Effect of Metrics-Based Simulation Training to Proficiency on Procedure Quality and Errors Among Novice Cardiac Device Implanters
The IMPROF Randomized Trial

Jorio Mascheroni, MSc; Martin Stockburger, MD; Ashish Patwala, MD; Archana Rao, MD; Hartwig Retzlaff, MSc; Christophe Garweg, MD, PhD; Tom Verbelen, MD, PhD; Anthony G. Gallagher, PhD, DSc

Abstract

IMPORTANCE In cardiac device implant training, there is no common system to objectively assess trainees’ ability to perform tasks at predetermined performance levels before in vivo practice; therefore, patients are potentially exposed to risks related to operators’ early learning curve.

OBJECTIVE To assess the effect on implant performance quality of novel metrics-based simulation training to proficiency (proficiency-based progression [PBP]) vs traditional simulation-based training (SBT).

DESIGN, SETTING, AND PARTICIPANTS In this prospective randomized trial, conducted between March 8, 2022 and November 24, 2022, novice implanters were randomized (blinded) 1:1 to participate in an SBT curriculum (procedural knowledge e-learning and in-person simulation training) at an international skills training center, with proficiency demonstration requirements at each training stage for advancing (PBP approach) or without the requirements. Ultimately, trainees performed a cardiac resynchronization therapy (CRT) implant using virtual reality simulation. The procedure was video-recorded and subsequently scored using previously validated metrics by 2 independent assessors blinded to group. Physicians who had already implanted more than 20 pacemakers or defibrillators and fewer than 200 CRT systems as the first operator were eligible. Thirty-two implanters from 10 countries voluntarily enrolled in the training program and were randomized; 30 (15 per group) started and completed training. Data analysis was performed from November 27 to December 22, 2022.

INTERVENTION Training with PBP vs SBT.

MAIN OUTCOME AND MEASURES The primary outcome comprised 4 objectively assessed performance metrics derived from the video-recordings: number of procedural steps completed, errors, critical errors, and all errors combined.

RESULTS Baseline experience of the 30 participants (19 [63%] male; mean [SD] number of years in implant practice, 2.0 [1.8]; median [IQR] number of implanted pacemakers or defibrillators, 47.5 [30.0-115.0]; median [IQR] number of implanted CRT systems, 3.0 [1.25-10.0]) was similar between study groups. Compared with the SBT group, the PBP group completed 27% more procedural steps (median [IQR], 31 [30-32] vs 24 [22-27]; P < .001) and made 73% fewer errors (median [IQR], 2 [1-3] vs 7 [5-8]; P < .001), 84% fewer critical errors (median [IQR], 1 [0-1] vs 3 [3-5]; P < .001), and 77% fewer all errors combined (errors plus critical errors) (median [IQR], 3 [1-3] vs 11 [8-12]; P < .001). One of the 15 PBP trainees (93%) demonstrated the predefined target performance level vs 0 of the 15 SBT trainees.

(continued)
CONCLUSIONS AND RELEVANCE  In this randomized trial, the PBP approach to novice implanter training generated superior objectively assessed performance vs SBT. If implemented broadly and systematically, PBP training may ensure safe and effective performance standards before trainees proceed to (supervised) in vivo practice. Future studies are needed to verify implications on procedure-related patient complications.

