

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<https://hdl.handle.net/2066/218674>

Please be advised that this information was generated on 2021-06-16 and may be subject to change.

Does a Long-Term Quality and Safety Curriculum for Health Care Professionals Improve Clinical Practice? An Evaluation of Quality Improvement Projects

Anne A. C. van Tuijl, MSc; Hiske Calsbeek, RN, PhD; Hub C. Wollersheim, MD, PhD;
Roland F. J. M. Laan, MD, PhD; Cornelia R. M. G. Fluit, MD, PhD; Petra J. van Gorp, MD, PhD

Introduction: Little is known about the effects on clinical practice of continuing education quality and safety curricula. The aim of this study is to gain insight into learning outcomes on the fourth level of the Kirkpatrick evaluation model for systematically deployed quality improvement projects performed by health care professionals during a Masters in Healthcare Quality and Safety in the Netherlands.

Methods: The researchers reviewed 35 projects led by health care professionals in 16 different hospitals to determine their scopes and effects. Afterward, professionals took an online survey to determine the extent of their project's sustainability and spread.

Results: Improving health care safety was the most prevalent quality dimension ($n = 11$, 31%). A positive change was measured by professionals for 64% ($n = 35$) of the primary outcomes. Statistical significance was measured in 19 (35%) of the outcomes, of which nine (47%) were found to have a statistically significant effect. A minority of professionals (17%) judged their project as sustained by the department, while some stated that the intervention (37%) or the results of the project (11%) had spread.

Discussion: Although most projects indicated an improvement in their primary outcomes, only a few resulted in statistically significant changes. Teaching professionals in using evaluation methods that take into account the complex context where these projects are performed and teaching them leadership skills is needed to reduce the likelihood of unmeasured outcomes. Analyzing learning experiences of professionals in performing the project is important to see what they learned from performing quality improvement projects, providing experiences that may lead to sustainable effects in future projects.

Keywords: continuing education in health care quality and safety, evaluating learning outcomes, Kirkpatrick evaluation model, quality improvement projects, effects on practice, sustainability and spread

DOI: 10.1097/CEH.0000000000000277

Improving quality and safety (Q&S) is a top priority for many health care systems.^{1,2} To prepare current and future health care professionals for their role in improving patient care, it is necessary to provide formal education on quality improvement.³⁻⁷ As a result, Q&S education has been introduced across the medical education continuum, from undergraduate programs to professional continuing education.⁶

A key factor for the success of Q&S curricula is the inclusion of learning principles and activities based on Kolb's experiential learning model.^{3,5,6,8,9} In experiential learning, the educational

process moves from the involvement of learners in new experiences to their reflection on that experience, after which they conceptualize and integrate their experience and implement their learning in practice.¹⁰ Performing workplace-related quality improvement projects (QIPs) is common and has been found to be an appropriate format for interpreting experiential learning in Q&S curricula.^{5,7,8}

QIPs, performed as an educational activity, must not only yield learning outcomes for the learner in terms of experiences, attitudes, knowledge, skills, and behaviors but should also have

Disclosures: The authors declare no conflicts of interest. This research study was carried out under the auspices of the Consortium Quality of Care of the Netherlands Federation of University Medical Centers (NFU), and the National Health Care Institute. The Dutch Association of Medical Education Ethical Review Boards approved this study (NERB number: 842).

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.jcehp.org).

Ms. van Tuijl: PhD candidate in the Department of IQ Healthcare, Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, the Netherlands. **Dr. Calsbeek:** Assistant professor of Quality Improvement Science and Junior Principle Lecturer in the Department of IQ Healthcare, Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, the Netherlands. **Dr. Wollersheim:** Associate professor of Quality Improvement Science in the Department of IQ Healthcare, Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, the Netherlands. **Dr. Laan:** Professor in Medical Education, and Director of the Radboudumc Health Academy of the Radboud University Medical Center, Nijmegen, the Netherlands. **Dr. Fluit:** Associate professor of Workplace Learning in Health Care, and Head of the Department for Research in Learning and Education of the Radboudumc Health Academy, Radboud University Medical Center, Nijmegen, the Netherlands. **Dr. van Gorp:** Internist in the Department of Internal Medicine and a Principle Lecturer at Radboud University Medical Center, Nijmegen, the Netherlands.

Correspondence: Anne A. C. van Tuijl, MSc, Radboud University Medical Center, Radboud Institute for Health Sciences, IQ Healthcare, Geert Grooteplein 21, P.O. Box 9101 (114), 6500 HB, Nijmegen, the Netherlands; e-mail: anne.vantuijl@radboudumc.nl.

