

BMJ Open Sustainability of performance improvements after 26 Kaizen events in a large academic hospital system: a mixed methods study

Erik Haapatalo ,¹ Elina Reponen ,² Paulus Torkki ^{1,3}

To cite: Haapatalo E, Reponen E, Torkki P. Sustainability of performance improvements after 26 Kaizen events in a large academic hospital system: a mixed methods study. *BMJ Open* 2023;**13**:e071743. doi:10.1136/bmjopen-2023-071743

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2023-071743>).

Received 19 January 2023
Accepted 15 August 2023



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Faculty of Medicine, Department of Public Health, University of Helsinki, Helsinki, Finland

²HUS Shared Group Services, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

³Department of Industrial Engineering and Management, Aalto University, Espoo, Finland

Correspondence to

Erik Haapatalo;
erik.haapatalo@helsinki.fi

ABSTRACT

Introduction Implementing Kaizen can improve productivity in healthcare but maintaining long-term results has proven challenging. This study aimed to assess improved performance achieved and sustained by Kaizen events and find explanatory factors for the persistence or decline of long-term results.

Methods Kaizen events were conducted in 26 specialised healthcare units in a large academic hospital system in southern Finland. Primary data for mixed methods analysis was collected from each unit with 21 semi-structured interviews, Kaizen report files and performance metrics.

Results Fifteen explanatory factors were found in this study. Work culture and motivation for continuous improvement stood out as the most important explanatory factor for the persistence of long-term results—lack of time for improvement activities and high workload for the decline. Success in preparation and follow-up was associated with sustained long-term results. Thirteen units achieved long-term results, three units could not sustain the performance improvements and five units struggled to make any improvements.

Conclusions This study explains the long-term sustainability of performance improvements, bringing new insights to Kaizen research. Our findings can guide organising successful Kaizen events. The events can be worth organising even though long-term performance improvements are not guaranteed. Units with supportive working culture and motivation for the Kaizen event will likely succeed. A unit should aim to create a supportive foundation for Kaizen before organising a Kaizen event. Units that lack the foundation can be identified, trained and guided to increase their chances of success. Pitfalls like high workload and insufficient follow-up should be proactively identified and appropriately managed by allocating the required time and resources for the development work.

INTRODUCTION

Lean is used to face healthcare systems' challenges with variable success. Lean originated in the Toyota Production System, and Kaizen is part of this system. The term Kaizen exhibits ambiguity and inconsistency in literature and practice.^{1,2} For example, Kaizen can be perceived as a management philosophy,

STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ Out of 26 units that participated in Kaizen events, we managed to recruit 21 interviews, mainly with chief physicians.
- ⇒ The interviews are triangulated with quantitative data from performance metrics used in Kaizen.
- ⇒ Evaluating Kaizen improvements, studying their long-term sustainability and recognising the associated factors are still rare in the healthcare literature.
- ⇒ Using a structured framework of sustainability levels to study Kaizen helps make this study repeatable.
- ⇒ Possible bias in data collection: difficulties recalling details, pressure to give positive results in the interview and the lack of qualitative data in some cases.

a component of Total Quality Management, and a theoretical principle for improvement methodologies and techniques.¹ However, some fundamental principles are commonly shared in each interpretation: Kaizen is continuous, incremental and participative.² Continuous improvement can be defined as 'a company-wide process of focused and continuous incremental innovation'.³ Continuous improvement is sustained when permanently integrated into the organisation's culture and daily activities. Consequently, it often delivers results in performance. Kaizen aims to initialise or enhance continuous improvement by challenging and empowering everyone in an organisation to find opportunities for improvement in daily work by implementing low-cost, low-risk changes in daily work to improve performance incrementally.^{4,5} The Kaizen cornerstones from the work of Suárez Barraza *et al*¹ describe the basic principles: teamwork, eliminating muda, gemba management, education and training, commitment from top management, proposing and applying improvements, focus on process and standards. The execution of the Kaizen in our study follows

these principles and is explained in more detail in the methodology section.

Kaizen is often introduced by Kaizen events: 5-day intensive workshops, during which a multidisciplinary team focuses on a predefined development target.^{6,7} The Kaizen event should be a catalyst for continuous improvement in healthcare units. However, there are sometimes challenges in organising Kaizen events in an ideal way. Usually, before the Kaizen event, there should be preparation and a follow-up period after the event.⁸ Successful Kaizen requires good teamwork and extensive involvement of employees in improvement activities.⁹ Therefore, Kaizen teams, their social dynamics and their internal processes are important in Kaizen.¹⁰ For a team to get the most out of Kaizen, their conversational space should encourage freedom of expression among individuals and openness of interaction during the Kaizen events.¹⁰

The performance improvements achieved by Kaizen events are significant to the success of a Lean transformation in an organisation.⁶ Still, maintaining long-term performance improvements has proven challenging, for example, in the industry where estimates of the benefits of Kaizen events are conflicting.¹¹ Results of Kaizen in public hospitals and private clinics include (1) decreased lengths of appointments and stays after surgery, (2) reduction in cost per operation, inventory of clinical stock and travel distance by the staff, (3) elimination of none-value-adding activities and variability and (4) improved process flow, patient outcomes and increased service quality.¹²⁻¹⁵

