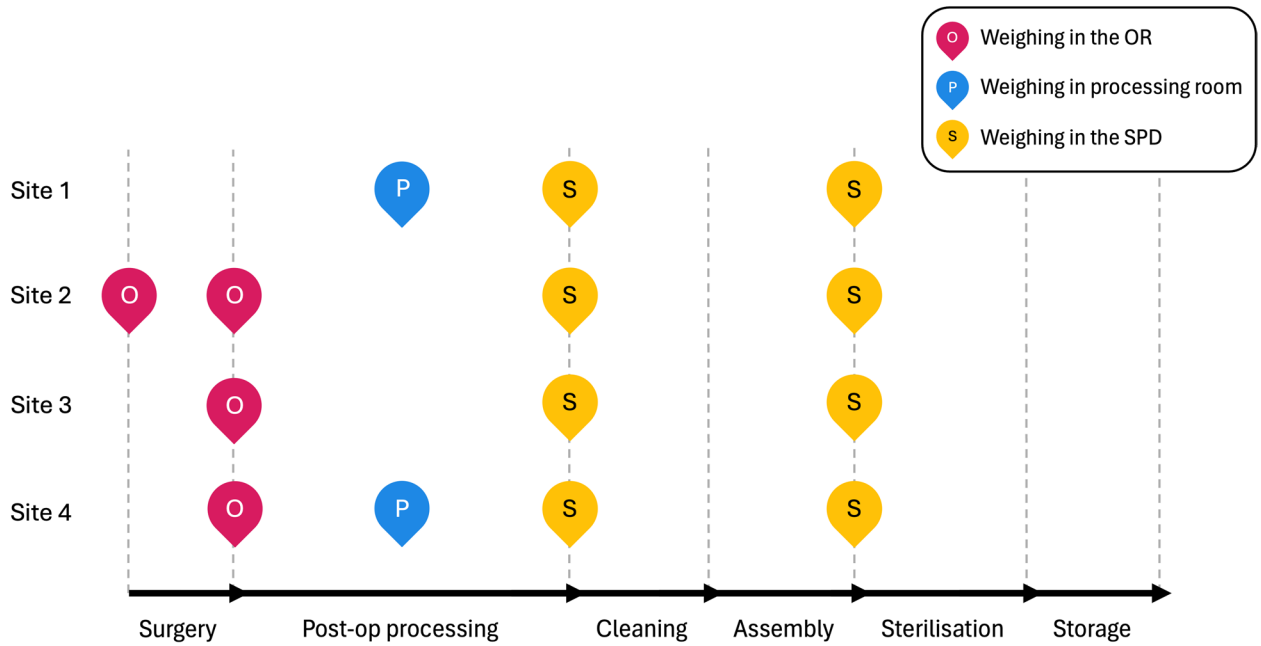

Fig. 2 Weighing workflows across studied hospitals (Site 1–4). Each marker denotes a weighing procedure. The location of the weighing system is denoted by the colour of the marker

instrument counting. Across all hospitals, weighing stations are set up in the OR and SPD, as well as the post-operative processing room in some cases. See Fig. 2 for a schematic overview of workflows.

Participants

All participants were part of senior staff, averaging 20 (7–36) years of experience in the current hospital department, being either the OR or SPD.

Fidelity and sustainment score

The level of fidelity was split evenly across the participating hospitals, with two sites having a high score, and two sites having a low score. As all hospitals showed similar scores for the Coverage and Frequency subcategories, the research project team agreed to differentiate the different hospitals based on their Content scores. Sustainment was equally high among all hospitals, with all showing features of niche saturation of the intervention; the third stage of sustainment according to Rabin et al. [10]. For this reason, factors influencing sustainment cannot be analysed in this study. However, sustainment-related findings in some cases did have an influence on the weighing system's fidelity, see the 'Adaptation of the innovation' section in the Discussion below.

Twenty-one total influencing factors were identified, categorised according to the i-PARIHS domains of Innovation (7), Recipients (4), Context (5), and Facilitation (5). Figure 3 shows these identified influencing factors by domain.

