The usability of ventilator maintenance user interface: A comparative evaluation of user task performance, workload, and user experience

Mingyin Jiang, Dongjie Sun, Qiang Li and Daoxiong Wang
Department of Medical Equipment, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Abstract
Poor usability designed of ventilator user interface can easily lead to human error. In this study, we evaluated the usability design of ventilator maintenance user interface and identified problems related to the usability of user interface that could easily cause human error. Sixteen respiratory therapists participated in this usability study. The usability of the ventilator maintenance user interface was evaluated by participants’ task performance (task completion time, task error rate), physiological workload (eye-fixation duration) and perceived workload (NASA-TLX), and user experience (questionnaire). For task performance, task completion time and task error rate showed significant differences. For task completion time, significant difference was found when conducting ventilator self-test ($p < 0.001$), replace the breathing circuit ($p = 0.047$), and check battery status ($p = 0.005$). For task error rate, the three ventilators showed significant difference ($p = 0.012$), and the Serov I showed a significantly higher task error rate than the Boaray 5000D ($p = 0.031$). For workload, the Serov I was associated with higher physiological and perceived workloads than other ventilators ($p < 0.05$). For user experience, the Boaray 5000D received better scores among the ventilators in terms of ease to maintain, friendly to maintain, and willingness to use ($p < 0.05$, respectively). Our study adds available literature for usability evaluation of...
ventilator maintenance user interface. The results indicate that the maintenance user interface of the Boaray 5000D performed better than the other two tested ventilators. Moreover, the study results also proved that eye-fixation duration can be a reliable tool for evaluating the usability of ventilator user interface.

**Keywords**
Critical care, equipment design, eye movement measurements, ventilator, user computer interface

**Introduction**
Ventilators are important medical devices in critical care set-ups, and approximately one-third to one-half of patients in intensive care units require ventilator support. Patients who require mechanical ventilation are usually fragile, and errors in ventilator use may increase death case. Ventilator use poses risks to patients, such as ventilator-associated pneumonia, atelectasis, infections, oxygen toxicity, and lung injury. Another significant risk with ventilator use is human error, which could be potentially harmful to a patient. A previous study analyzed ventilator-related adverse events in three incident reporting systems (the Pennsylvania Patient Safety Authority’s Patient Safety Reporting System, UHC’s Safety Intelligence Patient Safety Organization database, and the FDA’s Manufacturer and User Facility Device Experience database) and reported that more than 49% of ventilator-related adverse events (in the PSA database) were caused by human use errors.

However, adverse events as a result of human error are mainly blamed on medical personnel. previous studies reported poor designs of user interface of medical devices as the primary cause of adverse events. Therefore, a user-friendly of the design of the user interface of a mechanical ventilator may improve human-machine interactions and reduce use risks. Recently, the International Electrotechnical Commission and US Food and Drug Administration published standards and guidelines requiring medical device manufacturers to design medical devices with improve user interface thus reducing human use errors. The usability testing of medical devices is widely used to evaluate user interfaces thus reducing human use errors. Several studies have evaluated the usability of ventilator user interface, however, few studies report usability of ventilator maintenance user interface. Good ventilator maintenance is key in ensuring patient safety, therefore, maintenance user interfaces should have appropriate usability design to reduce human use errors which leads to potential risks to patients.

Currently, to our knowledge, no study has addressed usability of ventilator maintenance user interfaces, therefore, the aim of this study was to evaluate the usability of ventilator maintenance user interfaces through user task performance, workload and experience via qualitative and quantitative measurements.
Materials and methods

The nature of the study

This is a prospective usability study on performance of five maintenance tasks by respiratory therapists on three tested ventilators in a university affiliated hospital ICU treatment room in Wuhan, Hubei province, China. The aim of this study was to evaluate usability of ventilator maintenance user interface, rather than evaluation of participant performance.

Usability test ventilators

The three ICU ventilators selected for our usability study were, the Evita 4 (Draeger, Lubeck, Germany, software version: 04.24 07/12/11, manufacturing date: May 2012), Servo I (Maquet, Solna, Sweden, software version: v5.00.00, manufacturing date: July 2014), and Boaray 5000D (Probe, Shenzhen, China, software version: 0A_006_V06.10.02_151119, manufacturing date: October 2014). These ventilators were also tested in our previous usability studies. In addition, a test lung (Venti.Plus™, GaleMed, Taipei, Taiwan, China) was connected to each ventilator for ventilator maintenance user interface usability testing. For a detailed description of the test ventilators, please see the Supplemental Materials.