Previously identified factors associated with successful Kaizen implementation in healthcare include goal clarity and alignment,^{16,17} cross-functionality, stakeholder representation and commitment to change,¹⁶⁻¹⁸ general management support,^{16,17} organisational structure, culture and policies which provide an opportunity for effective communication within and across organisational hierarchies to enable sharing of learning and experience, local autonomy and empowerment.^{3,16,17} The Kaizen process is also affected by changes in environmental factors outside the organisation and beyond managerial control.¹⁸ Successful Kaizen requires cultural transformation,¹⁹ and employee attitude and change acceptance can potentially predict long-term performance.²⁰ However, such sociocultural factors are often less focused on, and more emphasis is put on tools and mechanisms of implementation.²¹ Factors that may inhibit the success of Kaizen in healthcare are distrust of Lean, low involvement, limited understanding of processes, failure of communication, difficulties in data collection and excess bureaucracy.²² A lack of understanding of Kaizen and its linkage to the organisation's overall goals may limit the sustainability of performance improvements.²³

The long-term results' persistence or decline can be described with sustainability levels, such as in the sustainability framework developed by Bateman and David.^{8,18} Our study has adapted sustainability levels from the framework with appropriate adjustments. The sustainability levels are presented in [table 1](#).

There are more case studies describing Kaizen events implementations or projects in healthcare than empirical investigations about Kaizen events' impact or critical success factors.²⁴ Furthermore, the research on Lean generally lacks studies on Kaizen and continuous improvement in healthcare.^{22,25} This study complements Kaizen literature in the public healthcare sector¹⁵⁻¹⁷ by providing insights into performance improvements, their long-term sustainability and the associated factors while using sustainability levels.

Organising Kaizen events is a significant investment for the healthcare organisation involved in this study. Still, a systematic evaluation of the performance improvements from the Kaizen events or their long-term sustainability has not been done. Therefore, it is essential to study Kaizen events' impact and identify explanatory factors for the persistence or decline of long-term results. Consequently, a healthcare organisation can make an evidence-based decision about initiating, continuing, abandoning or adjusting the Kaizen activities. Explanatory factors may also help in improving Kaizen events in the future.

Aims and objectives

This study was designed to investigate Kaizen events in detail to answer the following research questions:

RQ1. What immediate performance improvements did the teams that participated in the Kaizen events achieve?

RQ2. Have the teams been able to maintain or further improve the performance improvements in the longer-term?

RQ3. What factors explain the persistence or decline of long-term results after Kaizen events?

These research questions are answered in the results section of this study based on the findings from 26 cases of Kaizen implementation in specialised healthcare units. The structure of this research article has four main parts. (1) The first part aims to introduce the concept of Kaizen in healthcare, describe the sustainability levels that are an essential part of this study and present the research questions. (2) The second one describes the experimental set-up, data collection and analysis. (3) The third part shows the results achieved by the Kaizen implementation. (4) The fourth one concludes the research article by discussing the meaning of the results.

METHODS

Experimental set-up

Public healthcare services are the foundation of the Finnish healthcare system, and every permanent resident of Finland is entitled to these services. Specialised medical care, a part of public healthcare, is mainly provided by university hospitals and central hospitals.²⁶ The study was conducted for Finland's largest academic hospital system, The Hospital District of Helsinki and Uusimaa (HUS), to learn from Kaizen events by following the research questions. HUS is responsible for 2.2 million inhabitants and a staff of 27 000 employees.²⁷ HUS identified specialised

Table 1 Sustainability levels of performance improvements in Kaizen adapted from Bateman and David^{8 18}

Sustainability level	Sustainability of results	Immediate results	Type of long-term results	Follow-up carried out as planned
A	The unit sustains the results from Kaizen activities and procedures developed in the Kaizen event. Objectives are met or even exceeded. Further improvement is made by disseminating successful procedures in other units or applying the Kaizen philosophy to solve new problems.	Yes	Performance improvements, qualitative improvements, new procedures.	Yes
B	The unit sustains procedures and at least some results. Objectives are met, or at least some improvement in performance metrics has occurred. Significant further improvement is not made.	Yes	Performance improvements, qualitative improvements, new procedures.	Mostly yes
C	The unit makes qualitative improvements but cannot sustain or confirm them by performance metrics. Generally, the unit sustains at least some changes in procedures.	Yes	Qualitative improvements, new procedures.	No
D	The unit struggles to make any improvements and cannot sustain them. Minor or no changes in procedures are sustained.	No	At most, some new procedures.	No

Sustainability levels A, B, C and D describe the unit's ability to sustain long-term results. Performance improvements are the results of the Kaizen events, which are confirmed by performance metrics. Qualitative improvements are not monitored with performance metrics and thus are based on the subjective experiences of the interviewees. New procedures often appear as a change of practices in the process but can also be, for example, the recruitment of new staff, new appointment books, change of work rhythm, work instructions, induction frame or the introduction of daily management. Performance metrics are used to obtain data by monitoring a selected variable reflecting the unit's capability. Immediate results are usually achieved during or immediately after the Kaizen event. Immediate results can be performance improvements or qualitative improvements. Follow-up includes collecting and monitoring performance metrics and implementing and maintaining procedures after the Kaizen event.

medical care units suitable for the Kaizen events and organised Kaizen events one to two times a year. HUS organised the Kaizen events, collaborating with the units and their Kaizen teams, HUS Lean development unit and an external Lean consultant. A total of 26 of these Kaizen events were organised between 2017 and 2020. Specialised healthcare units of HUS participated in Kaizen events in 2017 (n=4), 2018 (n=11), 2019 (n=8) and 2020 (n=3). Therefore, this study uses retrospective data from multiple Kaizen events. Usage of the general principles of case study research improves substantiation for the findings of this study.²⁸ For example, triangulation with multiple data collection methods such as semi-structured interviews and qualitative data from performance metrics.