Key trends in influencing factors regardless of weighing system fidelity status

Three factors in the i-PARIHS domain of Innovation, one factor in the domain of Context, and three factors in the domain of Facilitation were relatively or strongly present in all hospitals with similar influence, regardless of weighing system fidelity status. They were the *Variation in size and weight of instruments and materials*, *Slim weight margins*, *Variation in tray wrapping paper*, *Integration into IT subsystems*, *Removal of additional materials from tray before weighing*, *Determining own weight margin*, and *Addition of weighing stations*. These factors will be discussed below.

Innovation-related factors All hospitals experience difficulties related to the weighing system itself. Especially the *Variation in size and weight of instruments and materials* is a prominent factor with a hindering influence: Participants mentioned that instruments of the same type can have slightly different shapes, sizes, and thus weights. Instrument tray wrapping paper also presents

i-PARIHS domain	Influencing factor	Site 1 (High fidelity)	Site 2 (High fidelity)	Site 3 (Low fidelity)	Site 4 (Low fidelity)
Innovation	Variation in size and weight of instruments and materials	+	++	++	++
	Soiling and/or moisture (on instruments)	+	N/m	++	N/m
	Only unsterile instruments can be weighed	-	N/m	++	++
	Slim weight margins	++	++	+	++
	Limited flexibility in case of unexpected events during procedure	N/m	N/m	N/m	++
	Variation in tray wrapping paper	N/m	++	N/m	++
	Room for error in weighing process	N/m	+	+	++
Recipients	Experience / Knowledge of instrument tray contents	++	++	++	++
	Attitude towards technology / Trust in technology	++	++	+	-
	Interdisciplinary coordination / Trust in other people's work	++	++	-	-
	Teamwork / Collaboration	++	+	N/m	N/m
Context	Integration into IT subsystems	++	++	++	++
	Large and complex instrument inventory	-	++	++	++
	Large hospital / complex surgical procedures	-	+	++	++
	Continually updated tray reference weight	N/m	+	-	++
	Being able to weigh in the operating room	N/m	N/m	++	++
Facilitation	Removal of additional materials from tray before weighing	++	++	++	++
	Determining own weight margin	++	++	++	++
	Only certain surgical procedures require weighing	N/m	++	N/m	N/m
	Limited consideration for weighing process during renovations	N/m	N/m	N/m	++
	Addition of weighing stations	N/m	++	++	N/m

++ Present
\+ Somewhat present
\- Minimally present
N/m Not mentioned

Enabling
Hindering
Neutral
Unclear/Neutral

Fig. 3 Influencing factors and their presence and type of influence on the fidelity of weighing systems use in surgical instrument counting

issues in some hospitals, as a participant mentioned that the amount of paper used depends on the preference of staff members. As a result, in hospitals with high fidelity, participants noted the need for human interpretation of weighing results, while participants in low fidelity hospitals place more emphasis on the variation of instruments as a flaw.

> *"One large-basic tray is not the other large-basic tray. One hook is narrower than the other, or it has a different handle." (Participant 2-1, high fidelity)*

> *"We use these blue paper sheets in the trays. These days, they make some of these trays way too big, or too high, they need more than one sheet. Some people use two sheets, but some people use one and a half sheet. That's a difference of 5 grams." (Participant 2-2, high fidelity)*

Slim weight margins were mentioned in all hospitals with a neutral to hindering influence on weighing system fidelity. While these margins are easily exceeded because of soiled or wet instruments, they cannot be increased, as some instruments only weigh a few grams.

Context-related factor Integration into IT subsystems is strongly present in all hospitals with an enabling influence. In every hospital, the weighing system is connected to the electronic patient file software. When a tray is weighed, it is automatically registered to its corresponding patient file. This traceability is an important component of the weighing system's sustained use:

> *"Yes, you always have to weigh, otherwise it [the tray] cannot advance in the system." (Participant 1-1, high fidelity)*

> *"For the SPD, the weighing system is a very important track & trace system. By weighing at specific moments in the process, we can accurately pinpoint where an instrument could have gone missing." (Participant 3-2, low fidelity)*

Facilitation-related factors In all hospitals, *Removal of additional materials from tray before weighing* is strongly present with a neutral or unclear influence. Excluding these items from weighing necessitates a manual count. Also strongly present in all hospitals is *Determining own weight margin*, with an enabling influence.