Copyright © 2019 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of The Alliance for Continuing Education in the Health Professions, the Association for Hospital Medical Education, and the Society for Academic Continuing Medical Education. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

an impact on health care in terms of changes to clinical processes or patient outcomes. To achieve this, QIPs should be evaluated on the fourth level of the Kirkpatrick¹¹ evaluation model (results on practice). This level is comparable with level 5 (performance) and level 6 (patient health) of Moore's Expanded Outcome Framework often used to evaluate continuing medical education activities.¹² An earlier systematic review⁶ using the Kirkpatrick fourth evaluation level showed that residents' involvement in experiential QIPs as part of a Q&S curriculum frequently led to significant improvements in the processes of care and, although less commonly measured, patient outcomes. Scientific studies of the learning outcomes of health care professionals performing QIPs as part of a Q&S curriculum are limited; however, one systematic review³ identified five Q&S curricula targeting health care professionals, all of which were relatively short-term programs with a limited description of their key features. These curricula included an evaluation of learning outcomes to provide data on the Kirkpatrick levels of knowledge, attitude, and behavior but did not measure the effects on practice. There is therefore a gap in the research regarding the success of QIPs in clinical practice conducted by health care professionals as part of long-term continuing education Q&S curricula.³

Therefore, we used the Kirkpatrick fourth level to investigate the learning outcomes of QIPs led by health care professionals as part of a 2-year Q&S postinitial (second Master after master degree) Masters curriculum in the Netherlands. In this program, health care professionals are trained to become leaders in the Q&S of health care. The purpose of this study was to evaluate the projects performed by health care professionals on their scope, effects, sustainability, and spread, and to contribute to more knowledge and understanding about the performance of QIPs in continuing education Q&S curricula. Using these insights, Q&S curricula being part of continuing education can be improved to better prepare health care professionals for performing QIPs.

METHOD

Context

In 2014, the 2-year Masters "Quality and Safety in Patient Care" began in collaboration with the eight university medical centers in the Netherlands. The curriculum is based on a constructivist vision on learning, with principles from experiential,⁸ adult,^{13,14} lifelong,¹⁵ and interprofessional learning.¹⁶ During the program, health care professionals with an academic background are trained to become leaders in the evidence-based quality improvement of health care. This means that they are able to strengthen their interprofessional collaboration in a team; initiate, support, and perform a successful QIP with the use of scientific methods; and strongly contribute to the spread of quality improvement both within and outside their own department and organization.

During the Master program, health care professionals are enrolled in 12 interactive learning modules focusing on many aspects of improving the quality of health care and personal leadership development. The educational format consists of lectures by leading experts, group discussions, role plays, group and individual assignments, and support from a coach, a methodological advisor, and their peers. Health care professionals work intensively on their professional development as a leader in quality improvement through individual

reflection, coaching, and group reflection meetings. Together, these educational interventions should guarantee an optimal preparation for health care professionals to lead and perform a QIP at their workplace during their Masters program. To demonstrate professional learning and development, professionals use a portfolio where they among other things reflect on their own performance in their project. Figure 1 shows the core elements of the master program.

In the first year of the Masters, the health care professionals begin their project by developing and refining their ideas. They can determine their own QIP within the broad requirements that the aim of the project fits with the organization's mission. Subsequently, they write a project plan, in which it must be clear that the project is manageable in terms of time, scale, and sphere of influence. Health care professionals must also carefully describe how they will manage the project and their methodological approach. After the project plan has been approved by the examiners, the health care professionals lead their QIP and write a scientific thesis in the second year of their Masters program. In the Masters thesis, professionals scientifically describe the performance and outcomes of their QIP. The SQUIRE (Standards for Quality Improvement Reporting Excellence) guidelines¹⁷ are used as a framework for reporting about the QIP in the thesis.

Participants

During the study period, 2014 to 2018, a total of 46 health care professionals with different backgrounds comprised the first two cohorts of the Masters program. Table 1 shows characteristics of the health care professionals from cohorts 1 and 2.

Data Collection and Analyses

From January 2017 until July 2018, a study was conducted to determine the scope, effects, sustainability, and spread of the QIPs. All first versions of the theses, submitted to the examiners, were included. The Masters theses from 20 (80%) health care professionals from cohort 1 and 18 (85%) from cohort 2 were included, for a total of 38 theses. Ethical approval was obtained from Ethical Review Board of the Dutch Association of Medical Education (ERB-842). Health care professionals of the master program were informed about this study and participated voluntarily as far as anonymous assessment data were not yet available for evaluation.

A document analysis¹⁸ was performed to determine the scope and effect of each QIP. First, the main researcher AvT became familiar with the content by reading all theses. Subsequently, the abstracts of the theses were thoroughly read. When an abstract provided too little information about the scope and effect of the project, specific information was looked up in the thesis. If any uncertainty arose, the thesis was further analyzed together with a second researcher (HC or HW), both of whom have extensive expertise on the quality improvement of health care.

To determine the scope of each project, their aims were categorized into the six improvement domains outlined by the Institute of Medicine (IOM)¹⁹ (safety, timeliness, efficacy, efficiency, equity, and patient-centeredness), while the primary outcome measures (POMs) were categorized into the five outcome levels used by Verweij²⁰ (professional knowledge, professional behavior, professional attitude, patient experience, and patient clinical outcome). To determine the effects on the POMs of the QIPs (eg, percentage of pain registration), we used the original data that the professionals used in their thesis.