TRIAL REGISTRATION  ClinicalTrials.gov Identifier: NCT05952908

Introduction

Every year, more than 1 million patients affected by cardiac rhythm disorders receive a cardiac implantable electronic device (CIED) in Europe and North America alone. 1-3 Although the benefits of device therapy largely outweigh risks in eligible patients, short-term and long-term complications represent a problem even in established treatments, such as transvenous permanent pacing, for which published 60-day and 3-year adverse event rates in new implants reach 12.4% and 18.3%, respectively.4,5 Evidence shows that events are often underreported in routine clinical practice.5-7 Complications related to CIEDs affect patient quality of life,8,9 considerably increase healthcare costs,5,9,10 and are associated with a higher risk of mortality.9,11 Most of the early complications in cardiac device procedures (eg, pneumothorax, pocket hematoma, infection, cardiac perforation, and lead dislodgement) derive from the surgical or implant technique applied rather than technology malfunction.4,5,12 The incidence is higher at the beginning of an operator’s learning curve, regardless of whether it is a junior implanter learning conventional therapies or an experienced implanter learning a new skill or technology.13-15 Despite the risks, procedure training for novice operators usually happens in vivo, on live patients, from the outset. The taught implant techniques often differ among teaching centers, and trainees’ skill assessment mostly relies on supervisors’ individual judgment. Acquisition of competence is classically associated with the number of cases performed and time spent in training16-19; after specialty training, operators’ skills are very rarely reassessed (or reinforced via training where needed) during the career span.

During the past 20 years, the apprenticeship model for medical education established by William Halsted20 more than a century ago has faced unprecedented challenges.21,22 Increasing clinical specialization, mandated reduced work hours for trainees, increased demands for operating room efficiency, and quantitatively demonstrated risk of medical errors significantly challenge current approaches to training. Trainees’ supervised and systematic exposure to presentations of clinical situations and tasks and access to unconstrained expert coaching have deteriorated over time. The system is overstretched. Nevertheless, patients should still be guaranteed a measurable, assured, and consistent quality of treatment, irrespective of physicians’ expertise. Discussions about how to reform medical education are ongoing across the globe,19,21 but a true paradigm shift has yet to happen. To ensure an objective, predetermined proficiency level is reached by a trainee before in vivo operations, training requires the following: (1) a specific system of measurement for performance (metrics) to be developed and adopted as a standard, (2) a risk-free environment to facilitate initial skills learning, and (3) faculty’s time and dedication for coaching to proficiency. This approach to training is known as proficiency-based progression (PBP),23-25 which builds on mastery learning26-32 and deliberate practice33-35 instructional methods. Published results have shown that PBP trainees perform approximately 60% better than their conventionally trained peers in various domains of procedural medicine,36 but the method has not yet been applied to cardiac interventional training.37 Intraoperative PBP metrics (procedural steps, errors, and critical errors) for CIED training have been reported for a reference implant procedure,38 derived from and with agreement by expert practitioners, comprehensively encapsulating performance of transvenous pacemaker, implantable
cardioverter defibrillator (ICD), and cardiac resynchronization therapy (CRT) system implantations. Each metric was written using unambiguous operational definitions (rather than descriptions) so that it could be objectively scored as either occurring or not occurring (yes/no) by independent raters with a high degree of reliability. The metrics have also demonstrated the ability to discriminate reliably among different skill levels. To investigate the value of the PBP approach in CIED training for novice implanters, we hypothesized that a novel simulation training curriculum based on the PBP method would be superior to a traditional simulation-based training (SBT) curriculum. We sought to evaluate the training effect on CIED implanters’ performance in a prospective randomized trial.

Methods

Study Design

The IMPROF (Implant Proficiency) study was an international, prospective, double-blinded (trainees and assessors) randomized (1:1) trial conducted at KULeuven, Belgium, in 2022. The study was approved by the local ethics committee, and written informed consent was obtained from all participants. Data were collected at a skill center in Tolochenaz, Switzerland, from March 31, 2022, to November 24, 2022. This study followed the Consolidated Standards of Reporting Trials (CONSORT) reporting guidelines. This trial was registered retrospectively at the request of the editors. The trial protocol is available in Supplement 1.

Participants

Participants were CIED-implanting physicians from Europe and Israel in the early stages of their careers who enrolled in a peer-to-peer simulation training curriculum that focused on CIED implantation techniques, as an optional supplement to their academic or institutional education. Inclusion criteria are illustrated in Figure 1.