> *"Cups, kidney dishes, cables, and lids aren't weighed. There's too much variation in size and weight." (Participant 3-2, low fidelity)*

"There's certain items that we don't weigh… kidney dishes, electrosurgery cables… Those are counted like usual." (Participant 1-1, high fidelity)

Hospitals in both the low and high fidelity group have added weighing stations somewhere in their workflow, with an unclear influence on fidelity.

"The reason we weigh twice is because in the past, we found that certain trays' weights were often incorrect." (Participant 2-1, high fidelity)

Key trends in influencing factors in low versus high fidelity hospitals

One factor in the Innovation domain, three factors in the Recipients domain, and three factors in the Context domain showed trends possibly related to fidelity status, namely: *Room for error in weighing process*, *Experience/Knowledge of instrument tray contents*, *Attitude towards technology/Trust in technology*, *Interdisciplinary coordination/Trust in other people's work*, *Large and complex instrument inventory*, *Large hospital/complex surgical procedures*, and *Continually updated tray reference weight*.

Innovation-related factor The factor *Room for error in weighing process* is somewhat present in low fidelity hospitals and minimally present in hospitals with a high fidelity, but with a hindering influence in all cases.

"In the SPD, they use a height-adjustable table on which the scale is placed. If this table is adjusted, the calibration of the scale is off. … Such an incorrect reference weight occurs multiple times a day and is very annoying." (Participant 4-3, low fidelity)

Recipients-related factors An important factor among OR nurses is their *Experience/Knowledge of instrument tray contents*, being strongly present in all hospitals. In hospitals with a low fidelity, this factor has a hindering influence, but in hospitals with a high fidelity, it has a neutral to enabling influence.

Participants from low fidelity hospitals found that the weighing of trays did not contribute to their knowledge of whether all instruments are complete. While participants in high fidelity hospitals share this sentiment, they acknowledge the added value of weighing.

"I know the contents of 200 different trays, so I already know if it's complete before weighing." (Participant 4-3, low fidelity)

"It is not for nothing, because you can really miss something. It's an extra check. Usually you get confirmation, and sometimes rejection, then you have to find out what's going on. Yes, that's a good thing." (Participant 2-1, high fidelity)

In low fidelity hospitals, the *Attitude towards technology/Trust in technology* is minimally to somewhat present with a hindering influence, while in hospitals with a high weighing system fidelity, it is strongly present with an enabling influence. Participants in low fidelity hospitals tended to see the limitations of technology in comparison to manual counting, while participants in high fidelity hospitals tended to see technology as a tool that supports them. However, interestingly, in both low fidelity hospitals, participants from the SPD considered the weighing system a valuable addition to their workflow.

"There will never be a technology that will replace manual counting in the OR. It will always remain a manual process." (Participant 4-3, low fidelity)

"Yes, I count way less now. I hardly count at all anymore. Because I have seen… I have trust in the system." (Participant 1-1, high fidelity)

"For the SPD the weighing system is an important track and trace tool. By weighing at certain points in the cycle we can pinpoint where an instrument might have gone missing." (Participant 3-2, low fidelity)

Another factor that differentiates low and high fidelity hospitals, is the level of *Interdisciplinary coordination/Trust in other people's work*. This factor is strongly present with an enabling influence in high fidelity hospitals, while it is minimally present with a neutral influence in low fidelity hospitals. In one high fidelity hospital, a participant mentioned the quick communication that was very standard for that hospital. In low fidelity hospitals, this coordination was less visible. Instead, the technology seemed to be used to alleviate a lack of interdisciplinary coordination:

"It has solved a lot of discussions." (Participant 3-2, low fidelity)

"It doesn't matter how a tray comes in… When it comes in, we weigh it, and that is our reference." (Participant 4-2, low fidelity)

Context-related factors *Large and complex instrument inventory* is strongly present in low fidelity hospitals with a hindering influence, while in high fidelity hospitals, this factor has a neutral to enabling influence when somewhat

present. Similarly, *Large hospital/Complex surgical procedures* is strongly present in low fidelity hospitals with a hindering influence, and minimally to somewhat present in high fidelity hospitals with a neutral to enabling influence.