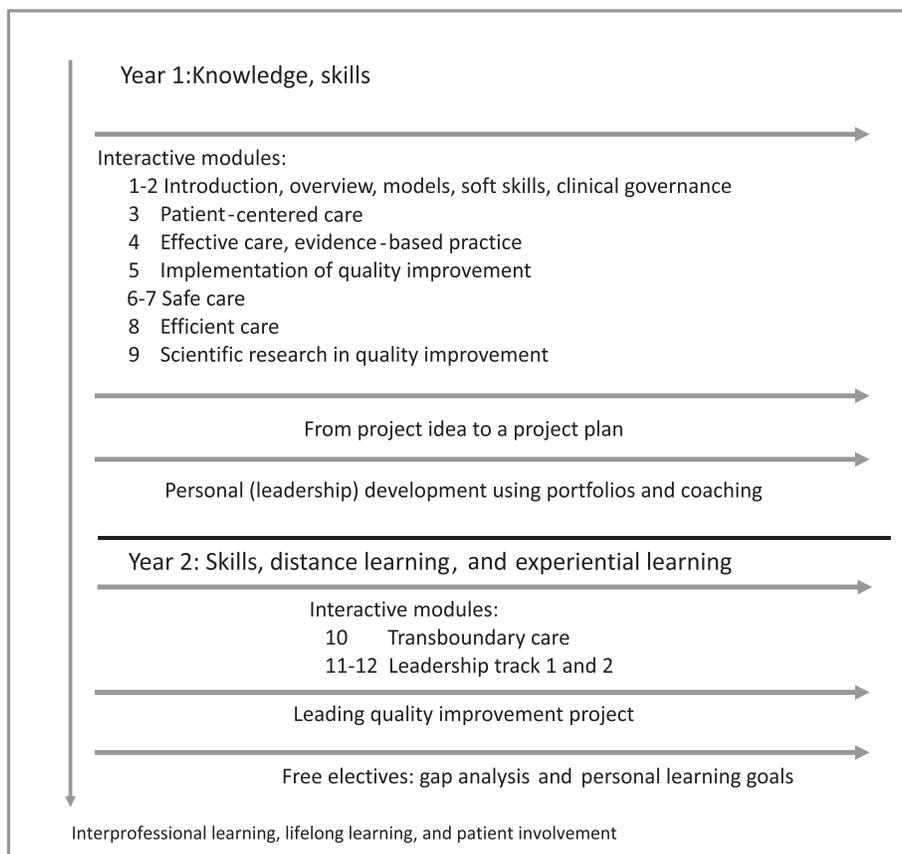


FIGURE 1. Core elements of the Dutch 2-year continuing education master program “Quality and Safety in Patient Care”

Because in every QIP unique outcomes, measurements, designs, and statistics were used, we calculated an effect size to be able to provide generic statements about the effects on the POMs. This effect size was expressed as a percentage or absolute difference between the baseline measurement before implementation and the postmeasurement after implementation. It was also possible that the postmeasurement was lacking. Based on the effect size, we defined three possible effects: no indication of improvement, an indication of improvement, and improvement was not

measured because the postmeasurement was lacking. Finally, we looked whether the professionals calculated statistical significance and whether the effect was found statically significant (SS) or not.

Data on the sustainability and spread of the projects were collected at least one year after the health care professionals had finalized their projects; therefore, only health care professionals from cohort 1 could be included. To determine the extent to which the QIPs were judged to have been sustained within the department, six subscales were taken from the short version of the Sustainability Instrument, the reliability and validity of which have been previously tested.²¹ Subscale routinization III was not included due to irrelevance of most of the items for these types of QIPs.

To determine the extent to which the new work method, developed as part of the QIP, had spread within and outside the organization, the Spread Instrument of Quality Improvement in Healthcare was used, the reliability and validity of which have previously been tested.²² The questionnaire consists of four subscales: spread of results, spread of work practices, action for results, and actions for the work practice.

Responses on the statements of both instruments were constructed using a five-point Likert scale, ranging from “*Totally agree*” to “*Totally disagree*,” including the option “*Don’t know*.” The answer categories from the subscales on the actions for results, and the actions for work practices were adapted into a dichotomous scale with the response options “*yes*” or “*no*,” including the option “*Don’t know*.” Higher scores indicated a greater sustainability of the project within

TABLE 1. Characteristics of Cohort 1 (N = 25) and Cohort 2 (N = 21) of the Masters Program

Characteristics	Cohort 1 (N = 25)	Cohort 2 (N = 21)
Background		
Physician	17 (68%)	8 (38%)
Nurse	5 (20%)	9 (43%)
Pharmacist	2 (8%)	1 (5%)
Allied health care professional	1 (4%)	1 (5%)
Health care scientist	—	1 (5%)
Health care jurist	—	1 (5%)
Involvement in (part-time) direct patient care		
Involved in (part-time) direct patient care	25 (100%)	19 (90%)
Master program		
Drop-out from the program	2 (8%)	1 (5%)
Not finished yet	3 (12%)	2 (10%)
Received Master title	20 (80%)	18 (85%)

the context in which it was performed, and a more effective spread of the QIP.