Figure 1. Inclusion Criteria and Study and Training Flowchart

<table>
<thead>
<tr>
<th>Inclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physicians actively practicing CIED implantation at time of enrollment</td>
</tr>
<tr>
<td>Minimum 20 pacemaker or ICD systems previously implanted as first operator</td>
</tr>
<tr>
<td>Minimum 3 CRT systems previously implanted at least as second operator</td>
</tr>
<tr>
<td>Familiarity with English language (written and spoken)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Having previously implanted ≥200 CRT systems as first operator</td>
</tr>
</tbody>
</table>

CIED indicates cardiac implantable electronic device; CRT, cardiac resynchronization therapy; ICD, implantable cardioverter defibrillator; PBP, proficiency-based progression; SBT, simulation-based training.
Interventions

After enrollment, participants were randomized 1:1 to follow 1 of 2 alternative training curricula covering a skin-to-skin, triple-chamber (ie, CRT) system implantation: SBT or PBP. Trainees were unaware of the type of simulation training received. Content, tools and resources, and schedule were the same for both curricula, which comprised 2 mandatory training modules: (1) an e-learning component illustrating the reference implant procedure, taking approximately 4 to 6 hours and concluded by a summative knowledge test; and (2) a peer-to-peer simulation-based component, consisting of hands-on implant coaching, lasting 2 days. The only difference between the 2 curricula concerned the way trainees advanced throughout the stages of training. Trainees in the PBP group were required to demonstrate a predefined level of performance (proficiency benchmark), quantitatively defined using previously tested metrics and verified by the instructors, to complete a training module and advance. Face, content, and construct validity had been previously established for the application of these metrics to the same purpose, context, and learner profile. The proficiency benchmarks had been calculated earlier as the mean of the scores obtained by experienced implanters performing the same assessment. Trainees in the SBT group instead were not given pass/fail thresholds. The study and training flowchart is illustrated in Figure 1.

The implant skills training was delivered by senior high-volume implanters all at consultant level who were working in Western European countries; they all had more than 10 years of device implant experience, were accustomed to coaching junior operators in their clinical practice, and had previously taught in more than 10 simulation training programs. The faculty had access to the e-learning module. Before the study began, the PBP group’s faculty were trained on how to apply the PBP method, whereas the SBT group’s faculty were not exposed to it. An instructor could teach only in 1 study arm to avoid bias related to PBP familiarity. The SBT faculty coached trainees and judged their progression based on general expert consensus on optimal implantation techniques as well as their long-term implant and teaching experience. In contrast, the PBP faculty coached and judged progression strictly based on the performance metrics and the proficiency benchmark. The skills training ended when a trainee was considered ready (within the maximum available time for practice) for a summative skills assessment, a decision based on the instructor’s personal judgment in the SBT group and on the demonstration of the quantitatively defined proficiency benchmark in the PBP group. For the PBP group, the proficiency benchmarks were a minimum of 83% correct answers in the online assessment for knowledge training and a maximum of 5 total errors and critical errors combined, with a maximum of 2 critical errors, in the simulated implantation of the 3 CRT system leads (right ventricular, right atrial, left ventricular) for skills training.

The summative skills assessment consisted of a solo performance of a 3-lead system implantation on the same full-physics virtual reality (VR) simulator used during coaching (CathLabVR, CAE Healthcare) and described previously. The given VR assessment case was a straightforward, uncomplicated patient, the same for every trainee, and had never been attempted before. The VR simulation performances were video-recorded in an anonymous fashion (participants were not recognizable) using a multicamera system connected to a laptop computer running OBS Studio software, version 27.2.3 (OBS Project), which simultaneously acquired and combined live streaming sources from 3 iPhones (Apple Inc) pointing at fluoroscopy, the patient’s electrocardiogram, and the operator’s hands.

Participation in the study ended with departure from the skill center. Before leaving, trainees were asked to fill in a training evaluation form to assess their level of satisfaction with the curriculum.