Participants

Mechanical ventilators in hospitals are usually maintained by respiratory therapists. Previous studies report that eight to twelve participants in usability study provide reliable findings. In this study, the participants included 16 respiratory therapists, who had experience in ventilator maintenance in our hospital. Before the formal usability test, all participants were taken through ventilator maintenance training courses by a clinical engineer familiar with the three tested ventilators. During the training course, the clinical engineer was available to answer the questions of the participants regarding ventilator maintenance. Moreover, each participant was given a learning goal that he or she should know how to maintain the three test ventilators. A preliminary test with three participants was given to improve the testing process and verify the reliability of the test data.

Usability study test tasks

Five maintenance tasks were selected to test the usability of each ventilator: (1) clean ventilator surfaces; (2) clean the air filter; (3) conduct the ventilator self-test; (4) replace the breathing circuit; and (5) check battery status. Participants performed the test tasks on each ventilator. For a detailed description of the test tasks, please refer to the Supplemental Materials.
Task performance

The task performance of each participant was evaluated through the task completion time and task error rate (the percentage of tasks that the participant failed to complete). A shorter task completion time and a lower task error rate showed a better task performance by the participants.

Workload measure

In this study, we evaluated the workload of each participant by measuring physiological and perceived workloads.

Physiological workload

Participant’s physiological workload was assessed by the duration of eye-fixation on a test task. Previous studies report that the duration of eye-fixation indicates the complexity of personal information processing or cognitive activities. A long eye-fixation duration indicates difficulties in information processing or cognitive activities, implying increasing workload on a person. The eye motion data of the participants were collected with a Tobii Glasses 2 Eye Tracker (Tobii Technology, Danderyd, Sweden). Before the test, the participant went through a pupil calibration process with the eye tracker. Data on a participant’s eye-fixation duration were collected with the Eye Tracker while the participant performing each task.

Perceived workload

The National Aeronautics and Space Administration Task Load Index (NASA–TLX) was used to assess perceived workload. Personal perceived workload was evaluated using the questionnaire based on six assessment facets (mental demand, temporal demand, physical demand, frustration, performance, and effort). The workload of each participant was assessed in two steps. First, participants went through six assessment metrics after completing all test tasks on each ventilator. Second, participant would be shown 15 pairs of comparison groups of rating scale for the workload assessment metrics (such as Performance vs Frustration). Participant was then allowed to choose the assessment aspect that mirrored their workload in performing tasks on ventilator. Finally, participant’s choices were used to calculate weighted combination of workload scores. The NASA-TLX questionnaire used in this study is presented in Supplemental Materials. Higher scores indicate a higher perceived workload for the participant. The NASA–TLX has been used for assessing the usability of ventilators.
User experience

A questionnaire assessing the ease to maintain (1 very difficult to 5 very easy), friendly to maintain (1 very bad to 5 very good) and willingness to use (1 certainly not to 5 certainly) was used to evaluate user experience of the participants. Higher scores indicate a better user experience.

Study procedure

This study was conducted in a university affiliated hospital ICU treatment room in Wuhan, Hubei province, China. Ethics Committee was obtained from Tongji Medical College, Huazhong University of Science and Technology Ethics Committee (IORG No: IORG0003571).

Each participant performed five test tasks on three tested ventilators. Furthermore, the three ventilators were randomly tested by each participant (for a detailed description of ventilator allocation, please see the Supplemental Materials, Table S1). A pupil calibration was carried out on each participant before performing each task. After calibration, a tester would stand next to the test ventilator and inform the participant of the test task. The participant would carry out a task after a test signal is given. Each participant was given only one chance to complete each task. Participant informed the tester when they completed each task. The tester confirmed the task completion status of the participant. Thereafter, the participant would perform next task. Each participant performed all five tested tasks on each tested ventilator. The usability testing of the three ventilators lasted about two hours. The participant could have a rest whenever they wanted.

Data collection

An eye tracking device was used to collected eye movement data for each participant when performing the tasks. The task completion time for each participant was recorded by the tester, and the task completion status (failure or success) confirmed. After completing the five tested tasks on a test ventilator, participant filled in the NASA–TLX questionnaire and the user experience questionnaire before performing the tasks on the next test ventilator.

