Simulation Training for Residents Focused on Mechanical Ventilation

A Randomized Trial Using Mannequin-Based Versus Computer-Based Simulation

**Introduction:** Advances in knowledge regarding mechanical ventilation (MV), in particular lung-protective ventilation strategies, have been shown to reduce mortality. However, the translation of these advances in knowledge into better therapeutic performance in real-life clinical settings continues to lag. High-fidelity simulation with a mannequin allows students to interact in lifelike situations; this may be a valuable addition to traditional didactic teaching. The purpose of this study is to compare computer-based and mannequin-based approaches for training residents on MV.

**Methods:** This prospective randomized single-blind trial involved 50 residents. All participants attended the same didactic lecture on respiratory pathophysiology and were subsequently randomized into two groups: the mannequin group (n = 25) and the computer screen–based simulator group (n = 25). One week later, each underwent a training assessment using five different scenarios of acute respiratory failure of different etiologies. Later, both groups underwent further testing of patient management, using in situ high-fidelity simulation of a patient with acute respiratory distress syndrome.

**Results:** Baseline knowledge was not significantly different between the two groups (P = 0.72). Regarding the training assessment, no significant differences were detected between the groups. In the final assessment, the scores of only the mannequin group significantly improved between the training and final session in terms of either global rating score (3.0 (2.5–4.0) vs. 2.0 (2.0–3.0), P = 0.005) or percentage of key score (82% vs. 71%, P = 0.001).

**Conclusions:** Mannequin-based simulation has the potential to improve skills in managing MV.


**Key Words:** Simulation training, mechanical ventilation, computer simulations, acute respiratory failure, critical care.

In recent years, advances in knowledge regarding mechanical ventilation (MV), in particular lung-protective ventilation strategies with low tidal volume, have been shown to reduce mortality. However, despite positive clinical trials and widespread access to recommendations, the translation of these advances in knowledge into better therapeutic performance in real-life clinical settings continues to lag. Indeed, acute respiratory distress syndrome (ARDS) is underestimated and protective ventilation is therefore underused. Needham et al recently found that only 41% of the observed ventilator settings in 13 intensive care units were compliant with low tidal volume ventilation guidelines for ARDS. Similarly, in a large observational study, Bellani et al found that ARDS was undertreated in terms of the use of recommended approaches to MV and the use of adjunctive measures and the in-hospital mortality is still unacceptably high. Even agreement among physicians seems to be low, when presented with the same patient cases, ten experienced Danish doctors selected different ventilator settings, and 33% considered the settings selected by other doctors to be unacceptable. These findings highlight the need for greater compliance to recommendations of the ARDS network and the importance of enhancing the training of clinical residents.

Simulation-based training may be a valuable addition to traditional didactic teaching. Several studies have described the benefits of simulation-based training on physician performance in clinical areas such as critical intraoperative events, airway management, noncardiac critical care ultrasonography, and obstetric emergencies. However, few studies have investigated the use of simulation training for teaching MV, and the existing studies report mixed results. Simulation-based education was shown to be more effective than traditional teaching and comparable with live-animal laboratory. However, not all the researchers reached the same conclusion because Corbridge et al did not find significant improvement when comparing simulation-based education and traditional didactic education. Moreover, it is not clear whether training...
performed using computer-based simulation provides the same benefit as that using the more expensive, lifelike, mannequin-based simulation.

The purpose of the present study was, therefore, to compare two approaches to simulation-based training of MV: computer and mannequin based. Our hypothesis was that the introduction of the mannequin in training would provide a more interactive learning experience compared with computer-based training, which would translate into better assessment and management performance. We also developed two new assessments for MV skills, a training assessment consisting of different clinical scenarios aimed at specific MV problems, and a final assessment representing a complex clinical case of ARDS. Instructors performed assessment during simulation by evaluating key actions and overall solving of the scenarios. For the final assessment of the performance of the two groups, we implemented a high-fidelity mannequin-based simulation system, using a mannequin with an integrated lung simulator capable of changing pulmonary resistance and compliance.