Kaizen activities in HUS can be explained in three main stages: preparation, Kaizen event and follow-up. Ideally, preparation includes determining the multidisciplinary Kaizen team, targets, objectives, performance metrics, timetable, understanding the problem and the data behind it, understanding the process, visualising demand and production: quantity, types and variation and identifying unfinished production and the demand it causes. During the 5-day Kaizen event, the team discusses the unit's background, the current state and the objective, brainstorms ideas, tests and improves the best ideas, plans and implements the changes, plans management and follow-up of Kaizen. The most intensive follow-up period lasts from approximately 1–3 months, during which the new procedures are stabilised, teams go through their

to-do lists, teams get management support in implementing changes and follow-up meetings are held. The effectiveness of the implemented changes is monitored with suitable performance metrics such as lead time or the number of patients treated. Ideally, these follow-up activities are continued after the most intensive period, but in reality, most units have trouble maintaining such intensive follow-up. In this study, the Kaizen teams consist of all roles involved in the unit's operation, including nurses, specialist doctor, senior doctor, ward manager, secretary, facility maintenance and representative of customers. The support of the head nurse, line manager and branch manager is available if needed. The quality manager and other experts participate if necessary. The Kaizen team is coached by an external Lean consultant.

Data collection

The data for this study was gathered with semi-structured interviews, Kaizen reports and performance metrics.

The semi-structured interview addressed the following themes:

1. Unit's background.
2. Preparing for the Kaizen event.
3. The Kaizen event.
4. The subject of development.
5. The implementation of changes.
6. Results and follow-up of Kaizen.
7. Persistence of changes and results.
8. Future of the improvement activities in the unit.

In addition to open questions, the interview framework had four semi-quantitative questions on a scale of 1–5. An English version of the interview framework is provided as a supplementary file. Interview invitations were emailed to the Kaizen team leaders in May 2021. The initial rate of interviews was 65%. We sent one reminder and eventually recruited 81% of the units for the interviews. The interviewees were Kaizen team leaders or representatives recommended by the team leaders. All but one interviewee had personal experience in their unit's Kaizen event. The interviewees were mainly occupied as chief physicians. We organised the interviews in the form of online meetings and held 21 semi-structured interviews in June, July and August of 2021. There were signs of saturation of the data during the last few interviews. All the interviews were audio recorded with the interviewees' consent. External transcription service Tutkimustie Oy transcribed the audio recordings of the interviews. In the early stages of the transcription process, we confirmed the quality of the transcription service by comparing the audio to the text. Transcription resulted in 21 text documents ranging from 2724 to 7699 words and an average length of 5157.

In connection with the email questionnaire and the interviews, 13 units shared Kaizen reports for research purposes and 9 of them sent the reports as attached data files. These reports contain quantitative data at two or more time points in a long-term follow-up period. Therefore, 13 units reported sufficient evidence of improvements in performance metrics, which are summarised in [table 2](#). Triangulation of data sources was done by reconciling qualitative data from the interviews and quantitative data from the Kaizen reports.²⁹ The final data used in this study: 21 interview transcripts, 13 Kaizen reports with qualitative data and 5 units participated in a Kaizen event but did not attend an interview.

Data analysis

Mixed methods were used to analyse the data. The interview transcriptions were analysed using ATLAS.ti research software (ATLAS.ti Scientific Software Development GmbH, Berlin, Germany). All the 21 text documents acquired by transcribing the interview audio files were imported to ATLAS.ti. Using ATLAS.ti, the text documents were scanned manually, and the relevant quotes from the text were given a code. The quotes and the codes were used to facilitate the analysis of the results and explanatory factors. The long-term results' persistence or decline was evaluated by determining sustainability levels for each unit. The sustainability levels were determined for each unit based on the criteria in [table 1](#) and the available information on Kaizen in each unit. A and B-level units confirmed the quantitative performance improvements with metric data in at least two time points. A-level units made progress beyond the objective, whereas B-level units mostly reached and sustained their objective during the period of data collection. C-level units did not reach their objective or could not confirm the progress

with metric data but mentioned qualitative improvements such as new procedures that make the daily work easier. D-level units struggled to make any improvements and could not sustain them.

Units' self-assessment of their success in Kaizen areas was done with the semi-quantitative questions: (1) Preparation—'On a scale of 1–5, how prepared were you for the Kaizen event?'. (2) Kaizen event—'On a scale of 1–5, how well did you get started in the Kaizen event's workflow?'. (3) Follow-up—the average numerical value of 'On a scale of 1–5, how well does continuous improvement take place in your unit?' and 'On a scale of 1–5, how optimistic are you about future development activities in your unit?' rounding to the nearest whole number.