Because some items of the Slaghuis questionnaire were not applicable for all the QIPs in this study, or too complicated for the context of the QIPs, 13 questions were added regarding the sustainability and spread of the QIPs. These additional questions (“expert questions”) were developed in a formal meeting with two experts (HC and HW) in the field of quality improvement. The complete survey is available as **Supplemental Digital Content 1** (see **Appendix**, <http://links.lww.com/JCEHP/A68>). Participation in this survey was voluntary. Informed consent was implied by the overt action of completing the online questionnaire after reading the information letter.

For the statements from the Slaghuis questionnaires, the individual scale score for each health care professional was determined by calculating an estimated score for the statements belonging to each scale. The professionals who scored a four or five on their scale score, corresponding to “*Agree*” and “*Totally agree*,” respectively, were calculated as the “agree” percentage on that scale. Respondents who answered “*Don’t know*” on one or more statements were not included in this calculation. For the self-constructed statements, professionals who scored a four or five were also interpreted as those who “agree” with that statement, and their percentage was calculated. IBM SPSS Statistics for Windows, version 25.0 (Armonk, NY), was used for the data analysis.

RESULTS

Scope and Effects

Of the 38 QIPs, we excluded three theses (8%) because their aims did not fit into one of the six IOM domains. This resulted in a total of 35 theses describing a QIP in patient care, with 18 (49%) from cohort 1 and 17 (51%) from cohort 2. The projects were performed in 16 different hospitals across the Netherlands, of which eight (50%) were university hospitals and eight (50%) were teaching hospitals. Table 2 presents the QIPs aims, if a postimplementation measurement was performed by professionals, if the effect size indicated an improvement on the POMs (eg, percentage of pain registration or percentage of patients that is correctly identified), and if statistical significance was measured and an SS result was reached.

The safety domain was the most prevalent IOM dimension, incorporating the aims of 11 QIPs (31%). The aims of nine QIPs (26%) fitted into the domain of patient-centeredness, eight (23%) focused on effectiveness, five (14%) had efficiency-related aims, and two (6%) were focused on the provision of timely care. None of the project aims fit into the IOM equity domain.

Most QIPs ($n = 23$, 65%) focused on health professionals, while 12 (35%) projects had patient-related outcomes. Of the projects that focused on professionals, 19 (83%) involved behavioral change, one (4%) aimed to improve attitudes, one (4%) focused on improving both knowledge and attitude, one (4%) aimed to improve both behavior and knowledge, and one (4%) focused on improving both behavior and attitude. Of the patient-focused projects, eight (67%) were related to patient experiences, three (25%) were focused on changing clinical outcomes, and one (8%) aimed to change both patient experiences and clinical outcomes.

In total, 55 POMs were included in the QIPs, of which 64% ($n = 35$) indicated an improvement based on the calculated effect size using the original data from the professionals in their thesis (see column 3 in Table 2). No improvement was recorded for 15% of the POMs ($n = 8$), while for one POM (2%; unplanned readmissions), a negative effect was found. For 20% ($n = 11$) of the POMs, it appeared not to be possible to perform a postimplementation measurement during the Masters program due to time constraints. Statistical significance was measured by the professionals for 19 of the POMs (35%), of which nine (47%) were found to have SS improvements and nine (47%) did not. One POM (5%) showed an SS deterioration. For 36 (65%) of the POMs, statistical significance was not measured due to the use of small sample sizes, follow-up measurements to directly followed after the implementation of the intervention, and no follow-up measurements.

Sustainability and Spread

All 18 health care professionals from cohort 1 received the questionnaire, of whom 14 (78%) responded. Table 3 (five-point scales) and table 4 (dichotomous scales) lists the results on sustainability and spread.

A minority of health care professionals (17%) agreed that everybody in their department who works with the new interventions know how to perform it (routinization I), and that the department is able to adapt the interventions to variations in practice (routinization II). On the self-constructed question regarding sustainability, half of the professionals (50%) agreed that the use of their QIP had been sustained within the department. When questioned about each of the supporting conditions, fewer than half of the professionals agreed that all conditions necessary to sustain the intervention were effectively present.

A total of 37% of health care professionals stated that the intervention is also being used in other departments within their organization (spread work practice). In response to the self-constructed question, a majority agreed that their QIP has spread within (75%) and outside (62%) the organization.

Few health care professionals (11%) agreed that the results of their QIP have been applied in other departments (spread of results); however, the responses to the self-constructed questions contrasted with these results, as more than half of the professionals stated their results had spread within (86%) and outside (57%) the organization. The majority of the professionals organized discussions of progress ($n = 11$, 85%) and training ($n = 12$, 86%) to spread the new way of working as part of their QIP (activities for work practice), while the results of the QIPs were mostly spread by presentations and informal discussions ($n = 12$, 86%) (activities for results).