Statistics

Task completion time, task error rate, eye-fixation duration, the NASA-TLX questionnaire, and user experience questionnaire data were analyzed using the Friedman nonparametric test by SPSS 20 (IBM Corporation, Armonk, New York). Data are expressed as median (interquartile range, IQR). $p < 0.05$ was considered statistically significant. In addition, we performed post hoc multiple
comparison tests with the Dunn-Bonferroni test to analyze the differences between the tested ventilators.34

Results

Task performance

Figure 1 shows the task completion time taken by participants to complete the five tasks on each ventilator. The task completion time for task 3, 4, and 5 showed significant differences for the three tested ventilators. Further, participants spend less time to conduct the ventilator self-test on the Boaray 5000D (25.00 (22.000–28.000)) than on the Evita 4 (107.500 (67.250–234.750), \(p = 0.008\)) and the Servo I (243.500 (112.250–313.250), \(p < 0.001\)). For the task of check battery status, participants spent more time to accomplish the task on the Evita 4 (28.500 (16.250–39.750)) than on the Boaray 5000D (5.500 (3.000–7.750), \(p = 0.004\)).

Table 1 shows the maintenance task failure rates for each task on all tested ventilators. Task failures were highest for clean the air filter (40.8% of all task failures), followed by check battery status (32.7% of all task failures) and conduct the ventilator self-test (26.5% of all task failures). Task error rates were analyzed for each of the tested ventilators, and results presented in Table 2. Task error rates were significantly different among the three ventilators \((p = 0.012)\). Post hoc multiple comparison tests showed that participants had fewer task failures on the Boaray 5000D than the Servo I \((p = 0.031)\).
Assessment of workload

**Physiological workload.** Table 3 presents the eye-fixation duration for each task for the five tasks on the three tested ventilators. A significant difference was recorded for conduct the ventilator self-test ($p < 0.001$) and check battery status ($p = 0.015$). Moreover, the total eye-fixation duration of all tasks was significantly different among the three tested ventilators ($p < 0.001$). A post hoc multiple comparison test of eye-fixation duration among the three tested ventilators was conducted, and the results are shown in Table 3. Five out of the 12 comparisons showed significant differences among the three ventilators ($p < 0.05$). Considering the total eye-fixation duration of all tasks, the Boaray 5000D (90.635 (65.810–96.785)) showed a lower eye-fixation duration with the Evita 4 (184.920 (153.635–284.055)) and the Servo I (281.920 (199.430–350.870), $p < 0.05$). This implies that participants had a lower physiological workload when performing tasks on the Boaray 5000D.

**Perceived workload.** Table 4 and Figure 2 show each facet of workload scores for the three tested ventilators. Significant differences were observed for the TLX scores and the temporal demand among the three tested ventilators ($p < 0.05$). The Servo
### Table 3. Eye-fixation duration for each task and the results for the eye-fixation duration from post hoc multiple comparisons.

| Task | Eye-fixation duration (second) | p       | Post hoc tests with Bonferroni correction | p       |
|------|-------------------------------|---------|------------------------------------------|---------|
|      | Evita 4                        | Servo I | Boaray 5000D                             |         |
|      | Median (IQR)                   | Median (IQR) | Median (IQR) | p       | Post hoc tests with Bonferroni correction | p       |
| Task 1: clean ventilator surfaces | 31.560 (21.515–45.825) | 39.520 (23.585–52.090) | 24.560 (22.745–33.860) | 0.269   | Boaray 5000D | Evita 4 | 0.004   |
| Task 2: clean the air filter     | 10.300 (9.275–20.980)    | 11.520 (7.080–14.760) | 16.970 (11.460–18.620) | 0.368   | Boaray 5000D | Servo I | <0.001  |
| Task 3: conduct the ventilator self-test | 108.340 (64.685–182.605) | 196.530 (80.450–227.840) | 16.890 (15.420–19.080) | <0.001  |
| Task 4: replace the breathing circuit | 12.380 (8.750–19.165)   | 17.910 (13.175–25.220) | 11.040 (7.420–24.055) | 0.305   |
| Task 5: check battery status     | 20.670 (13.050–35.540)   | 19.090 (8.365–37.110) | 4.710 (3.930–8.080)   | 0.001   |
| Total fixation duration of all tasks (second) | 184.920 (153.635–284.055) | 281.920 (199.430–350.870) | 90.635 (65.810–96.785) | <0.001  |

The bold entries are statistically significant results (P < 0.05).
I had the highest TLX score (41.670 (18.330–61.500), \( p = 0.047 \)), and temporal demand score (8.000 (6.168–10.670), \( p = 0.023 \)) compared to those of the other ventilators (Figure 2, Table 4). Post hoc multiple comparison tests were performed (Table 4), significantly higher temporal demand scores were observed for the Servo I (8.000 (6.168–10.670)) compared with the Boaray 5000D (4.665 (2.168–5.330), \( p = 0.031 \)).