The RQ3 was approached by identifying the explanatory factors through content analysis and coding. The codes include information on whether the factor was explanatory for persistence (+) or decline (–) of results, a brief description of the factor and the level of sustainability of the subject unit: A, B, C or D. For example, A (+): culture, traditions and atmosphere. Fifteen general themes were identified. The themes are used as explanatory factors for the persistence or decline of the results to answer RQ3. The factors were categorised into five upper categories based on the five-factor Lean Healthcare Implementation Self-Assessment Instrument, which has been used to evaluate Lean implementation in HUS. The five categories are leadership, commitment, standard work, communication and daily management system.³⁰

Patient and public involvement

The research questions and outcome measures were designed to improve the quality of healthcare and promoting the patients' and the staff's priorities, experience and preferences are in the essence of this study. The study's data is anonymous and previously used locally in the units for management purposes. Therefore, the patients are indirectly involved in this study because the data is initially collected from the records of ordinary, guideline-following patient visits. Patients were not involved in the recruitment and conduct of the study. The results of this study will be disseminated to the research community and the participants through publication.

Ethical considerations

This study did not involve any questions of a sensitive nature or, for example, related to the interviewee's state of health, nor questions that could have had adverse effects on the interviewee's employment relationship if answered. Informed consent was obtained from all subjects. All methods were carried out following relevant guidelines and regulations.

RESULTS

Of the 21 interviewed units, 62% achieved sustainability levels A or B. The units reaching sustainability level A or B achieved immediate results and eventually sustained

Table 2 Long-term performance improvements and their sustainability

Unit's description	Sustainability level	The main objective and the performance metric	The unit reached the objective	Long-term performance improvement
Clinical microbiology, diagnostic	A	Timely and high-quality samples (%)	Yes	95% of sample processing from approximately 4 days to less than 2 days.
Gynaecology and obstetrics, ward	A	Earlier mobilisation after the procedures (hour)	Yes	Baseline 6 hours and after Kaizen 3 hours.
Gynaecology and obstetrics, unit of operations/obstetrics	A	Reduction of delivery room closures (the number of)	Yes	In 2017: 64 closures, in 2018: 39 closures, in 2019: 29 closures, in 2020: 60 closures.
Gynaecology and obstetrics, unit of operations	A	Reduction of operation time (hour)	Yes	Baseline 2.5 hours and after Kaizen 1.18 hours.
Youth psychiatry, outpatient clinic	A	Shortening of the examination periods (day)	Yes	Baseline 2–4 months and after Kaizen 3–14 days.
Dentistry, unit of operations	A	Shortening the queue (person)	Yes	Baseline 150 people and after Kaizen 50 people.
Cancer diseases, ward	A	Reduction of the time of admission from ER (hour)	Yes	Baseline 0–36 hours and after Kaizen 0–4 hours.
Gynaecology and obstetrics, outpatient clinic	B	Shortening the queue (month)	Yes	Baseline 4 months and after Kaizen 1.5 months.
Assistive equipment centre	B	Shortening the queue (day)	Yes	Baseline 254 days and after Kaizen 14 days.
Psychiatry, outpatient clinic	B	To increase the share of patients with a time of admission under 14 days (%)	No	Baseline: 6% of patients had a time of admission under 14 days, and after Kaizen, 9%. The objective is 80%.
Rheumatology, outpatient clinic	B	Reduction of the proportion of people waiting for a first appointment for more than 1 month (%)	Yes	Baseline 60% and after Kaizen 10 %.
Internal diseases, ward	B	To increase the share of on-time discharges (%)	Yes	Baseline 34% and after Kaizen 61 %.
Ear, nose and throat diseases, outpatient clinic	B	Reduction of the lead time (appointment, guidance, dictation) (min)	Yes	Baseline 94 min and after Kaizen 44 min.
Cardiac and thoracic surgery, ward	C	Earlier discharge from the ward (hour)	Yes, momentarily	
Infectious diseases, ward	C	Earlier discharge from the ward (hour)	Yes, momentarily	
Cardiology, unit of operations	C	To increase the number of operations (the number of)	No	
Biobank, research laboratory	D	To increase the number of samples (pcs/month)	No	
Pathology, diagnostic	D	To increase the flow of samples (pcs)	No	
Ear, nose and throat diseases, outpatient clinic	D	Enhancing the work process and reducing cancellations (the number of)	No	
Emergency department	D	Reduction of the lead time (hour)	No	

Continued

**Table 2** Continued

Unit's description	Sustainability level	The main objective and the performance metric	The unit reached the objective	Long-term performance improvement
Cancer diseases, outpatient clinic	D	Reduction of the time of admission (hour)	No	

Sustainability levels A, B, C and D describe the unit's ability to sustain long-term results. The baseline before Kaizen and the situation after Kaizen describe long-term performance improvements. They are average values from performance metrics. The unit's description shows the area of specialty and the type of the unit.

Sustainability levels: (A) The unit sustains the results from Kaizen activities and procedures developed in the Kaizen event. Objectives are met or even exceeded. Further improvement is made by disseminating successful procedures in other units or applying the Kaizen philosophy to solve new problems. (B) The unit sustains procedures and at least some results. Objectives are met, or at least some improvement in performance metrics has occurred. Significant further improvement is not made. (C) The unit makes qualitative improvements but cannot sustain or confirm them by performance metrics. Generally, the unit sustains at least some changes in procedures. (D) The unit struggles to make any improvements and cannot sustain them. Minor or no changes in procedures are sustained.