**MATERIALS AND METHODS**

**Study Design**

A prospective randomized trial was performed at the Simulation Centre of the University of Ferrara (Italy) from January 2013 to April 2013, evaluating the training outcome of 53 anesthesiology residents. As summarized in Figure 1, all participants attended the same didactic lecture session on respiratory pathology and respiratory failure applied to MV, which was specifically aimed at creating a homogeneous background of knowledge. During the didactic lecture, the learning objectives for all the study period were explained; of note, these learning objectives were the main topics of the didactic lecture (management of COPD, Near Fatal Asthma, and ARDS). No further learning objectives were given for the final assessment compared with the beginning, to avoid influence by solo studying of the residents. Residents were then randomized into two groups: the mannequin group (MG), to be trained on the mannequin-based simulator, and the computer group (CG), to be trained on the computer screen–based simulator. One week later, each group performed simulation training with both the MG and CG being exposed to five different training scenarios (see Appendix, Supplemental Digital Content (SDC) 1, http://links.lww.com/SIH/A335, which contains all scenarios explained in detail). During performance of the training scenarios, the instructors assessed the skills of each participant (training assessment). The training was followed by a debriefing session for both groups, aimed at discussing correct actions and best approaches to solving each scenario. Three weeks later (with no interval clinical or theoretical training), both groups underwent a final skills assessment, using in situ high-fidelity simulation (HFS), administered on an individual basis. All participants were challenged on the same ARDS scenario, this time in a vacant real room in the intensive care area of the emergency department. The study protocol was approved by our institutional ethics committee (number of the protocol: 161088), and residents provided written informed consent before participating.

The study cohort comprised anesthesiology residents from junior to senior year and clinical anesthesia (CA) levels CA-1 to CA-4 (Table 1). Every resident had a practice of 2 months in intensive care unit (ICU) for each year of residence. All the trainees knew that the purpose of the training was to ensure an improvement in the ability of managing MV. Participation was not mandatory, but all recruited residents participated in the study. However, the three residents recruited for pilot testing were not included in the study, leaving 50 residents. Randomization to the MG or CG was performed by computer-generated randomization based on year of residency, to ensure

---

**FIGURE 1.** Flowchart of the study.

---

**TABLE 1.** Distribution of participants by training modality and level.

<table>
<thead>
<tr>
<th>Level</th>
<th>Computer Group</th>
<th>Mannequin Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA-1</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>CA-2</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>CA-3</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>CA-4</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>

---

**Enrollment**

- Assessed for eligibility (n=53)
  - Excluded (n=4)
  - Recruited for pilot testing (n=50)

**Allocation**

- Randomized (n=50)
  - Allocated to Mannequin Group (n=25)
    - Received 22 questions written test (n=25)
  - Allocated to Computer Group (n=25)
    - Received 22 questions written test (n=25)

**Initial evaluation**

- Allocated to Mannequin Group (n=25)
  - Received 5 scenarios training session (n=25)
  - Allocated to Computer Group (n=25)
  - Received 5 scenarios training session (n=25)

**Final assessment**

- Allocated to Mannequin Group (n=25)
  - Performed final assessment on ARDS scenario (n=25)
  - Allocated to Computer Group (n=25)
  - Performed final assessment on ARDS scenario (n=25)
Excel 2007 tool developed by Iotti et al. Virtual MV allowed Virtual MV v4.0 (Mechanical Ventilation Simulator), a Microsoft Simulation Training com/SIH/A336, for the complete questionnaire). principles of MV (see Appendix, SDC 2, http://links.lww. residents had equal distributions of knowledge and training. The we administered a test to assess whether the two groups of res- Simonini and R.R.), with each scenario covering tasks reflecting relevant areas in best practice for MV. All the instructors involved in the study were anesthetists with at least 10 years of experience and participated in the respective debriefing. Scenarios were presented to each group in random order, and 15 minutes were allowed for completion, followed by a 15-minute debriefing. The presentation of each scenario given by the instructors is presented in Appendix 1. After the presentation, each scenario was performed by residents as individuals. All the residents performed all the five scenarios and participated in the respective debriefing. Scenarios were specifically designed to cover most of the items necessary for the final assessment (ie, titration of fraction of inspired oxygen [FIO₂], increase oxygenation, titration of optimal PEEP to ensure alveolar recruitment and stabilization, calculation of static compliance (PPlat, PEEPi), usage of recruitment maneuvers to open collapsed part of the lung and to improve oxygenation) (Table 2).

The performance of each resident during the training was assessed by a key action score and a global rating score (GRS).

### Baseline Assessment

After the didactic lecture, but before simulation training, we administered a test to assess whether the two groups of residents had equal distributions of knowledge and training. The test consisted of 22 multiple choice questions that covered principles of MV (see Appendix, SDC 2, http://links.lww.com/SIH/A336, for the complete questionnaire).