**User experience evaluation**

User experience results for each ventilator as evaluated by the participants are presented in Table 5. Significant differences were observed in ease to maintain (\( p = 0.007 \)), friendly to maintain (\( p = 0.027 \)), and willingness to use (\( p = 0.039 \)) among the three tested ventilators. The Boaray 5000D demonstrated a better user experience compared with the other tested ventilators. Post hoc multiple comparison tests were also performed (Table 5), and the results showed that the Boaray 5000D (5.000 (4.000–5.000)) was easier to maintain than the Servo I (4.000 (4.000–4.000), \( p = 0.040 \)).

**Participant demographics**

Demographic data for the 16 participants were collected before formal study. Half of the participants were male (\( n = 8 \)) and the other half were female (\( n = 8 \)). The median ages for participants were 27 (27–30) years old. Two participants had no use experience of all tested ventilators (but had other ventilator use experience), 11 participants had prior use experience of the Evita 4, and nine participants had the Servo I and Boaray 5000D use experience (Supplemental Materials Table S2). Notably, 68.75% of the participants had one or more years of use experience with
Table 4. NASA-TLX workload for ventilators and the results of the NASA-TLX workload from post hoc multiple comparisons.

| NASA-TLX workload scores | Evita 4 Median (IQR) | Servo I Median (IQR) | Boaray 5000D Median (IQR) | p | Post hoc tests with Bonferroni correction | p |
|--------------------------|----------------------|----------------------|---------------------------|---|------------------------------------------|---|
| TLX (task load index)    | 23.000 (20.000–39.668) | 41.670 (18.330–61.500) | 24.335 (17.503–25.165) | **0.047** | Boaray 5000D Evita 4 | 1.000 |
| Mental demand            | 3.665 (0.835–7.668) | 2.670 (1.498–12.998) | 3.665 (2.000–7.668) | 0.952 | Boaray 5000D Servo I | 0.065 |
| Physical demand          | 1.330 (0.000–4.000) | 0.665 (0.000–6.335) | 0.665 (0.000–2.670) | 0.657 | Evita 4 Servo I | 0.155 |
| Temporal demand          | 6.670 (4.835–11.500) | 8.000 (6.168–10.670) | 4.665 (2.168–5.330) | **0.023** | Boaray 5000D Evita 4 | 0.155 |
| Performance              | 5.330 (3.003–8.498) | 3.665 (2.670–9.500) | 4.000 (4.000–6.503) | 0.732 | Boaray 5000D Servo I | **0.031** |
| Effort                   | 4.000 (2.670–6.670) | 10.670 (2.333–15.500) | 2.670 (1.333–4.998) | 0.285 | Evita 4 Servo I | 1.000 |
| Frustration              | 0.665 (0.000–3.330) | 2.335 (0.000–5.330) | 2.670 (0.000–4.000) | 0.094 | |

The bold entries are statistically significant results (P < 0.05).
Table 5. User experience for each ventilator.

| Post test questionnaire          | Evita 4                  | Servo I                  | Boaray 5000D              | p          | Post hoc tests with Bonferroni correction | p          |
|---------------------------------|--------------------------|--------------------------|----------------------------|------------|------------------------------------------|------------|
|                                 | Median (IQR)             | Median (IQR)             | Median (IQR)              |            |                                          |            |
| Ease to maintain: 1 (very difficult) to 5 (very easy) | 4.000 (4.000–5.000)      | 4.000 (4.000–4.000)      | 5.000 (4.000–5.000)       | 0.007      | Boaray 5000D                            | Evita 4    | 0.933                                    |
|                                 |                          |                          |                            |            | Boaray 5000D                            | Servo I    | **0.040**                                |
|                                 |                          |                          |                            |            | Evita 4                                 | Servo I    | 0.399                                    |
| Friendly to maintain: 1 (very bad) to 5 (very good) | 4.000 (4.000–4.750)      | 4.000 (3.000–4.000)      | 4.000 (4.000–4.750)       | **0.027**  | Boaray 5000D                            | Evita 4    | 1.000                                    |
|                                 |                          |                          |                            |            | Boaray 5000D                            | Servo I    | 0.155                                    |
|                                 |                          |                          |                            |            | Evita 4                                 | Servo I    | 0.399                                    |
| Willingness to use: 1 (certainly not) to 5 (certainly) | 4.000 (3.000–4.750)      | 4.000 (3.000–4.000)      | 4.000 (4.000–4.000)       | **0.049**  | Boaray 5000D                            | Evita 4    | 1.000                                    |
|                                 |                          |                          |                            |            | Boaray 5000D                            | Servo I    | 0.155                                    |
|                                 |                          |                          |                            |            | Evita 4                                 | Servo I    | 0.648                                    |