ER, emergency room; pcs, pieces.

long-term performance improvements. The Kaizen events of these units can be considered a success. Level A units also disseminated successful changes to other units or solved new problems in their unit with Kaizen. One unit with sustainability level B achieved performance improvements and sustained them but could not reach the initial objective. Level C units (14%) achieved immediate results but could not sustain them. The success of level C units is below expectations or unclear due to the lack of quantitative data. Five units (24%) fall into level D, meaning they were struggling to make any improvements; thus, the impact of their Kaizen events was limited or non-existent. Five units participated in a Kaizen event and were invited to an interview but did not respond or attend it. Estimating these units belong to level D, sustainability levels C and D would then contain 50% of 26 Kaizen units and levels A and B the other half.

Every unit aimed to improve processes as the primary objective. Accordingly, the performance metrics monitored time or quantity. The performance metrics were monitored during the Kaizen events to evaluate the effectiveness of interventions. However, the data of these immediate results is unavailable because it was not preserved for research purposes. To answer RQ1, we must rely on the interview data and assume that units had immediate results before long-term performance improvements. Levels A, B and C achieved immediate results—qualitative or quantitative. Examples of qualitative immediate results are (1) facilitation of daily work by better communication or management of items, (2) elevated team spirit by enhancing multidisciplinary teamwork and aiming towards a common goal and (3) improved patient experience by providing faster, better and more available service. Two level C units reached objectives momentarily, indicating that they had quantitative immediate results but could not sustain them. Quantitative immediate results of A and B level units are likely similar to the long-term performance improvements.

The sustainability levels in [table 2](#) contribute to RQ2 by showing which units sustained performance

improvements in the longer term. [Table 2](#) also shows what kind of long-term performance improvements units achieved.

Level A units evaluated their success in Kaizen phases significantly higher than the others ([table 3](#)). Follow-up is the most significant difference between successful A and B-level units and unsuccessful C and D-level units. A and B-level units self-assessed their success in follow-up two times higher than C and D-level units. Units with level D sustainability self-assessed their success in preparation as about 30% lower than the others on average.

The quotes explaining the persistence or the decline of long-term performance improvements are summarised into 15 explanatory factors ([table 4](#)). The factors are divided into five categories: commitment, daily management system, communication, standard work and leadership, adapted from Reponen *et al.*³⁰ Work culture and motivation for continuous improvement stood out for the persistence of long-term results—lack of time for improvement activities and high workload for the decline.

Table 3 Units' self-assessment of success in Kaizen phases

Sustainability levels	Units	Preparation	Kaizen event	Follow-up	Sum
A	7	4.3	4.9	4.9	14.1
B	6	3.2	3.5	3.8	10.5
C	3	3.7	4.3	2.0	10
D	5	2.6	3.2	2.4	8.2

Units assessed their success in Kaizen phases on a scale of 1–5. Five means very successful, and one means unsuccessful. Sustainability levels A, B, C and D describe the unit's ability to sustain long-term results. Levels A and B sustained long-term results, and levels C and D did not sustain long-term results or struggled to make any improvements.

Table 4 Explanatory factors for the persistence and decline of long-term results

Explanatory factor	Quotes	Persistence (%)	Decline (%)	Category
Work culture and motivation for continuous improvement	32	66	34	Commitment
Employee turnover	6	0	100	Commitment
Commitment to Kaizen	4	50	50	Commitment
Lack of time for improvement activities and high workload	16	0	100	Daily management system
Daily or otherwise regular meetings and events	10	100	0	Daily management system
Working time is devoted to improvement activities	4	100	0	Daily management system
Usage of principles of continuous improvement	4	100	0	Daily management system
Collecting improvement ideas	4	100	0	Daily management system
Contacts, multidisciplinary cooperation and communication	11	64	36	Communication
Clarity of the objectives and purpose of Kaizen among staff	11	45	55	Communication
Monitoring and reporting of processes and performance metrics	11	18	82	Standard work
Clarity of areas of responsibility and responsible persons	4	75	25	Standard work
Leadership supports continuous improvement	8	88	12	Leadership
Challenges related to the Kaizen event	5	0	100	Leadership
Challenges in implementing changes	3	0	100	Leadership

This table shows explanatory factors for the persistence and decline of long-term performance improvements achieved by Kaizen. The factors consist of interview quotes explaining the persistence or the decline. For example, the explanatory factor 'Work culture and motivation for continuous improvement' was discussed in 32 quotes, and 66% of them explained persistence. Respectively 34% of the quotes explained decline, meaning that work culture and motivation were not supportive and it caused the results to decline.