### Simulation Training

The CG was trained on a free online software program, Virtual MV v4.0 (Mechanical Ventilation Simulator), a Microsoft Excel 2007 tool developed by Iotti et al. Virtual MV allowed residents to set up the ventilator in presence of various diseases and to modify the parameters of respiratory mechanics, such as resistance, compliance, tidal volume, respiratory rate, and inspiratory and expiratory pause. Any changes in the respiratory variables, including pressure, flow, and volume traces, were displayed on the computer screen. This allowed residents to practice calculating dynamic and static compliance, resistance, and intrinsic positive end-expiratory pressure (PEEP) by simply pushing the hold button on the worksheet.

Simulation for MG was performed using a SimMan mannequin and Advanced Video System (Laerdal, Stavanger, Norway) and a Servo 300 ventilator (Siemens-Elena, Solna, Sweden).

### Training Assessment

Performances in management of the training scenarios (training assessment) were evaluated by two instructors (C.V. and R.R.), with each scenario covering tasks reflecting relevant areas in best practice for MV. All the instructors involved in the study were anesthetists with at least 10 years of experience in ICU. Scenarios were presented to each group in random order and 15 minutes were allowed for completion, followed by a 15-minute debriefing. The presentation of each scenario given by the instructors is presented in Appendix 1. After the presentation, each scenario was performed by residents as individuals. All the residents performed all the five scenarios and participated in the respective debriefing. Scenarios were specifically designed to cover most of the items necessary for the final assessment (ie, titration of fraction of inspired oxygen [FIO₂], increase oxygenation, titration of optimal PEEP to ensure alveolar recruitment and stabilization, calculation of static compliance (PPlat, PEEPi), usage of recruitment maneuvers to open collapsed part of the lung and to improve oxygenation) (Table 2).

The performance of each resident during the training was assessed by a key action score and a global rating score (GRS).

The design of the key action score was inspired by the specific items investigated by Cox et al. These actions reflecting guidelines that in general are universally and well as locally accepted. The score reflected the number of correct actions performed (diagnosis, initial treatment, and final treatment). Key actions, that is, those needed to achieve successful patient management, were identified by an expert panel providing a checklist used to create training scenarios assessing resident skills. The checklist for each scenario included ten items that instructors registered as done or not done, yielding one point or no point, respectively. The SDC shows the key action score checklist for the five training scenarios.

The GRS was used to assess overall performance (both technical and nontechnical skills) with instructors using a 1 to 4 scale (1 = poor: problem not identified; 2 = marginal: problem identified but not solved; 3 = acceptable: problem identified and partially solved; and 4 = good: problem identified and completely solved). As an overall score, the GRS addressed the participant’s ability, attitude, and confidence in identifying and solving the problem.

### Final Assessment

Three weeks after performing the training scenarios, participants from both groups underwent a final assessment using HFS in situ, as individual 20-minute sessions. All the residents continued their usual clinical practice between training and final assessment, without other theoretical or practical training. In situ HFS was conducted with a SimMan (Laerdal Corporation, Stavanger, Norway) mannequin placed in an intensive care unit bed, within a fully equipped intensive care room. To obtain an HFS environment, we equipped the mannequin with a lung simulator (Adult/Pediatric Lung Model; Pulmonary Mechanics Graphics Module PMG 3000; and SB2000 Spontaneous Breathing Module, Pittsburgh, Pa) (Fig. 2). The simulator allowed changes in compliance and resistance, and it could trigger the ventilator and display flow, volume, and pressure curves by connection via RS-232 to a personal computer.