The bold entries are statistically significant results (P < 0.05).
the tested ventilators. Further, 56.25% of the participants had four or more years of work experience as respiratory therapists. (Supplemental Materials Table S3 and S4).

A perfect balance of participant work experience level with each tested ventilator was not possible as participant experience on tested ventilator in our department of critical care medicine was not evenly distributed. Using the demographic data collected from participants, we performed multiple regression models with collected data all variables in this study. The results showed that participant work experience affected several variables for the Boaray 5000D and the Servo I. Notably, participant work experience affected task 2 completion time for the Boaray 5000D, \( (F(1,14) = 9.060, p = 0.009, \text{adj. } R^2 = 0.350, \beta = 0.627) \). Further, participant work experience affected task 3 completion time for the Servo I, \( (F(1,14) = 4.836, p = 0.045, \text{adj. } R^2 = 0.204, \beta = 0.507) \). In addition, participant work experience affected the willingness to use for the Boaray 5000D \( (F(1,14) = 6.646, p = 0.022, \text{adj. } R^2 = 0.350, \beta = 0.567) \). Further, participant work experience affected the friendly to maintenance for the Boaray 5000D \( (F(1,14) = 8.621, p = 0.011, \text{adj. } R^2 = 0.337, \beta = 0.617) \), and for the Servo I \( (F(1,14) = 4.885, p = 0.044, \text{adj. } R^2 = 0.206, \beta = -0.509) \). Participant work experience affected the task error rate for the Boaray 5000D \( (F(1,14) = 12.741, p = 0.003, \text{adj. } R^2 = 0.439, \beta = 0.690) \). Notably, data for the other variables were not affected by ventilator use experience in this study.

**Discussion**

The ergonomic design of medical devices \(^{35,36}\) and the potential for usability problems both contribute to adverse events, which has been reported in several studies. \(^{12,37}\) In this study, we report that ventilators used in our local medical institution in China have poorly designed maintenance user interface, and have a high task error rate (11.3%–28.8%, Table 2).

The poor usability design of ventilator maintenance user interface was demonstrated by several usability problems observed in this study. Notably, the self-test user interface for the Evita 4 was poorly designed. The self-test user interface for this ventilator was below the center of the screen, making it difficult for the participant to check the information. Further, the self-test feedback information was displayed in a small font at the bottom of the user interface, making it difficult to read. In addition, the translation of the self-test information was not accurate, which was confusing to the participants. These usability problems increase the mental workload of users, and important self-test information can easily be overlooked, resulting in operational failures. This may explain why the participants who performed the ventilator self-test on the Evita 4 had 10 task failures (Table 1).

The clean air filter task showed a high error rate (41.7%, Table 1), and participants reported difficulties in disassemble and installation of the air filter for the Servo I and the Boaray 5000D. This usability problem increased task completion time for ventilator, resulting in task failures (Table 1).
Furthermore, the design of the ventilator power switch showed a usability limitation. A previous study reported that about 91% of Chinese people are right-handed.\(^{38}\) In our study, nearly all participants used their right hand to search for the power switch on the ventilator. The Boaray 5000D and the Servo I had the power switches on the left-hand side. Moreover, the Servo I power switch was concealed behind a sliding cover, thus participants wasted time searching for the switch and starting the machine. This shortcoming increased task completion time (Figure 1) and physiological workload of participants (Table 3). However, the Evita 4 had the power switch on the right side. In a previous study, Vignaux et al. reported difficulties in switching on and off the Servo I which increased operational time.\(^{25}\)

Further, Gonzalez-Bermejo et al. also reported challenges when powering on the Servo I.\(^{39}\) Jiang et al. reported that Chinese users had challenges in operating ventilators with a power switch on the left back side.\(^{19}\) Operation habits of users should be considered when designing ventilator user interface.