The higher the sustainability level of units, the more units mentioned factors explaining the persistence of the results. For example, level A units mentioned factors for persistence nine times more than level C or D units and, respectively, level B units four times more. Factors in commitment and daily management system categories explained the persistence with twice as many quotes than the other categories on average. The finding indicates that the units with supportive work culture and motivation, regular meetings, the devotion of working time for improvement activities and which collect improvement ideas perform better. Furthermore, the units that perform the best likely have clear objectives, areas of responsibility, multidisciplinary cooperation, and leadership support. The lower the sustainability level of units, the more units mentioned factors explaining the decline of the results. Level D units mentioned explanatory factors for the decline over two times more than the other units. Level D units had five times more negative factors in commitment than other units on average and, respectively, four times more in communication. The finding indicates that high employee turnover, challenges in communication, lack of commitment and the work culture and motivation degrade long-term performance improvements. The units with poor improvement will likely lack clarity of the objective and have trouble monitoring the performance metrics.

DISCUSSION

Supportive working culture and motivation create a foundation for Kaizen. Successful follow-up is essential for the persistence of long-term results. High workload and negative sociocultural factors related to commitment to Kaizen decline long-term results. The long-term quantitative results of Kaizen events in this study are improvements in processes measured by performance metrics. Twelve out of 21 interviewed units (57%) reached their objective and achieved long-term results. Most of the units reached immediate results, qualitative or quantitative.

Our findings can support making evidence-based decisions for initiating, continuing, abandoning or adjusting the Kaizen activities. This study can guide organising future Kaizen events to improve the selection of the units and assessment of effectiveness. Explanatory factors for persistence can be used to find units with the best prerequisites for a successful Kaizen event. The units where the factors are already present can be identified as low-hanging fruits, in which performance could likely be improved with Kaizen. Units, where the factors lean toward decline rather than persistence are less likely to succeed in Kaizen. It should be carefully considered whether the factors can be enhanced to support Kaizen in such cases. An evaluation of whether a unit has factors that favour success in Kaizen can be done, for example, with an expert assessment, survey or interview. Ideally,



a supportive foundation for Kaizen should be created before organising a Kaizen event.

Table 3 indicates that the higher the unit's self-assessment of success in the Kaizen phases, the better the units sustain performance improvements. Therefore, a high self-assessment of success in the Kaizen phases is associated with sustained long-term results after the Kaizen event. The finding contributes to RQ3 by suggesting that self-assessed success in different Kaizen phases can explain or predict the persistence or decline of long-term results. Based on the findings in table 3, units' success in Kaizen events can be improved by dedicating more effort to preparation, the Kaizen event itself and especially follow-up. Successful and sustained follow-up is essential to the success of Kaizen. The findings can motivate units to focus on these phases more. Preparation could be done with instructions from the Kaizen event's organiser. For example, by mapping the unit's processes in advance, thinking of a suitable performance metric and gathering suggestions for changes in the unit. Follow-up includes collecting and monitoring performance metrics and implementing and maintaining changes after the Kaizen event. Daily management, such as regular meetings, devoting working time for improvement activities and collecting improvement ideas has worked well in follow-up. Sufficient follow-up guides in continuing or adjusting Kaizen implementation. Ideally, a unit monitors the performance metrics continuously and evaluates performance improvements with at least three reports: baseline, improvements during or immediately after Kaizen and after several months or years.

This study is one of the first to use sustainability levels (adapted from Bateman and David⁸) to evaluate Kaizen improvements and explain their long-term sustainability in healthcare. Continuous improvement (level A) is a core idea of Lean and Kaizen, and therefore, it is important to separate from those achieving only one-off improvements (level B), which are commonly reported in the literature. Using sustainability levels in categorising units and communicating the types of improvements achieved can be practical both in research and management. Our results, in certain respects, align with the findings of healthcare Kaizen described in international literature.^{8 12–18 23} Therefore, our study may be transferable and relevant to Kaizen activities in other healthcare contexts, such as the USA. Both our study and other studies on healthcare Kaizen found that the reported results are primarily improvements in processes such as lengths of appointments and stays after operations.^{12–15} This indicates that monitoring performance metrics related to processes is a popular and relatively easy way to evaluate Kaizen. Processes include areas like access, efficiency, quality and operational effectiveness. Process improvements can likely cause improvements in other areas, such as patient experience, staff well-being, clinical outcome and effective resource use. These improvements were mentioned in this study's interviews but are more challenging to measure and thus are qualitative

improvements. Adaptation of the sustainability framework^{8 18} contributed significantly to this study as it was found suitable for Kaizen events in HUS with some adjustments.

Factors associated with successful Kaizen implementation in healthcare have similarities with the explanatory factors for the persistence of the results found in this study. The main similarities are culture and motivation, cooperation and communication and management support.^{3 16–19} The findings indicate that these kinds of sociocultural factors are the essence of Kaizen. Our findings of the explanatory factors for the decline also share common themes with the previous literature. In addition to the lack of sociocultural factors, difficulties in data collection and lack of clarity inhibited the success.^{22 23} A significant difference in the findings is that in our study, the lack of time for improvement activities and high workload were the main factors for the decline. However, this has not been emphasised in other studies. Environmental factors outside the organisation, such as the coronavirus pandemic and introducing a new electronic health record system, affected data collection, comparability of the results and implementation of changes. These environmental factors affected every unit in this study. Our findings support statements of Kaizen literature suggesting that Kaizen events can significantly contribute to the success of overall lean change in an organisation while maintaining long-term performance improvements is challenging.^{11 31} Kaizen's conflicting benefits are visible in the distribution of the sustainability levels of the units involved in this study. The conflicting results could hypothetically be reduced by selecting the most suitable units for Kaizen or improving Kaizen events.