DISCUSSION

In this study, we found that QIPs performed by health care professionals during a two-year continuing education master curriculum in health care Q&S were most often focused on changing the behavior of colleague health care professionals. The most prevalent improvement theme was to make health care safer. The exploration of learning outcomes using the Kirkpatrick fourth evaluation level revealed that most primary outcomes of the QIPs led to indications of improvement based on the original data from the theses. The statistical significance

TABLE 2.
Categorization of the QIPs (*N* = 35) by Their Improvement Aim, Performance of a Postimplementation Measurement, Indication of Improvement, Performance of Test of Statistical Significance, and Statistical Significance

Improvement Aim	Postimplementation Measurement Performed by Professionals (No. of POMs)	Indication of Improvement (No. of POMs)*	Test of Statistical Significance Performed by Professionals (No. of POMs)	Statistical Significance (No. of POMs)
Improving pain registration	Yes	Yes	No	NM
Improving safety of preoperative plans	Yes	Yes	Yes	SS
Improving shared decision-making	Yes	Yes	Yes	SS
Reduction of deep postoperative wound infections	Yes	Yes	Yes	NSS
Improving quality of pediatric discharge	Yes (2)	Yes (2)	No (2)	NM (2)
Improving sleep in intensive care	Yes	Yes	No	NM
Improving patient involvement in the referral process from secondary to primary care	Yes (5)	Yes (1) No (4)	No (5)	NM (5)
Improving system awareness, communication skills, empathy, and professional identity	No (3)	NM (3)	No (3)	NM (3)
Reduction of intravenous medication errors	Yes	Yes	Yes	SS
Improving nurses as a role model in evidence-based practice	Yes (3)	Yes (3)	Yes (3)	NSS (3)
Improving wound care after hand operations	Yes (6)	Yes (6)	No (6)	NM (6)
Improving the use of the patient portal by patients	Yes	Yes	No	NM
Improving the use of policy regarding antithrombotic therapy	No	NM	No	NM
Improving patient identification and verification procedures	Yes	Yes	Yes	NSS
Improving the pharmaceutical consultation	Yes	No	No	NM
Improving patient satisfaction by the substitution of care to outreach clinics	No (2)	NM (2)	No (2)	NM (2)
Improving the medical handover in complex cases	Yes	Yes	No	NM
Improving the transfer of care responsibilities	No	No	No	NM
Improving the collection of patient experiences	Yes	Yes	No	NM
Improving the use of a pain behavior observation tool	Yes (3)	Yes (2) No (1)	Yes (3)	SS (2) NSS (1)
Improving the use of the intrahospital checklist when transporting patients	Yes	Yes	Yes	SS
Improving resuscitation and nontechnical team skills of nurses	Yes	Yes	Yes	SS
Reduction of medication administration errors	Yes	Yes	Yes	NSS
Reduction unnecessary laboratory testing	Yes	Yes	Yes	SS
Improving patient throughput from intensive care	Yes (2)	Yes (2)	Yes (1) No (1)	SS (1) NM (1)
Improving shared decision-making between parents and health care professionals	Yes	Yes	Yes	NSS
Improving the use of the Utrecht Symptoom Dagboek	Yes	No	No	NM
Improving participation in shared decision-making and patient information	No (2)	NM (2)	No (2)	NM (2)
Reduction of acetylcysteine and/or salbutamol without a strict clinical indication	Yes	Yes	No	NM
Reduction of the number of unplanned readmissions	Yes	No†	Yes	SS
Improving professional skills of nurses in using ultrasound	Yes	Yes	Yes	NSS
Improving shared decision-making	Yes	Yes	No	NM
Improving early detection of deterioration among children in general hospitals	Yes	Yes	No	NM
Improving hand hygiene compliance	No	NM	No	NM
Reduction of pain in children below the age of 10 years	No (2)	NM (2)	No (2)	NM (2)

NM indicates not measured; NSS, not statistically significant; QIPs, quality improvement projects; SS, statistically significant.

*Based on the effect size, we calculated (a percentage or absolute difference between the baseline measurement and the postimplementation measurement) using the original data from the professionals in their theses.

†An increase in the number of unplanned readmission after implementation of the intervention was found.

of the improvements was only measured by professionals in a minority of primary outcomes. According to the results of the Slaghuis questionnaires, a minority of health care professionals reported that their projects and projects' results had been sustained and spread; however, health care professionals were

more positive about this when responding to the additional constructed questions.

None of the QIPs' aims fitted into the equity domain of the IOM. A rationale for this finding could be that the Dutch health care system is recognized with a relatively low level of health

TABLE 3.
Sustainability and Spread of the QIPs (N = 14) Measured Using a Five-Point Likert Scale

	N	% Agree
Sustainability		
Routinization I (subscale)	12	16.7
Routinization II (subscale)	12	16.7
Your QIP is sustained within the department in which it was performed (SC)	12	50.0
Institutionalization of		
Skills (subscale)	6*	16.7
Documentation Materials (subscale)	8*	50.0
Practical Materials (subscale)	5*	20.0
Team reflection (subscale)	7*	28.6
Spread		
Spread work practices (subscale)	11	36.4
The new work practice has spread within your organization (SC)	14	76.9
The new work practice has spread outside the organization (SC)	14	61.5
Spread of results (subscale)	9	11.1
The results have spread within the organization (SC)	14	85.7
The results have spread outside the organization (SC)	14	57.1

QIPs indicates quality improvement projects; SC, self-constructed statements.