### TABLE 2. Scoring System for the Final Assessment Test Scenario on Refractory Hypoxemia

<table>
<thead>
<tr>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Evaluate acute respiratory distress. Make decisions relative to</td>
</tr>
<tr>
<td>clinical and laboratory assessment, pathophysiology, and oxygen</td>
</tr>
<tr>
<td>administration (eg, titration of Fio₂).</td>
</tr>
<tr>
<td>2. Exclude misplacement of the tracheal tube. Consider a pneumothorax</td>
</tr>
<tr>
<td>after auscultation and percussion of the lungs. Perform a chest x-ray</td>
</tr>
<tr>
<td>and computed tomography evaluation.</td>
</tr>
<tr>
<td>3. Interpret ventilator waveforms and respiratory mechanics parameters.</td>
</tr>
<tr>
<td>4. Set ventilator parameters. Ventilator adjustments are made with</td>
</tr>
<tr>
<td>the primary goal of keeping plateau pressure (measured during an</td>
</tr>
<tr>
<td>inspiratory hold of 3–5 s) &lt; 30 cm H₂O.</td>
</tr>
<tr>
<td>5. Troubleshoot the patient/ventilator system when the patient is</td>
</tr>
<tr>
<td>dys-synchronous with the ventilator.</td>
</tr>
<tr>
<td>6. Ensure adequate sedation and consider continuing neuromuscular</td>
</tr>
<tr>
<td>blockade if plateau pressures remain elevated.</td>
</tr>
<tr>
<td>7. Implement a plateau pressure-based strategy with permissive</td>
</tr>
<tr>
<td>hypercapnia (eg, decrease tidal volume/increase respiratory rate).</td>
</tr>
<tr>
<td>8. Calculate optimal PEEP (increment PEEP/titrate PEEP).</td>
</tr>
<tr>
<td>9. Perform recruitment maneuvers and consider placing the patient in</td>
</tr>
<tr>
<td>the prone position for severe hypoxemia.</td>
</tr>
<tr>
<td>10. Consider the use of ECMO if the protective lung ventilation is</td>
</tr>
<tr>
<td>not sufficient to avoid ventilator-induced lung injury.</td>
</tr>
</tbody>
</table>

ECMO indicates extracorporeal membrane oxygenation.
The system was equipped with software for data acquisition and could display loops and waveforms. As such, the feedback from the mannequin equipped with the lung simulator was new for both MG and CG participants. The mannequin’s clinical parameters (blood pressure, electrocardiography, oxygen saturation, and expired carbon dioxide) were displayed on a monitor (Elo Touch Solutions, Milpitas, Calif). To avoid favoring the MG, who had trained on the Servo 300 ventilator, the ICU bed was equipped with a Dräger Evita XL ventilator (Dräger Medical AG & Co, Lübeck, Germany).

Before assessment, residents were allowed to familiarize themselves with the equipment. The assessment started with briefing the resident on patient history and physical condition (see Appendix, SDC 3, http://links.lww.com/SIH/A337, for description of the scenario for the final assessment). During the assessment, the instructors were in the room. All the residents received the instruction to ask for every examination, analysis, or device not present in the room. Key action and GRS were used for assessing resident performance as during the training scenarios (Table 2). All residents were evaluated by two raters (S.S. and E.M.), both blinded to the study group. The scoring was supported by digital audio and video recording of each participant’s performance using a four-quadrant screen with two different video views of the provider and the mannequin. The raters were able to watch the videos recorded for assessing the final score.

Resident Evaluation of Training

Resident’s opinions about the usefulness of the simulation were assessed using an online survey at the end of the final assessment. To avoid influences, opinions were taken before the residents were told about the key action score and the GRS score. Responses were recorded on a Likert-type scale (1 = strongly disagree, 2 = disagree, 3 = uncertain, 4 = agree, and 5 = strongly agree). A Likert-type scale was also used to assess perceived satisfaction with type of simulation training.

Statistical Analysis

A convenience sample of 50 residents was chosen, including all anesthesiology residents from our department. The normality of the variables was tested using the Chen-Shapiro test. All variables were expressed as mean(SD) or as median (interquartile range), when appropriate. A P value of 0.05 or less was considered statistically significant for all tests.

At baseline assessment, the difference in knowledge between CG and MG was evaluated using two-tailed independent-samples t tests. Key action score and GRS were compared between groups using independent-samples t tests or Mann–Whitney tests for normally and nonnormally distributed data, respectively. Paired t test analysis (for normally distributed data) or Wilcoxon signed-rank tests (for nonnormally distributed data) were performed to check for any improvement in key action score and GRSs that occurred from the training to the final assessment within the MG and CG.