Strengths and limitations of this study

Strengths of this study include the availability of data on 26 Kaizen events organised in a relatively standard way, which is rare in the field. This study mainly uses interviews of chief physicians. The data collected from 21 interviews showed signs of saturation indicating that the data is extensive. The data used to determine long-term results and their sustainability is based on monitoring variables at different points in time and triangulated, bringing credibility to the evaluation. The difficulty of collecting data and potential sources of bias in data collection and analysis create limitations for this study. For example, many of the events were organised years before the interview, so the interviewee may have had difficulties recalling details of the Kaizen event. Some units may have had pressure to give positive reports and interview responses about Kaizen. Eventually, the findings are the researcher's interpretation of the available data. The ability to answer RQ1 was limited due to the lack of documentation of immediate performance improvements available for this research. Even though monitoring performance metrics occurred during the Kaizen events, most data was unavailable for research purposes. The lack of documented performance metrics made it necessary to rely

on subjective reports from the interviews in some cases. It is difficult to determine to what extent the environmental factors affected the performance metrics of this study. Therefore, factors outside the scope of this study have affected the persistence or decline of performance improvements. Differences in healthcare systems and units should be considered when comparing findings.

Even though Kaizen events are primarily used to enhance the performance of a unit, it is important to organise them with sufficient documentation and data for systematic long-term evaluation of effectiveness. Future research could focus on investigating themes arising from the findings, such as: (1) Investigating whether selecting units that are likely to succeed based on the explanatory factors can increase the success of units. (2) Improving the quality of Kaizen activities by dedicating special effort to preparation and follow-up. (3) Predicting the persistence of performance improvements with self-assessment of success in preparation, Kaizen event and follow-up. (4) Finding out what performance metrics would be suitable for monitoring sociocultural improvements. (5) To investigate whether improvements in processes also improve performance in sociocultural metrics.

Conclusions

This study is one of the few explaining the long-term sustainability of performance improvements of Kaizen in healthcare. The findings indicate that Kaizen events can be worth organising even though long-term performance improvements are not guaranteed. Units that lack factors explaining the persistence of results can be identified, trained and guided to increase their chances of success. Chances of success may be increased with the proper selection of the units and by careful planning of the Kaizen events.

Acknowledgements We would like to express our gratitude to Tommi Jokiniemi and Ritva Jokela from the Lean unit of Helsinki University Hospital, who participated in guiding this project. We thank Helsinki University Hospital for its financial support during this project.

Contributors EH wrote the main manuscript text, and prepared tables 1–4 and supplementary files. Supervisors ER and PT provided guidance and comments throughout the project. ER and PT modified the text when necessary. All authors reviewed the manuscript. EH is responsible for the overall content as guarantor.

Funding This study is funded by Helsingin ja Uudenmaan Sairaanhoidopiiri (HUS). HUS paid a salary for the research work of EH. HUS paid Tutkimustie Oy for transcription work. The authors have no affiliations to Tutkimustie Oy. Open access funded by Helsinki University Library.

Competing interests EH: Salary from Helsingin ja Uudenmaan Sairaanhoidopiiri (HUS) for the research work. All other authors declare no conflicts of interest.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval The Research Ethics Committee of the Helsinki and Uusimaa Hospital District evaluated the study protocol and decided that this study is exempt, and that the ethics approval is not needed. The approval is not needed due to the provisions of the Medical Research Act (488/1999), General Data Protection Regulation (GDPR) Art. 6 (1.e) and Data Protection Act 4 §. GDPR Art. 6 (1.e) and Data Protection Act 4 § apply to this study and they have been followed. Medical Research Act (488/1999), does not apply to this study because this study is not medical research but management research based on the definition of medical

research in the act. Legislation of healthcare research in Finland differs from many other countries and there is no relevant legislation other than the Medical Research Act (488/1999), GDPR Art. 6 (1.e) and Data Protection Act 4 §. The Research Ethics Committee of the Helsinki and Uusimaa Hospital District's letter of inquiry about the need for an ethical review: 'In Finland, medical research involving invasive studies on humans must comply with the provisions of the Medical Research Act (488/1999). These provisions stipulate that the research must receive a favourable opinion from an ethics committee, if the research (a) will bring more information on health, the causes of diseases, symptoms, diagnostics, treatment, prevention, or the essence of diseases in general and (b) will involve interfering with the integrity of a human or human embryo or fetus. If these criteria are not met, there is no need for the ethical review by the HUS Regional Ethics Committees. Moreover, research based purely on documentation or registered materials does not need to be reviewed by the regional ethics committees.' The interview did not ask questions of a sensitive nature or, for example, related to the interviewee's state of health, nor questions that could have had negative effects on the interviewee's employment relationship if answered. Informed consent was obtained from all subjects. All methods were carried out in accordance with relevant guidelines and regulations.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. The corresponding author has the data used in this article. The data is not publicly archived. Permission from the parties involved should be obtained before possible data sharing. Data availability is checked with The Hospital District of Helsinki and Uusimaa (HUS) upon reasonable request. The data is in Finnish.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Erik Haapatalo <http://orcid.org/0000-0001-5708-8478>