Moreover, power-related challenges are the most common adverse events during ventilator use in the FDA’s MAUDE database in 2010, and most are battery-related issues.\(^{9}\) The battery usability problem was also observed in this study. For the Boaray 5000D, the battery status was easily obtained from the user interface, while the battery status for the Servo I was via a menu, thus locating it was challenging, resulting in high task failure rates (Table 1). However, the only way to see the battery status for the Evita 4 was to turn off the power, which may harm patients. Several studies have reported battery failures in hospitals. In a study by Amagasa et al., due to a temporary interruption of electrical supply in the intensive care unit in their hospital, a ventilator shut down instantly without any alarm.\(^{40}\) Fortunately, the nurse noticed this and acted fast, leaving the patient free of injury. Similarly, in the US, a ventilator-dependent patient under home-base care died as a result of power failure.\(^{41}\) Therefore, it is important for medical staff to conveniently and quickly check the battery status. Our study showed that the battery status display should be improved.

In this study, eye motion data (eye-fixation duration) was used as an objective indicator of physiological workload to evaluate the usability of the ventilator maintenance user interface. Eye-fixation duration is related to the complexity of personal information processing or cognitive activities.\(^{27–30}\) A longer eye-fixation duration means an increased physiological workload.\(^{31,32}\) The eye-fixation duration results for each task among the three tested ventilators are shown in Table 3. Significant differences were demonstrated in tasks 3 and 5. For task 3, shorter eye-fixation duration was observed for the Boaray 5000D (16.890 (15.420–19.080)) compared with that for the Evita 4 (108.340 (64.685–182.605), \(p = 0.004\)) and the Servo I (196.530 (80.450–227.840), \(p < 0.001\)). For task 5, longer eye-fixation duration was observed for the Servo I (19.090 (8.365–37.110)) compared with that for the Boaray 5000D (4.710 (3.930–8.080), \(p = 0.014\)). Moreover, significant differences were observed for the total eye-fixation duration of all tasks among the three ventilators (\(p < 0.001\)). The eye-fixation duration for the five test tasks for the Boaray 5000D (90.635 (65.810–96.785)) was shorter compared with that for the
Evita 4 (184.920 (153.635–284.055), \( p = 0.004 \)) and the Servo I (281.920 (199.430–350.870), \( p < 0.001 \)). This finding implies that participants had a lower physiological workload when performing tasks on the Boaray 5000D compared with the Evita 4 and the Servo I. Similar findings have been reported in previous studies.\(^{31,32}\) Furthermore, longer eye-fixation and higher NASA-TLX scores were positively correlated with less user experience.

In this study, we observed that participant work experience affects variable data collected. Usability of ventilator maintenance user interface between experienced respiratory therapists versus less experienced respiratory therapists should be evaluated to improve ventilator usability design.

In this study, a tester stood next to participant when carrying out tasks. This may make the participant nervous when performing test tasks. To avoid this, the tester should keep some distance from the participant and avoid distracting the participant during the test. Further, in real life, respiratory therapist would carry out maintenance work one type. In this study, respiratory therapists were required to perform maintenance tasks on ventilator all three ventilator one after the other. This process may increase participant mental workload. During the research, participants were allowed to rest whenever they wanted, to reduce participant mental stress during the study.

Our study showed that eye-fixation duration is effective in evaluating the usability of ventilator maintenance user interface design. Further, the findings of this study offer information for manufacturers of ventilators to improve the usability of ventilator maintenance user interface design.

**Limitations**

This study had a few limitations. First, the study only tested a sub-set of the required maintenance tasks performed by respiratory therapists. Therefore, the study findings may not be applicable to other user groups, such as the Bio-Medical Equipment Technicians and clinical engineers. Second, the study was carried on three ventilators thus the findings are not be representative for all ventilators. However, the tested ventilators are commonly used in our local medical institutions and were available for our usability research. Third, our participants were respiratory therapists, so the results of this study may not be extrapolated to other types of end-users of ventilators. Finally, the tested ventilators are not the latest ventilator.

**Conclusion**

This study provided available literature for usability evaluation of ventilator maintenance user interface. Notably, maintenance-related usability problems were identified, such as the location of the power switch, battery status, and how to disassemble and install the air filter, which led to a high operational error rate and high mental workload. Based on our results, we can infer that the usability of the
Boaray 5000D performed better than the other tested ventilators. Furthermore, eye-fixation duration is an effective indicator for evaluation of the usability of ventilators. Future studies should include more types of new generation ventilators for the findings to be more representative.