Validity Evidence for Training and Final Assessments

As pointed out before, we developed two new assessments, the training and final assessments. We gathered the following validity evidence in accordance with Kane’s framework. The purpose, as also described in SDC 1, was a correct diagnosis and treatment using appropriate MV settings. We aimed to establish scoring validity by the description of the procedures involved, the descriptions of key action and GRSs, as well as the level of expertise of the instructors performing the rating. These details are provided in the methods section as well as in SDC 1 and 3 for training and final assessment, respectively. Furthermore, we established generalization validity by our sampling strategy and empirically evaluating interrater reliability between the two training assessment instructors and final assessment raters GRS. The interrater reliability of the raters’ GRSs was evaluated using the intraclass correlation coefficient (ICC) and Cohen’s κ, in accordance with Cicchetti, as well as Cronbach’s α coefficient. We considered ICCs greater than 0.75 as excellent agreement, while κ had to be 0.80 or higher for a scale to be considered reliable, in accordance with Altman. We established extrapolation evidence by providing authentic clinical cases in training and final assessment scenarios, with the authenticity by design being different between MG and CG. We gathered empirical evidence for extrapolation evidence by linear regression analysis. Univariate linear regression was performed to assess the influence of the training scenarios (1–5) on key action score and GRS at the training assessment, as well as the influence of year of residency (CA1–CA4) on key action score and GRS at the training.
and final assessment. Pass/fail considerations were outside the scope of our study; hence, we collected no evidence for implication validity.

**RESULTS**

**Resident Characteristics and Baseline Knowledge Assessment**

All 50 residents consented to participate in the study and completed the protocol. Demographic characteristics of participants are summarized in Table 1. A written test assessing baseline knowledge did not reveal differences between the two groups ($P = 0.72$).

**Training Assessment**

No significant differences were detected between the two groups both in terms of GRS ($P = 0.59$) or percentage of key action score ($P = 0.74$), as shown in Table 3. The univariate linear regression across both groups showed statistically significant relationships between the score and the year of training ($P < 0.001$), a strong effect size with $\beta$ value of 0.47 ($P < 0.05$) for key action score, and a moderate effect size with $\beta$ value of 0.36 for global score. There were no significant linear relationships between training scenario and key action score or GRS ($P > 0.05$).

The interrater reliability between the GRSs of the two instructors was excellent, with an ICC of 0.87 [95% confidence interval (CI) = 0.74–1.00]. Cohen $\kappa$ showed the evaluators fully agreed on the classification 87.8% of the time. The coefficient of agreement was very good ($\kappa = 0.82$, $z = 8.14$, $P < 0.001$).

**Final Assessment HFS in situ for ARDS Scenario**

When comparing the ability of the two groups to deal with the same scenario in HFS, a statistically significant difference was found both in GRS [MG: (3.0 2.5–4.0) vs. CG: 2.0 (2.0–3.0), $P = 0.005$] and key action score (MG: 82% vs. CG: 71%, $P = 0.001$) (Figs. 3A, B). Linear regression between year of training and final assessment score was not statistically significant after the simulation training (GRS: $P = 0.9$; key action score: $P = 0.8$). Only residents in the MG group showed significant improvement in GRS ($P = 0.003$) and key action score ($P = 0.007$) from training to final assessment (Fig. 3). The interrater reliability between the GRSs of the two raters was excellent, with an ICC of 0.89 (95% CI = 0.77–1.00). Cronbach $\alpha$ coefficient for the checklist was 0.73, indicating good internal consistency. Resident evaluation of training was in general high, indicating good satisfaction scores for both MG and CG (Table 4). However, the MG group had a significantly higher overall satisfaction score ($P = 0.015$).

**TABLE 3.** Key Action and Global Rating Scores for Training Assessment

<table>
<thead>
<tr>
<th>Scenario</th>
<th>MG Key Action, %</th>
<th>CG Key Action, %</th>
<th>P</th>
<th>MG GRS (n)</th>
<th>CG GRS (n)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>86 (12)</td>
<td>86 (12)</td>
<td>0.95</td>
<td>2.0 (1.5–3)</td>
<td>2.0 (2.0–3.0)</td>
<td>0.51</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>81 (10)</td>
<td>71 (21)</td>
<td>0.53</td>
<td>1.5 (1.0–2.5)</td>
<td>2.0 (2.0–2.5)</td>
<td>0.38</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>62 (12)</td>
<td>77 (16)</td>
<td>0.20</td>
<td>2.0 (1.0–2.5)</td>
<td>1.0 (1.0–1.5)</td>
<td>0.28</td>
</tr>
<tr>
<td>Scenario 4</td>
<td>81 (10)</td>
<td>91 (13)</td>
<td>0.41</td>
<td>2.5 (2.0–3.0)</td>
<td>2.0 (2.0–2.4)</td>
<td>0.24</td>
</tr>
<tr>
<td>Scenario 5</td>
<td>61 (27)</td>
<td>58 (30)</td>
<td>0.90</td>
<td>1.0 (0–2.5)</td>
<td>2.0 (1.0–2.0)</td>
<td>0.89</td>
</tr>
<tr>
<td>Overall</td>
<td>76 (16)</td>
<td>73 (18)</td>
<td>0.74</td>
<td>2.0 (1.0–3.0)</td>
<td>2.0 (1.5–3.0)</td>
<td>0.59</td>
</tr>
</tbody>
</table>

Date are presented as mean (SD) or median (interquartile range), as appropriate.