Outcomes

The primary outcome comprised 4 performance metrics, objectively assessed from each trainee’s video-recording: number of procedural steps completed, errors, critical errors (independent and scored separately), and errors all combined (all deviations from optimal performance). The primary analysis compared the performances of the SBT and PBP groups for each of the 4 variables separately. The prespecified secondary outcomes and comparisons between the 2 groups were as...
follows: (1) procedural measures derived from the VR simulator (procedure time, fluoroscopy time, cineradiography acquisition time, and amount of contrast media injected); (2) score (expressed as percentage of correct answers) obtained by trainees in the online test; (3) number of trainees per study group demonstrating the proficiency benchmark in the video-recorded performance; and (4) participants' satisfaction derived from the training evaluation forms.

Outcomes assessment started when all participants completed the study. All the videos were reviewed and scored by 2 independent assessors with more than 20 years of experience in CIED therapy, blinded to the groups and distinct from the teaching faculty, using the performance metrics. The number of scorable metrics for a given performance depended on the implant tools the operator chose (e.g., implanting active vs passive fixation leads). The maximum possible number of scorable procedure metrics for the VR performance in this study was 52, the ones best differentiating among levels of performance in the construct validity study.39 For each video, the assessors' scores were eventually compared, and the interrater reliability, expressed as the percentage of agreement between 0 (no agreement) and 1 (perfect agreement), was calculated.

**Sample Size, Randomization, and Blinding**
The sample size calculations were based on the results of the metrics construct validity study39 and of the transfer of training study.40 It was hypothesized that for the SBT group a mean (SD) of 15.6 (5.9) procedural errors would be observed on their video-recorded cases.39 For the PBP group, an error reduction of 42% was hypothesized.40 Therefore, with an α = .05 and a β = 0.20, it was determined that at least 14 individuals were required in each group to demonstrate statistical significance.

Participants were randomized at enrollment to follow 1 of the 2 alternative training curricula (SBT or PBP). The training curriculum manager generated a set of random numbers using online software (Research Randomizer, version 4.0, Social Psychology Network) and sequentially revealed them when a training application arrived. Participants were blinded to their training group, which was possible because the structure, content, and agenda of the 2 alternative training curricula were the same. Assessors of the video-recordings were blinded to the trainees and their training group, which was possible because the videos were recorded anonymously.

**Statistical Analysis**
Data analysis was performed from November 27 to December 22, 2022. Differences between groups for each primary and secondary outcome variable were individually tested for statistical significance. A Mann-Whitney U test $P < .05$ (2-sided) was considered statistically significant for each comparison. Data were analyzed with SPSS software, version 27 (IBM Inc).

**Results**
**Participants and Baseline Data**
Thirty-two novice CIED implanters were enrolled and randomized in the study. One participant per group could not participate, leaving 30 participants (19 [63%] male and 11 female [37%]; mean [SD] number of years in implant practice, 2.0 [1.8]; median [IQR] number of implanted pacemakers or defibrillators, 47.5 [30.0-115.0]; median [IQR] number of implanted CRT systems, 3.0 [1.25-10.0]).

Fifteen trainees per group received training, completed the study, and were analyzed for primary and secondary outcomes on an intention-to-treat basis. The study flow diagram is presented in Figure 2. Participants were from Austria (n = 2), Belgium (n = 1), Germany (n = 6), Greece (n = 2), Israel (n = 1), Italy (n = 5), the Netherlands (n = 2), Spain (n = 3), Switzerland (n = 2), and the United Kingdom (n = 6). Participants' baseline characteristics by group are given in Table 1; background implant experience was similar between the study groups.
Outcomes

Thirty videos (mean [SD] length, 34.7 [15.7] minutes) of the 3-lead VR implant procedure were independently scored by the 2 assessors. The mean number of metrics was 50.6 (95% CI, 50.3-51.0). All the metrics associated with a chosen implant approach were scored (no blanks). The mean interrater reliability between the assessors in the scoring was 0.94 (95% CI, 0.93-0.96).

The performances of the SBT and the PBP groups were compared for each variable measured. The primary outcomes are shown in Figure 3. The objective assessment using validated metrics showed that PBP trainees performed consistently better than the SBT trainees in the 3-lead implantation. Compared with the SBT group, the PBP group completed 27% more procedural steps and made 73% fewer errors, 84% fewer critical errors, and 77% fewer errors all combined (errors plus critical errors). All the differences were statistically significant, and the effect size was large (Table 2).