Acknowledgements

We acknowledge the support of the respiratory therapists from the Department of Critical Care Medicine, who made this study possible. We are also thankful to the clinical engineers from the Department of Medical Equipment, who helped us complete the entire study.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by the Research Fund of Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology.

Ethical approval

Ethical approval for this study was obtained from The Ethics Committee of Medical College, Huazhong University of Science and Technology (IORG NO: IORG0003571).

Informed consent

Written informed consent was obtained from all subjects before the study.

Supplemental material

Supplemental material for this article is available online.

References

1. Blanch PB. An evaluation of ventilator reliability: a multivariate, failure time analysis of 5 common ventilator brands. *Respir Care* 2001; 46(8): 789–797.
2. Pogorzelska M, Stone PW, Furuya EY, et al. Impact of the ventilator bundle on ventilator-associated pneumonia in intensive care unit. *Int J Qual Health Care* 2011; 23(5): 538–544.
3. Michaud G and Cardinal P. Mechanisms of ventilator-induced lung injury: the clinician’s perspective. *Crit Care* 2003; 7(3): 209–210.
4. Gilstrap D and MacIntyre N. Patient-ventilator interactions: implications for clinical management. *Am J Respir Crit Care Med* 2013; 188(9): 1058–1068.
5. Chastre J and Fagon JY. Ventilator-associated pneumonia. *Am J Respir Crit Care Med* 2002; 165: 867–903.
6. Chatwin M, Heather S, Hanak A, et al. Analysis of home support and ventilator malfunction in 1,211 ventilator-dependent patients. *Eur Respir J* 2010; 35(2): 310–316.

7. Gravenstein JS. How does human error affect safety in anesthesia?. *Surg Oncol Clin N Am* 2000; 9(1): 81–95.

8. Auriant I, Reignier J, Pibarot M, et al. Critical incidents related to invasive mechanical ventilation in the ICU: preliminary descriptive study. *Intensive Care Med* 2002; 28(4): 452–458.

9. Pham JC, Williams TL, Sparnon EM, et al. Ventilator-related adverse events: a taxonomy and findings from 3 incident reporting systems. *Respir Care* 2016; 61(5): 621–631.

10. Crowley JJ and Kaye RD. Identifying and understanding medical device use errors. *J Clin Eng* 2002; 27(3): 188–193.

11. Israelski EW and Muto WH. Human factors risk management as a way to improve medical device safety: a case study of the therac 25 radiation therapy system. *Jt Comm J Qual Saf* 2004; 30(12): 689–695.

12. Bogner MS. There is more to error in healthcare than the care provider. *Proceedings Hum Fac Erg Soc Annual Meeting* 2005; 49(11): 952–954.

13. Richard JCM and Kacmarek RM. ICU mechanical ventilators, technological advances vs. User friendliness: the right picture is worth a thousand numbers. *Intensive Care Med* 2009; 35(10): 1662–1663.

14. IEC. *IEC 62366-1 medical devices – part 1: application of usability engineering to medical devices*. Geneva: International Electrotechnical Commission, 2015.

15. IEC. *IEC 62366-2 medical devices – part 2: guidance on the application of usability engineering to medical devices*. Geneva: International Electrotechnical Commission, 2015.

16. IEC. *Medical electrical equipment – part 1–6: general requirements for basic safety and essential performance*. Geneva: Usability International Electrotechnical Commission, 2013.

17. FDA (US). *Applying human factors and usability engineering to medical devices*. U.S. food and drug administration. Washington: Food and Drug Administration, 2016.

18. Wiklund ME, Kendler J and Strochlic AY. *Usability testing of medical devices*. Florida: Taylor & Francis, 2011.

19. Jiang M, Liu S, Gao J, et al. Comprehensive evaluation of user interface for ventilators based on respiratory therapists’ performance, workload, and user experience. *Med Sci Monit* 2018; 24: 9090–9101.

20. Jiang M, Liu S, Feng Q, et al. Usability study of the user-interface of intensive care ventilators based on user test and eye-tracking signals. *Med Sci Monit* 2018; 24: 6617–6629.

21. Morita PP, Weinstein PB, Flewwelling CJ, et al. The usability of ventilators: a comparative evaluation of use safety and user experience. *Crit Care* 2016; 20: 263.