> *"Technology … could work for focus clinics and regional hospitals, but here we only perform highly complex surgeries, with many different trays… too many." (Participant 4-3, low fidelity)*

> *"We're not very big, we don't have many sets. So if you make some hours, generally speaking, you know the contents of the trays." (Participant 1-1, high fidelity)*

An interesting factor that influenced weighing system fidelity was a *Continually updated tray reference weight*. In high fidelity hospitals, the reference weight is updated for each tray as it is assembled in the SPD. In low fidelity hospitals, there is variation. In one hospital, the reference weight is only updated when a tray's contents are permanently changed:

> *"Another reason that weighing is not used for instrument counting, is because trays don't always have the same content, due to repairs or maintenance. But the reference weight stays the same… it's only updated after a permanent change."(Participant 3-1, low fidelity)*

However, a participant from the SPD considered this an extra safety measure:

> *"It's an indication for the OR nurses that there's something different about the tray… that they should pay extra attention to counting." (Participant 3-2, low fidelity)*

In the other low fidelity hospital, the reference weight is updated for each tray. However, in the OR staff's experience, this is not always done:

> *"It starts in the SPD… the reference weight gets updated for each tray, but this is sometimes forgotten." (Participant 4-3, low fidelity)*

Influencing factors identified only for hospitals with high or low fidelity

Two factors were identified in only high fidelity hospitals. They were the factors *Teamwork/Collaboration* in the i-PARIHS domain of Recipients, and *Only certain procedures require weighing* in the Facilitation domain. Two factors were identified only in hospitals with a low fidelity status. They were *Only unsterile instruments can be weighed* in the Innovation domain and *Limited consideration for weighing during OR renovations* in the Facilitation domain.

Innovation-related factor With a strong presence and hindering influence, the fact that *Only unsterile instruments can be weighed* is deemed a major limitation of the weighing system in low fidelity hospitals:

> *"If you can weigh earlier in this process, when the tray is still sterile, you might be able to use the weight for counting instruments." (Participant 3-1, low fidelity)*

Recipients-related factor While *Interdisciplinary coordination/Trust in other people's work* summarises the coordination that occurs between departments, *Teamwork/Collaboration* encompasses individuals working together. It is somewhat to strongly present in hospitals with high fidelity, with an enabling influence.

Facilitation-related factors One high fidelity hospital facilitated the weighing system by excluding complex surgical trays from weighing. The influence on weighing system fidelity is unclear or neutral.

One low fidelity hospital renovated their OR complex in recent years. Even though weighing was part of the process for many years already, there was limited consideration for it during these renovations. Currently, weighing stations are placed on a small shelf in each OR, where space to adequately weigh trays is lacking.

Discussion

Nationwide, around one quarter of hospitals have a weighing system in place, initially intended for surgical instrument counting. In this study, we examined factors influencing the fidelity of these weighing systems across four hospitals. We found that the fidelity level varied between hospitals, often with many factors influencing it. These factors, whether innovation-related, recipients-related, or context-related, seem to be closely related or dependent on each other, with facilitation-related factors taking a central position.

The weighing system as a tool to count surgical instrument is flawed, which is highlighted by the widespread presence of hindering Innovation-related factors within the i-PARIHS framework. Most importantly, all participants mentioned the variation in surgical instruments. As hospitals in general can process over 100,000 instrument trays on a yearly basis, it is unrealistic to assume all instruments are in similar condition due to maintenance or repairs, or, in the case of a single type of instrument, even of the same specification [12–14]. Variations of a

few grams in a single instrument are amplified in a tray that could contain a hundred instruments.

Due to these common issues with the innovation itself, efforts were made in all studied hospitals to facilitate it. For example, each hospital determined a weight margin of error based on their own instrument inventory and situation. Similarly, all hospitals removed highly variable items, such as kidney dishes or cables from trays before weighing to increase accuracy. These facilitation efforts were done to increase the usability of the weighing system on the organisation-level (i.e. the hospital). However, some facilitation efforts took place on the department level (i.e. the OR or SPD). Both an OR and an SPD (in different hospitals) added a weighing station at the beginning of their workflow as to be less dependent on reference weights determined at the other department, as they were often found to be inaccurate. This highlights an organisational complexity where interdepartmental coordination is limited. Rather than holistically solving inaccuracies in the weighing process, each department, through facilitation, takes steps to mitigate the effect of inaccurate weighing of the other department.