Procedure, fluoroscopy, cineradiography times, and contrast media injected (secondary outcomes) are reported in Table 2. The differences between groups for these measures were not statistically significant, and the observed effect size was generally small.

In the online summative knowledge test, the PBP group’s score was 32% higher than SBT group’s score (median [IQR], 93.3 [90.0-95.5] vs 70.8 [61.7-76.7]; P < .001). All the PBP trainees demonstrated the proficiency benchmark (as required by their study group); all the SBT trainees’ scores resulted below that benchmark (no threshold was given to the SBT group).

In the objectively assessed video-recorded implant performance, 14 of the 15 trainees (93%) in the PBP group but none in the SBT group demonstrated the proficiency benchmark. In the anonymous training evaluation forms, the level of satisfaction was similar in both groups. All 30 trainees agreed or strongly agreed that the overall quality of the training was excellent and rated their faculty excellent (93%) or good; on a scale of 1 to 10, the mean (SD) probability that they would recommend the curriculum to a peer was 9.6 (0.6) for the traditional simulation-based group and 9.6 (0.8) for the proficiency-based progression group.

---

**Table 1. Baseline Characteristics of the Study Participants**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>SBT group (n = 15)</th>
<th>PBP group (n = 15)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex, No. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>11 (73)</td>
<td>8 (53)</td>
</tr>
<tr>
<td>Female</td>
<td>4 (27)</td>
<td>7 (47)</td>
</tr>
<tr>
<td>Time in device implant practice after specialty training, mean (SD), y</td>
<td>1.9 (2.1)</td>
<td>2.1 (1.6)</td>
</tr>
<tr>
<td>No. of pacemaker or ICD implants as first operator, median (IQR)</td>
<td>50 (35-125)</td>
<td>45 (30-110)</td>
</tr>
<tr>
<td>No. of CRT implants as first operator, median (IQR)</td>
<td>3 (1-5)</td>
<td>5 (2-15)</td>
</tr>
</tbody>
</table>

Abbreviations: CRT, cardiac resynchronization therapy; ICD, implantable cardioverter defibrillator; PBP, proficiency-based progression; SBT, simulation-based training.
Discussion

In this study, we sought to evaluate the effect of a PBP simulation training approach on novice cardiac device implanters’ performances compared with an SBT approach. Virtual reality 3-lead CIED implantations video-recorded at the end of the training were objectively assessed using validated...
performance metrics. The metrics scores showed that the PBP-trained group completed significantly more procedure steps and made significantly fewer objectively assessed procedural errors (ie, errors, critical errors, and errors all combined) than the SBT group. The high interrater reliability levels also confirmed that such metrics could be scored with a high degree of reliability. Similar to a previous study, differences in procedural measures (ie, time, duration of fluoroscopy, and amount of contrast agent used) derived from the VR simulator were not statistically significant, and our results substantiate that they are weak indicators of skill.

The metrics characterize an optimal implant performance, and they were previously demonstrated to reliably distinguish among operator skill levels. Therefore, the difference in the group scores reported here suggest that the training provided to the PBP study group was more effective in enhancing operators’ skills than SBT. Such difference can reasonably be attributed to the PBP method, considering that our study minimized the differences in the training curriculum between the 2 groups by design, that participants’ baseline characteristics were similar, and that simulation guaranteed uniform conditions to every operator for the purpose of comparing skills objectively, irrespective of patient anatomies.

The PBP approach appeared also more efficient than SBT: all PBP trainees except 1 demonstrated the proficiency benchmark at the end of the instruction compared with none in the SBT group. Efficiency may derive from 3 components: (1) systematicity of the PBP instruction, (2) procedural knowledge consolidated before the skills training, and (3) objective, explicit, and formative performance feedback through metrics during coaching.

The e-learning component plays an important role in the overall performance result. Every student completed the online lessons and took the summative test: the different scores in the groups may indicate that the PBP students dedicated more time and attention to their e-learning, probably because an explicit target performance (proficiency benchmark) was required from them before training progression.