**DISCUSSION**

The main finding of our study was that among anesthesiology residents, mannequin-based simulation seemed more effective than computer-based simulation for improving knowledge and skills related to MV. Indeed, the MG, when compared with the CG, improved on both the key action score and GRS.

To our knowledge, there are no studies on the effectiveness of meeting learning objectives for MV using a computer screen–based simulator compared with a mannequin-based simulator. Our results seem to support the conclusion that mannequin-based training contributes toward strengthening residents’ skills in treating patients in difficult clinical scenarios. The difference between the two groups may be due to the manner of interaction with the simulators, in that the MG participants actively used the ventilator, whereas the CG trained conceptualized actions on the computer screen. Furthermore, the observed difference could be related to the ability of the MG to deal with a “real” patient, including observations of continuous variations in clinical parameters. In contrast, the CG simply interacted with a computer screen.

**Validity of Training and Final Assessments**

Scoring validity of the assessments was supported by the use of a fair evaluation in terms of standardized key actions and use of experienced raters blinded to study groups in their evaluation. Videotaping of sessions was furthermore employed to improve accuracy of the rating. The finding of no relation between training assessment scenario and evaluation also indicates that evaluation was similar across scenarios. Both training and final assessment scenarios were developed to cover a broad section of required skills of MV management with the final assessment scenario representing a significantly complex problem drawing upon skills developed from multiple training scenarios. The broad scope of the scenarios support generalization validity, which is further supported by the very good to excellent interrater reliability observed during both training and final assessments. Scenarios were developed by instructors with certificated experience in MV and simulator programming with scoring checklists provided by an expert panel to allow the evaluation as closely as possible to reflect an authentic clinical context supporting an establishment of extrapolation validity. The relationship between experience level and performance during the simulation training assessment indicates that the scenarios provided a valid method to discriminate between different skill levels supporting extrapolation validity. However, future studies are required to investigate whether a score during the training and final assessments would correlate with performance at the bedside. We found the combined validity evidence in favor of rating and comparing MV skills in the investigated resident sample. However, we did not consider thresholds for passing/failing the assessments nor any consequent actions, and further studies are therefore required to investigate the implication validity of the presented training and final assessments.

**Relation Between Experience and Performance**

This study included residents with different levels of experience. The four levels (CA-1 to CA-4) were equally distributed between the two study groups. Our results showed a significant
linear relationship between the grade year and both GRS and key action score at the training assessment. However, no such relationship was observed at the final assessment. The relationship between experience level and performance during simulation training indicates that traditional didactic lectures alone do not sufficiently raise skill levels of less experienced residents. In contrast, mannequin-based simulation seemed to improve the skills of residents of all experience levels in managing a complex MV scenario. Training assessment probably reflects the baseline knowledge (experienced residents performed better); however, the MG performed better, irrespectively of the year of training, after the training assessment. We believe that this result can be due to the more realistic practice during the training assessment. Previous studies have reported mixed results, either observing no difference in performance between different experience levels or reporting a difference between the least and most experienced levels. The controlled environment in the simulator center helps me focus on clinical education problems and clinical self-confidence. With no risk of harming the patient, I receive useful educational feedback from the training sessions. Practice with the simulator boosts my skills for performing these competencies in the ICU. Increasing the difficulty of simulated clinical problems helps me become a better doctor. The controlled environment in the simulator center helps me focus on clinical education problems.

**The Role of High-Fidelity Simulation**

We chose an ARDS scenario representative of a complex clinical education problem. Acute respiratory distress syndrome is characterized by a sharp reduction of compliance, increased airway resistance, and lung inhomogeneity. The ARDS scenario is intended to simulate a patient with a complex medical condition that requires multiple interventions. The scenario is designed to challenge the residents' ability to manage the patient's clinical condition and to test their ability to make decisions based on real-time feedback.