22. Spaeth J, Schweizer T, Schmutz A, et al. Comparative usability of modern anaesthesia ventilators: a human factors study. *Br J Anaesth* 2017; 119(5): 1000–1008.

23. Marjanovic NS, De Simone A, Jégou G, et al. A new global and comprehensive model for ICU ventilator performances evaluation. *Ann Intensive Care* 2017; 7: 68.

24. Marjanovic N and L’Her E. A comprehensive approach for the ergonomic evaluation of 13 emergency and transport ventilators. *Respir Care* 2016; 61(5): 632–639.

25. Vignaux L, Tassaux D and Jolliet P. Evaluation of the user-friendliness of seven new generation intensive care ventilators. *Intensive Care Med* 2009; 35(10): 1687–1691.
26. Wiklund M, Kendler J and Stroclic A. Usability testing of medical devices. 2th ed. Boca Ration: CRC Press, 2011.
27. Kules B and Xie B. Older adults searching for health information in MedlinePlus – an exploratory study of faceted online search interfaces. Proceedings Am Soc Info Sci Technol 2011; 48(1): 1–10.
28. Strandvall T. Eye tracking in human-computer interaction and usability research. Lect Notes Comput Sci 2009; 5727: 936–937.
29. Barkana DE, Açık A, Duru DG, et al. Improvement of design of a surgical interface using an eye tracking device. Theor Biol Med Model 2014; 11(1): 48.
30. Liu CJ, Kemper S and Mc Dowd J. The use of illustration to improve older adults’ comprehension of health-related information: is it helpful?. Patient Educ Couns 2009; 76(2): 283–288.
31. Asan O and Yang Y. Using eye trackers for usability evaluation of health information technology: a systematic literature review. JMIR Hum Factors 2015; 2(1): e5.
32. Meghanathan RN, van Leeuwen C and Nikolaev AR. Fixation duration surpasses pupil size as a measure of memory load in free viewing. Front Hum Neurosci 2015; 8: 1063.
33. Hart SG and Staveland LE. Development of NASA-TLX (task load index): results of empirical and theoretical research. Hum Mental Workload 1988; 52(6): 139–183.
34. Dunn OJ. Multiple comparisons using rank sums. Technometrics 1964; 6(3): 241–252.
35. Buckle P, Clarkson PJ, Coleman R, et al. Patient safety, systems design and ergonomics. Appl Ergon 2006; 37(4): 491–500.
36. Martin JL, Norris BJ, Murphy E, et al. Medical device development: the challenge for ergonomics. Appl Ergon 2008; 39(3): 271–283.
37. Horsky J, Zhang J and Patel VL. To err is not entirely human: complex technology and user cognition. J Biomed Inform 2005; 38(4): 264–266.
38. Li X. The distribution of left and right handedness in Chinese people. Acta Psychol Sinica 1983; 15(3): 268–76 [in Chinese]
39. Gonzalez-Bermejo J, Laplanche V, Husseini F, et al. Evaluation of the user-friendliness of 11 home mechanical ventilators. Eur Respir J 2006; 27(6): 1236–1243.
40. Amagasa S, Igarashi A, Yokoo N, et al. Backup failure of an adjuvant battery in an Evita 4® ventilator. Anesthesiology 2008; 108(4): 763–764.
41. Lechtzin N, Weiner CM and Clawson L. A fatal complication of noninvasive ventilation. N Engl J Med 2001; 344(7): 533.

Author biographies

Mingyin Jiang received the B.S. degree in Biomedical Engineering from Hubei University of Science and Technology, China, in 2014, and his Ph.D. degrees in Biomedical Engineering from Huazhong University of Science and Technology, China, in 2019. He is an Engineer in Department of Medical Equipment, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology. His research interests include medical device management, ergonomics, and usability engineering of medical device.

Dongjie Sun received the B.S. degree in engineering from Wuhan University of Technology, China, in 1989. He is a Professor in Department of Medical Equipment, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology. His current research interests include medical device management and design.
Qiang Li received his B.S. degree in Biomedical Engineering from Huazhong University of Science and Technology, China, in 2006. He is an Engineer in Department of Medical Equipment, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology. His research interests include medical device management and training.

Daoxiong Wang received the B.S. degree in Clinical Medicine from Jianghan University, China, in 2001, and the master’s degree in Engineering from Central China Normal University, China, in 2020. He is a Professor in Department of Medical Equipment, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology. His research interests include medical device management, training and quality control of medical device.