To our knowledge, our study is the first quantification of the effect of PBP training in the cardiac interventional workspace. Similar studies conducted in other interventional disciplines have shown that the PBP training approach generates significantly better (ie, 60%) procedural performance compared with traditional training, affecting, in particular, objectively assessed performance errors. Likewise, our study reports a significant reduction in performance errors (ie, 77%). We chose to compare the PBP approach with a robust, although traditionally delivered, SBT approach, rather than with a traditional apprenticeship training model. The differences in performances may have been even larger in the latter case, as previously reported. Unfortunately, SBT before in vivo practice still represents more the exception than the rule in current academic and institutional training for CIED implantation. The choice of training method also affects skills retention, because clinical skills acquired to high proficiency standards are more likely to be retained than those acquired through traditional training. A final consideration concerns the perceived quality of a training program. In both study groups, participants provided excellent appraisal of the curriculum and faculty; however, the different VR performance scores indicated that great satisfaction does not necessarily translate into great skills improvement.

**Limitations**

This study has some limitations. A first limitation is that trainees were video-recorded in a simulated environment rather than in clinical practice, and no patient outcomes could be measured. The relationship between errors and complications will need to be confirmed in large-scale clinical trials, but the association between skills and outcomes had been shown before and appears strong and substantial. The transfer of training from simulation to clinical practice has already been quantified.

A second limitation is that the VR simulation videos did not allow or cover the surgical parts of the procedure. The authors acknowledge the importance of surgical skills in the device implantation.
process, but a large part of a CIED procedure was still included, and the scored metrics were sufficient to demonstrate the effect of PBP compared with traditional SBT.

A third limitation is that the faculty were different in the 2 study groups. Although their seniority and their implant and teaching experience were similar, we cannot exclude that their teaching style, their interaction with the trainees, and some details in their implant technique may have differed. Because PBP instructors were trained on the PBP method before the study commenced, their modified teaching approach would almost certainly have affected the SBT too if applied. This would have potentially introduced a considerable confounding bias (no more traditional teaching) than using different, but equally skilled, faculty for the 2 groups.

Conclusions

In this randomized trial, a PBP approach to novice cardiac device implanter training resulted in superior objectively assessed operator performance compared with a traditional SBT approach. If applied systematically, the PBP model may represent an alternative to the Halstedian training framework and ensures trainees reach a predefined, objectively assessed performance level before proceeding to (supervised) in vivo practice. Future studies in CIED implantation will quantify the effect of such an approach on the reduction of procedure-related patient complications.
Conflict of Interest Disclosures: Mr Mascheroni reported receiving nonfinancial support from KULeuven and being an employee of Medtronic during the conduct of the study. The present research project has been conducted as part of Mr Mascheroni's PhD studies in Biomedical Sciences at KULeuven. Dr Stockburger reported receiving personal fees from Medtronic during the conduct of the study and personal fees from Biotronik Research Collaboration outside the submitted work. Dr Patwala reported receiving honoraria from Medtronic during the conduct of the study. Dr Mont reported receiving personal fees from Medtronic during the conduct of the study and outside the submitted work. Dr Rao reported receiving grants from Liverpool Heart and Chest Hospital and personal fees from Medtronic during the conduct of the study. Mr Retzlaff reported receiving personal fees from Medtronic during the conduct of the study and outside the submitted work. Dr Garweg reported receiving grants from Medtronic, Abbott, and Biotronik outside the submitted work. Dr Verbelen reported receiving personal fees from Medtronic outside the submitted work. Dr Gallagher reported receiving personal fees from Medtronic during the conduct of the study.

Funding/Support: This study was supported by Medtronic, which provided the training infrastructure and organization.

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and the decision to submit the manuscript for publication.

Data Sharing Statement: See Supplement 2.

REFERENCES


**SUPPLEMENT 1.**

Trial Protocol

**SUPPLEMENT 2.**

Data Sharing Statement