The presence of an organisation-wide facilitator is an integral part of implementing an innovation [19]. In the absence of organisation-wide facilitation efforts, more inaccuracies in the weighing process occur, negatively impacting staff attitude towards it. Where participants in high fidelity hospitals acknowledge the technology as a helpful aide in their work, participants in low fidelity hospitals experience a more flawed technology that does not contribute to their manual instrument counting. This is a remarkable difference, since all hospitals experience similar issues relating to the innovation itself.

Adaptation of the innovation

As stated earlier, weighing surgical trays to determine whether all instruments are present is a flawed system. While some OR departments have found ways to make use of the system, its value is not apparent in all OR departments. In SPDs however, the weighing system is consistently deemed a valuable part of 'track and trace' systems on the instrument tray level. SPD participants mention that the various weighing stations allow them to determine where an instrument tray currently is, and, if an instrument is lost, pinpoint where it might have been lost. Thus, the weighing system has been adapted to a different use-case which allowed for its sustained use up until today. This shows that in some cases, in order to achieve a high sustainment, an innovation's fidelity is reduced. This is in line with literature, where adaptation of the innovation leads to long-term sustainment [20–22].

Lessons learnt for implementing future technologies

We have investigated a technology that was implemented in various hospitals around twenty years ago. Several general lessons can be extracted from this long-term post-implementation study, which are applicable to future implementation efforts of new innovations.

First, we have identified many influencing factors across all four i-PARIHS domains of Innovation, Recipient, Context, and Facilitation. These factors often relate to each other or are interdependent on each other. To drive new implementation efforts, it is essential to be aware of the presence of these factors. By using an established implementation framework, influencing factors can be identified at an early stage, allowing adjustments to the implementation process as needed.

Second, we have seen that as hospitals get larger and more complex, interdepartmental coordination lags behind. When a technology is used across multiple departments, a lack of coordination can leave procedural errors unresolved. This results in a lack of trust in the technology among users. As the innovation studied in this work was implemented around twenty years ago, we were unable to determine whether a facilitator was present then. A designated facilitator could be helpful in creating awareness of how specific procedures affect other users, as well as advocate for the needs of specific user groups to decision-makers. For example, one participating hospital renovated their OR department without providing a designated space for the weighing stations in the new ORs, which lead them to be placed on a small shelf that is inconducive to an efficient weighing process.

Third, despite using a similar weighing system, all hospitals adapted their workflow to the technology in a different way. These adaptations, however necessary for local implementation, hinder knowledge-sharing between organisations and thus broader implementation of innovations. A collaborative approach among facilitators from different organisations could improve knowledge sharing and reduce local variations, improving innovation dissemination.

We do note that these insights were obtained from a qualitative study in a limited number of hospitals. However, the current study already showcases a wide range of practices which underlines our conclusions.

Conclusions

While novel developments might have a positive effect on the healthcare labour shortage, implementing these developments warrants extended efforts. As our study shows, practices and sentiments diverge significantly

even in the case of straight-forward technology such as a weighing scale. Gaining an overview of the influencing factors involved is essential to an efficient practice and positive user sentiments. For this, large hospitals with a complex organisational structure could benefit from a designated facilitator.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s43058-025-00808-8.

Supplementary Material 1.

Supplementary Material 2.

Acknowledgements
The authors would like to thank Benno Hendriks, Teddy Vijfvinkel, Rick Butler, and Emanuele Frassini for their input during the study's consensus meetings. The authors would also like to thank Julian Houwen for providing feedback on the selection of implementation outcome framework.

Authors' contributions
AK is responsible for conceptualisation, writing the research protocol, formulating the interview questions, carrying out the interviews, proposing intitial findings, and writing the manuscript. ME and JD contributed to the final manuscript and provided project supervision. All authors read and approved the final manuscript.

Funding
The corresponding author is funded by the Phoenix Foundation (Delft, the Netherlands), which has no influence on this research.

Data availability
Interview questions are available in Appendix A. The datasets generated and analysed during the current study are not publicly available due to the sensitivity of hospital data but are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
Ethics approval was obtained from the TU Delft's Human Research Ethics Committee (application number 4055). Informed consent was obtained from study participants.

Consent for publication
Consent was obtained from study participants for publication of the findings of this study.

Competing interests
The authors declare that they have no competing interests.