Elina Reponen <http://orcid.org/0000-0003-2974-9683>

Paulus Torkki <http://orcid.org/0000-0002-1127-4205>

REFERENCES

- Suárez Barraza M, Ramis-Pujol J, Kerbache L. Thoughts on Kaizen and its evolution: three different perspectives and guiding principles. *Int J Lean Six Sigma* 2011;2:288–308.
- Brunet AP, New S. Kaizen in Japan: an empirical study. *Int J Operat Product Manag* 2003;23:1426–46.
- Bessant J, Caffyn S, Gilbert J, *et al*. Rediscovering continuous improvement. *Technovation* 1994;14:17–29.
- Healthcare kaizen. Engaging front-line staff in sustainable continuous improvements. 2012.
- Jacobson GH, McCoin NS, Lescallete R, *et al*. Kaizen: A method of process improvement in the emergency Department. *Acad Emerg Med* 2009;16:1341–9.
- Laraia. The Kaizen blitz. 1999.
- Melnik SA. Short-term action in pursuit of long-term improvements: introducing Kaizen events. *Product Invent Manag J* 1998;39:69–76.
- Bateman N, David A. Process improvement programmes: a model for assessing Sustainability. *Int J Operat Product Manage* 2002;22:515–26.
- Singh J, Singh H. 5 1 Kaizen philosophy: a review of literature Kaizen philosophy: a review of literature. *IUP J Operat Manage* 2009.
- Suárez-Barraza MF, Lingham T. Kaizen within Kaizen teams: continuous and process improvements in a Spanish municipality. *Asian Journal on Quality* 2008;9:1–21.



- 11 Glover W, Farris J, Van Aken E, *et al.* Kaizen event result Sustainability for lean enterprise transformation. *J Enterprise Transformat* 2013;3:136–60.
- 12 Sladen DP, Olund AP, Schmitt M, *et al.* Application of Kaizen principles to a large cochlear implant practice: A continuous quality improvement initiative at Mayo clinic. *Otol Neurotol* 2019;40:e592–9.
- 13 Iannettoni MD, Lynch WR, Parekh KR, *et al.* Kaizen method for Esophagectomy patients: improved quality control, outcomes, and decreased costs. *Ann Thorac Surg* 2011;91:1011–7;
- 14 Nino V, Claudio D, Valladares L, *et al.* An enhanced Kaizen event in a sterile processing Department of a rural hospital: A case study. *Int J Environ Res Public Health* 2020;17:8748.
- 15 Suárez-Barraza MF, Miguel-Davila JA. Kaizen-Kata, a problem-solving approach to public service health care in Mexico. A multiple-case study. *Int J Environ Res Public Health* 2020;17:3297.
- 16 Gonzalez-Aleu F. n.d. Conceptual framework to estimate continuous improvement project success in hospitals.
- 17 Gonzalez-Aleu F, Van Aken EM, Cross J, *et al.* Continuous improvement project within Kaizen: critical success factors in hospitals. *TQM* 2018;30:335–55.
- 18 Stelson P, Hille J, Eseonu C, *et al.* What drives continuous improvement project success in Healthcare *Int J Health Care Qual Assur* 2017;30:43–57.
- 19 Womack JP, Jones DT. Lean Thinking—Banish waste and create wealth in your corporation. *J Operat Res Soc* 1997;48:1148.
- 20 Glover WJ, Farris JA, Van Aken EM, *et al.* Critical success factors for the Sustainability of Kaizen event human resource outcomes: an empirical study. *Int J Product Econ* 2011;132:197–213.
- 21 Radnor ZJ, Holweg M, Waring J. Lean in Healthcare: the unfilled promise? *Soc Sci Med* 2012;74:364–71.
- 22 Lima RM, Dinis-Carvalho J, Souza TA, *et al.* Implementation of lean in health care environments: an update of systematic reviews. *IJLSS* 2021;12:399–431.
- 23 Mazzocato P, Stenfors-Hayes T, von Thiele Schwarz U, *et al.* Kaizen practice in Healthcare: a qualitative analysis of hospital employees' suggestions for improvement. *BMJ Open* 2016;6:e012256.
- 24 Gonzalez Aleu F, Van Aken EM. Continuous improvement projects: an authorship Bibliometric analysis. *Int J Health Care Qual Assur* 2017;30:467–76.
- 25 Aleu FG, Van Aken EM, Keathley-Herring H. *Bibliometric analysis of research design characteristics in continuous improvement projects in hospitals.*
- 26 Health, M.o.S.a.a. Health services. 2022. Available: <https://stm.fi/en/health-services>
- 27 HUS - about us. 2022. Available: <https://www.hus.fi/en/about-us>
- 28 Eisenhardt KM. Building theories from case study research. *Acad Manage Rev* 1989;14:532.
- 29 Patton MQ. Enhancing the quality and credibility of qualitative analysis. *Health Serv Res* 1999;34:1189–208.
- 30 Reponen E, Jokela R, Blodgett JC, *et al.* Validation of the lean Healthcare implementation self-assessment instrument (LHISI) in the Finnish Healthcare context. *BMC Health Serv Res* 2021;21:1289.
- 31 Doolen TL, Van Aken EM, Farris JA, *et al.* Kaizen events and organizational performance: a field study. *Int J Product Perform Manag* 2008;57:637–58.