**TABLE 4. Resident Satisfaction About Training**

<table>
<thead>
<tr>
<th>Likert Scale</th>
<th>CG (M)</th>
<th>MG (M)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practice with the simulator and mechanical ventilator boosts my skills for</td>
<td>3.8 (0.4)</td>
<td>4.6 (0.6)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>performing these competencies in the ICU</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With no risk of harming the patient, I feel more self-confident in making</td>
<td>4.5 (0.2)</td>
<td>4.3 (0.4)</td>
<td>0.89</td>
</tr>
<tr>
<td>decisions during simulation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I receive useful educational feedback from the training sessions</td>
<td>4.1 (0.4)</td>
<td>4.4 (0.5)</td>
<td>0.09</td>
</tr>
<tr>
<td>Practice with the simulator boosts my clinical self-confidence</td>
<td>3.8 (0.4)</td>
<td>4.6 (0.7)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Simulator sessions should be a required component of residency education</td>
<td>3.9 (0.7)</td>
<td>4.6 (0.6)</td>
<td>0.002</td>
</tr>
<tr>
<td>Increasing the difficulty of simulated clinical problems helps me become</td>
<td>4.6 (0.4)</td>
<td>4.7 (0.4)</td>
<td>0.54</td>
</tr>
<tr>
<td>a better doctor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The controlled environment in the simulator center helps me focus on</td>
<td>4.1 (0.5)</td>
<td>4.0 (0.7)</td>
<td>0.97</td>
</tr>
<tr>
<td>clinical education problems</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>4.1 (0.5)</td>
<td>4.5 (0.6)</td>
<td>0.015</td>
</tr>
</tbody>
</table>

Data are expressed as mean(SD).

FIGURE 3. A and B, Global rating score and key action score for training assessment and high-fidelity simulation.

Recently, Khanduja et al advocated for more widespread use of simulations. Regardig simulations in MV, the Harmonising Education in Respiratory Medicine for European Specialists Task Force has begun working toward a standard training curriculum for European clinicians in respiratory critical care, establishing a training framework recommending simulation as an assessment tool. Indeed, there is general agreement that assessment through simulation should be viewed as a learning tool, in which effective real-time feedback from the instructor supports the learning process and translates to improved clinical skills. Hence, our model of HFS could play a fundamental role, providing a valid training and assessment experience at a relatively affordable cost for university departments and clinical facilities. Moreover, HFS could evidence errors in the clinical practice and could help focus residents teaching on the issues that they were not able to accomplish. Even if our study was not designed to evidence specific gaps, two key items were missed frequently in the MG group (ie, “troubleshoot the patient/ventilator system when the patient is dys-synchronous with the ventilator” and “perform recruitment maneuvers and consider placing the patient in the prone position for severe hypoxemia”). Our results could suggest a more specific teaching of these skills.

**Limitations**

Our study has some limitations. First, the evaluation of skills in the final assessment scenario may not predict performance in other scenarios. However, the final assessment scenario was intentionally designed to have greater complexity and to require additional actions that were not covered by any of the training scenarios. Second, this was a single-center study, and the performance and responses to scenarios may be representative of the teaching strategies of our department. Third, we use a novel outcome tool GRS and key action score to evaluate the residents; even if interrater agreement between the two raters was excellent, we cannot exclude that the novelty of our rating tools could influence the reliability of our results. Finally, although we changed the ventilator, scenario and fidelity between the training and final assessment, the MG might have had a test-retest advantage as compared with the simulation group. However, this is one of the justifications for using...
HFS to teaching residents in the field of MV; moreover, the test–retest advantage, if present, might be limited because both groups interacted with the same mannequin in previous simulation teaching experience in different critical settings (ie, airway management), as usual practice in our department.

CONCLUSIONS

Our results indicate that as compared with computer-based training, mannequin-based training provides skills that transfer better to the case of an in situ high-fidelity simulated patients with ARDS. Residents also reported increased satisfaction with mannequin-based training. Our findings support the idea that mannequin-based simulation has the potential to improve MV skills. Future studies should address whether this also translates into better performance in the clinical setting.

ACKNOWLEDGMENTS

The authors thank the Ferrara University anesthesiology and intensive care medicine residents for their dedication to patient care and education.

REFERENCES


10. Park CS, Rochlen LR, Yaghmour E. Acquisition of critical intraoperative event management skills in novice anesthesiology residents by using high-fidelity simulation-based training. Anesthesiology 2010;112:202–211.