Received: 13 May 2025 Accepted: 11 October 2025
Published online: 10 November 2025

References

1. Olivere LA, Hill IT, Thomas SM, et al. Radiofrequency identification track for tray optimization: an instrument utilization pilot study in surgical oncology. J Surg Res. 2021;264:490–8.
2. Toti G, Garbey M, Sherman V, et al. A smart trocar for automatic tool recognition in laparoscopic surgery. Surg Innov. 2015;22(1):77–82.
3. Hanada E OA, Hayashi M, Sawa T. Improving efficiency through analysis of data obtained from an RFID tag system for surgical instruments. In: IEEE 5th International Conference on Consumer Electronics. Berlin: IEEE; 2015. p. 84–7. https://doi.org/10.1109/ICCE-Berlin.2015.7391339.
4. Kang HS, Khoraki J, Gie J, et al. Multiphase preclinical assessment of a novel device to locate unintentionally retained surgical sharps: a proof-of-concept study. Patient Saf Surg. 2023;17(1):10.
5. Hill I, Olivere L, Helmkamp J, et al. Measuring intraoperative surgical instrument use with radio-frequency identification. JAMIA Open. 2022;5(1):ooac003.
6. Nakano A, Nagamune K. A development of robotic scrub nurse system - detection for surgical instruments using faster region-based convolutional neural network –. J Adv Comput Intell Intell Inform. 2022;26(1):74–82.
7. Meijsen P, Meers E. Uitgeteld?2006. https://sterilisatievereniging.nl/wp-content/uploads/2015/08/2006-Onderzoek-naar-compleetheidscontroles-van-instrumenten.pdf. Accessed 11–08–2023.
8. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health. 2011;38(2):65–76.
9. Carroll C, Patterson M, Wood S, et al. A conceptual framework for implementation fidelity. Implement Sci. 2007;2:40.
10. Rabin B, Brownson R, Haire-Joshu D, et al. A glossary for dissemination and implementation research in health. J Public Health Manag Pract. 2008;14:117–23.
11. George RE, Bay CC, Shaffrey EC, et al. A day in the life of a surgical instrument: the cycle of sterilization. Ann Surg Open. 2024;5(1):e381.
12. Stockert EW, Langerman A. Assessing the magnitude and costs of intraoperative inefficiencies attributable to surgical instrument trays. J Am Coll Surg. 2014;219(4):646–55.
13. Acland R. Notes on the care restoration and repair of microsurgical instruments. Indian J Plast Surg. 2006;39(1):51–6.
14. Webster RJ, Yaniv ZR, Glaser B, et al. Surgical instrument similarity metrics and tray analysis for multi-sensor instrument identification. In: Medical imaging 2015: image-guided procedures, robotic interventions, and modeling. 2015.
15. Kim B, Sullivan JL, Ritchie MJ, et al. Comparing variations in implementation processes and influences across multiple sites: what works, for whom, and how? Psychiatry Res. 2020;283:112520.
16. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ)- a 32-item checklist for interviews and focus groups. Int J Qual Health Care. 2007;19(6):349–57.
17. Harvey G, Kitson A. Parihs revisited: from heuristic to integrated framework for the successful implementation of knowledge into practice. Implement Sci. 2016;11:33.
18. Proudfoot K. Inductive/deductive hybrid thematic analysis in mixed methods research. J Mixed Methods Res. 2022;17(3):308–26.
19. Kim B, Sullivan JL, Brown ME, et al. Sustaining the collaborative chronic care model in outpatient mental health: a matrixed multiple case study. Implement Sci. 2024;19(1):16.
20. Olsson TM, von Thiele Schwarz U, Hasson H, et al. Adapted, adopted, and novel interventions: a whole-population meta-analytic replication of intervention effects. Res Soc Work Pract. 2023;34(8):860–72.
21. Hailemariam M, Bustos T, Montgomery B, et al. Evidence-based intervention sustainability strategies: a systematic review. Implement Sci. 2019;14(1):57.
22. Nevedal AL, Widerquist MAO, Reardon CM, et al. Understanding pathways from implementation to sustainment: a longitudinal, mixed methods analysis of promising practices implemented in the Veterans Health Administration. Implement Sci. 2024;19(1):34.

Publisher's Note