*Students who were eligible for these subscales.

disparities,²³ making the focus on this improvement domain less urgent. The system obliges everyone living in the Netherlands to purchase basic health insurance. Health insurers in turn are obliged to offer basic health insurance to everyone. Dutch citizens pay an income-dependent contribution for their insurance. Thereby, the Netherlands has a well-organized network of health care suppliers, and remote areas hardly exist.²³

Our study shows that the measurement of statistically significant effects performed by the health care professionals was difficult due to problems with sample sizes and the restricted period. The problems with sample sizes and the classical statistical techniques highlight the need for the teaching and use of different methods for analyzing and evaluating the outcomes of QIPs. These alternatives would be beneficial because the target population of the improvement project is part of a complex adaptive system, and because it is practically impossible to

control for context, the evaluation of QIPs using experimental methods such as a randomized control trial is not suitable.²⁴

A before/after design with interrupted time series analyses²⁵ or a controlled before/after design with statistical process control²⁶ are sound methodological approaches for improvement assessment.²⁷ Most of the health care professionals in this study indeed used a nonrandomized before/after measurement design to evaluate their projects, which does not fit with classical statistical analysis based on “time static” statistical tests, requiring large sample sizes. Also, using nonparametric tests that do not make stringent assumptions about the population and the data²⁸ is valuable for analyzing the statistical significant effect of the QIPs. With the use of these statistics incremental changes in small populations can be measured, and if significant, these changes are valuable because they provide information regarding the QIP’s potential sustainability and spread.

TABLE 4.
Sustainability and Spread of the QIPs (N = 14) Measured Using a Dichotomous Scale

	Yes Responses	No Responses
Sustainability		
The routine performance of the new work practice developed during the QIP needs to be supported by (SC)		
Skills	7	7
Documentation Materials	8	6
Practical Materials	5	9
Team Reflection	7	7
Spread		
Spread activities for work practices		
Discussions of progress	11	2
Courses	12	2
Guest visits	8	6
Spread activities for work practices		
Presentations	12	2
Reports	4	9
Informal discussions	12	2
Courses	7	5

QIPs indicates quality improvement projects; SC, self-constructed statements.

However, only using quasiexperimental methods makes an evaluation of effects susceptible to bias and, moreover, provides no understanding about the context-specific drivers behind implementation outcomes. Using mixed-method designs where quantitative and qualitative methods are combined provides a more comprehensive and richer understanding of the research issue than one approach alone. This triangulation of research methods facilitates to measure outcomes and to understand the process of implementation, to examine both the intervention content and context, and to compensate for one set of methods by the use of another set of methods.^{29,30} Using analytical techniques that are associated with realist evaluation are particularly useful for evaluating QIPs because these are able to adequately link mechanisms, context, and outcomes to explain and understand interventions.^{31,32} Using a realist evaluation method to evaluate a QIP helps to interpret the findings in light of other interventions and settings and gives answer to the question what works, for whom, and in what circumstances. Teaching health care professionals in Q&S curricula about these methods and designs to evaluate their QIP will hopefully reduce the likelihood of unmeasured improvement aims and a lack of evidence related to patient outcomes.

Based on the results of the effects of the QIPs, we also feel the need to pay more attention to leadership skills in improving the safety and quality of health care in these continuing education curricula. As previous research suggests, health care professionals' leadership in quality improvement plays a crucial role in the success of improvement projects.^{33–35} Competencies as organizational awareness, communication skills, and life-long learning behaviors³⁶ need to be taught in these curricula. In our Master program, we now try to put more emphasis on these leadership skills by using coaching and portfolios.

Only a minority of the outcomes of the QIPs were reported to be sustained in departments, possibly because the necessary supporting conditions, such as the availability of materials and documents, were suboptimal. A minority of professionals stated that the new work practice and the results of their QIPs had spread to other departments. The assessment of the sustainability and spread of the projects, as part of the official examination of the thesis, were among the poorest/worst-scored items. Surprisingly, the findings from the Slaghuys questionnaires contrast those of the additional expert-constructed questions, in which respondents seemed to be more positive about the sustainability and spread of their QIPs. This discrepancy could be attributed to professionals' overestimation of their own performance³⁷ or to the low content validity of these questions because we did not include a control for their validity. Future research with more focus on the correlation between the outcomes on the POMs and the sustainability and spread data will be needed to determine which project aims were not only achievable but more likely sustainable.

Our study has several limitations. First, this is a single educational program in one country involving two cohorts of health care professionals, which may limit its generalizability. The QIPs were performed in a wide range of hospitals across the Netherlands, however, and all academic hospitals were included. Another limitation is that we reviewed health care professionals' theses about their QIPs to obtain information about the effects on the POMs and did not observe or analyze these effects ourselves. To be more certain about the effects of the projects, triangulation

with observational and interview data is required. Third, we did not validate our additional expert-constructed questions regarding the sustainability and spread of the QIPs, collecting this information only from the health care professionals themselves, which adds uncertainty to our conclusions about the results of these questions. The inclusion of feedback from other people, such as coworkers or patients, would enhance our understanding of the effects of the QIPs. Finally, we only collected data about the short-term effects of the QIPs. The assessment of the long-term effects was beyond the scope of this study, but would help to elucidate whether the new working methods had been effectively deployed.

We have learned several important lessons from this study. The most important is that, although the success of the performed QIPs seems rather disappointing, these projects did not fail and still have value. The QIPs involve an intervention with learning outcomes in practice but are also learning processes in their own right. Using the fourth level of the Kirkpatrick model as a way of evaluating the learning outcomes, we were unable to gain insights into the learning experiences of the health care professionals. When the projects are defined as an experiential learning intervention involving the different phases of the learning cycle of Kolb, however, we can explore what health care professionals have learned from performing the project and how they reflect on these experiences. These reflections could bring about changes in their behavior or performance in this or future projects, which may lead to sustainable effects in practice.

We have also learned that, although sustainability and the spread of the projects may not be a realistic goal because of the 2-year period of the Masters, it is important that health care professionals learn about sustaining and spreading their project during the program. Using this knowledge, health care professionals can work on sustaining and spreading their QIP after finishing their Masters to make a significant and long-lasting effect on practice.

Another important lesson is that we should pay more attention in our Q&S curricula to the assessment and evaluation of the QIPs using methods that take into account the complex adaptive contexts where these projects are performed. Professionals in our Masters seem to be too focused on using the traditional before/after designs to find postimplementation effects. The question how, why, and in which context these are the outcomes seems to be inferior to them. The Masters program seems to maintain these classical views on outcome assessment by putting great emphasis on performing an effect evaluation. Relatively little attention is paid to the process evaluation of the project. The power of integrating effect evaluations and process evaluations by using mixed-method designs and principles from a realist evaluation remain relatively unknown to professionals. By strengthening the use and value of less classical designs and system-level evaluations, we hope to train professionals with the necessary skills and a broad vision on evaluating QIPs.

CONCLUSIONS

Although most projects resulted in an improvement on their POMs, only a few were able to measure whether the effects were statistically significant. To reduce the likelihood of unmeasured aims and a lack of evidence regarding patient outcomes, it is important to teach professionals in using less classical designs

and evaluation methods that take into account the complex contexts where these projects are performed. We also suggest to teach health care professionals' leadership skills such a lifelong learning behaviors and communication skills. Despite the limited sustained effects, the learning experiences of the health care professional performing the project have a major value in itself. Further studies, including the analysis of less classical and realist evaluations of QIPs and professionals' learning experiences of the project through Kolb's learning cycle, are needed to provide a broader and more in depth view of the learning outcomes.

Lessons for Practice

- QIPs performed by health care professionals in the context of a continuing education Masters indicate improvements on practice but only a few showed statistically significant effects
- We should pay more attention in our Q&S curricula to the evaluation of the projects using methods that take into account the complex adaptive contexts where QIPs are performed
- We suggest to put emphasis on teaching health care professional leadership skills in Q&S improvement
- Further study is required to analyze the outcomes of QIPs using less classical designs and evaluation methods and to determine learning experiences of health care professionals in performing QIPs during continuing educational programs

ACKNOWLEDGMENTS

The authors acknowledge the health care professionals from cohorts 1 and 2 of the Master program.

REFERENCES

1. Ferlie EB, Shortell SM. Improving the quality of health care in the United Kingdom and the United States: a framework for change. *Milbank Q*. 2001;79:281–315.
2. Scott I. What are the most effective strategies for improving quality and safety of health care? *Intern Med J*. 2009;39:389–400.
3. Boonyasai RT, Windish DM, Chakraborti C, et al. Effectiveness of teaching quality improvement to clinicians: a systematic review. *JAMA*. 2007;298:1023–1037.
4. Kaminski GM, Britto MT, Schoettler PJ, et al. Developing capable quality improvement leaders. *BMJ Qual Saf*. 2012;21:903–911.
5. Ogrinc G, Headrick LA, Mutha S, et al. A framework for teaching medical students and residents about practice-based learning and improvement, synthesized from a literature review. *Acad Med*. 2003;78:748–756.
6. Wong BM, Etchells EE, Kuper A, et al. Teaching quality improvement and patient safety to trainees: a systematic review. *Acad Med*. 2010;85:1425–1439.
7. Zenlea IS, Billett A, Hazen M, et al. Trainee and program director perceptions of quality improvement and patient safety education: preparing for the next accreditation system. *Clin Pediatr (Phila)*. 2014;53:1248–1254.
8. Goldman J, Kuper A, Wong BM. How theory can inform our understanding of experiential learning in quality improvement education. *Acad Med*. 2018;93:1784–1790.
9. Batalden P, Davidoff F. Teaching quality improvement: the devil is in the details. *JAMA*. 2007;298:1059–1061.
10. Kolb DA. *Experiential Learning: Experience as the Source of Learning and Development*, 2nd ed. Upper Saddle River, NJ: Pearson Education, Inc; 2015.
11. Kirkpatrick DL, Kirkpatrick JD. *Evaluating training programs. The four levels*. 3rd ed. San Francisco, CA: Berrett-Koehler; 2006.
12. Moore DE Jr, Green JS, Gallis HA. Achieving desired results and improved outcomes: integrating planning and assessment throughout learning activities. *J Contin Educ Health Prof*. 2009;29:1–15.
13. Brookfield SD. *Understanding and Facilitating Adult Learning: A Comprehensive Analysis of Principles and Effective Practices*. Milton Keynes, UK: Open University Press; 1986.
14. Collins J. Education techniques for lifelong learning: principles of adult learning. *Radiographics*. 2004;24:1483–1489.
15. Jennett PA, Swanson RW. Lifelong, self-directed learning: why physicians and educators should be interested. *J Contin Educ Health Prof*. 1994;14:69–74.
16. Thistlethwaite J, Moran M; Education WHOSGoL, Practice C. Learning outcomes for interprofessional education (IPE): literature review and synthesis. *J Interprof Care*. 2010;24:503–513.
17. Davidoff F, Batalden P, Stevens D, et al. Publication guidelines for quality improvement studies in health care: evolution of the SQUIRE project. *BMJ*. 2009;338:a3152.
18. Bowen GA. Document analysis as a qualitative research method. *Qual Res J*. 2009;9:27–40.
19. Berwick DM. A user's manual for the IOM's "Quality Chasm" report. *Health Aff*. 2002;21:80–90.
20. Verweij LM, Baines R, Friele RD, Wagner C. Implementation of effective interventions requires knowledge of the practice, attention to context and flexibility: evaluation of the ZonMw subprogramme Implementation [in Dutch: Implementatie van doelmatige interventies vraagt kennis van de praktijk, aandacht voor de context en flexibiliteit: evaluatie van het ZonMw deelprogramma Implementatie]. Utrecht: 2015. Available at <https://www.nivel.nl/sites/default/files/bestanden/Rapport-deelprogramma-implementatie.pdf>. Accessed 2019.
21. Slaghuis SS, Strating MM, Bal RA, et al. A framework and a measurement instrument for sustainability of work practices in long-term care. *BMC Health Serv Res*. 2011;11:314.
22. Slaghuis SS, Strating MM, Bal RA, et al. A measurement instrument for spread of quality improvement in healthcare. *Int J Qual Health Care*. 2013;25:125–131.
23. Kroneman M, Boerma W, van den Berg M, et al. Netherlands: health system review. *Health Syst Transit*. 2016;18:1–240.
24. Øvretveit J, Gustafson D. Evaluation of quality improvement programmes. *BMJ Qual Saf*. 2002;11:270–275.
25. Penfold RB, Zhang F. Use of interrupted time series analysis in evaluating health care quality improvements. *Acad Pediatr*. 2013;13:S38–S44.
26. Benneyan JC, Lloyd RC, Plsek PE. Statistical process control as a tool for research and healthcare improvement. *Qual Saf Health Care*. 2003;12:458–464.
27. Fretheim A, Tomic O. Statistical process control and interrupted time series: a golden opportunity for impact evaluation in quality improvement. *BMJ Qual Saf*. 2015;24:748–752.
28. Pett MA. *Nonparametric Statistics for Health Care Research: Statistics for Small Samples and Unusual Distributions*, 2nd ed. Newbury Park, CA: SAGE Publications; 2015.
29. Palinkas LA, Aarons GA, Horwitz S, et al. Mixed method designs in implementation research. *Adm Policy Ment Health*. 2011;38:44–53.
30. Aarons GA, Fettes DL, Sommerfeld DH, et al. Mixed methods for implementation research: application to evidence-based practice implementation and staff turnover in community-based organizations providing child welfare services. *Child Maltreat*. 2012;17:67–79.
31. Nurjono M, Shrestha P, Lee A, et al. Realist evaluation of a complex integrated care programme: protocol for a mixed methods study. *BMJ Open*. 2018;8:e017111.
32. Pawson R, Tilley N. *Realist Evaluation*. London, UK: SAGE Publications Ltd; 1997.
33. Kaplan HC, Provost LP, Froehle CM, et al. The Model for Understanding Success in Quality (MUSIQ): building a theory of context in healthcare quality improvement. *BMJ Qual Saf*. 2012;21:13–20.
34. Shah P, Cross V, Sii F. Sailing a safe ship: improving patient safety by enhancing the leadership skills of new consultant specialist surgeons. *J Contin Educ Health Prof*. 2013;33:190–200.
35. Donaldson LJ. Safe high quality health care: investing in tomorrow's leaders. *Qual Health Care*. 2001;10(suppl 2):ii8–12.
36. Garman A, Scribner L. Leading for quality in healthcare: development and validation of a competency model. *J Healthc Manag*. 2011;56:373–384.
37. Gude WT, Roos-Blom MJ, van der Veer SN, et al. Health professionals' perceptions about their clinical performance and the influence of audit and feedback on their intentions to improve practice: a theory-based study in Dutch intensive care units. *Implement Sci*. 2018;13:33.