

---

Article

Not peer-reviewed version

---

# Using Visual Patient to Show Vital Sign Predictions, a Computer-Based Mixed Quantitative and Qualitative Simulation Study

---

Amos Malorgio , David Henckert , Giovanna Schweiger , Julia Braun , [Kai Zacharowski](#) , [Florian J. Raimann](#) ,  
Florian Piekarski , [Patrick Meybohm](#) , [Sebastian Hottenrott](#) , Corinna Froehlich , [Donat R. Spahn](#) ,  
[Christoph B Nöthiger](#) , [David W. Tscholl](#) , [Tadzio R. Roche](#) \*

Posted Date: 5 October 2023

doi: [10.20944/preprints202310.0249.v1](https://doi.org/10.20944/preprints202310.0249.v1)

Keywords: avatar; machine learning; monitoring; predictive models; visual patient; vital sign predictions



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

## Article

# Using Visual Patient to Show Vital Sign Predictions, a Computer-Based Mixed Quantitative and Qualitative Simulation Study

Amos Malorgio <sup>1</sup>, David Henckert <sup>1</sup>, Giovanna Schweiger <sup>1</sup>, Julia Braun <sup>2</sup>, Kai Zacharowski <sup>3</sup>, Florian J. Raimann <sup>3</sup>, Florian Piekarski <sup>3</sup>, Patrick Meybohm <sup>4</sup>, Sebastian Hottenrott <sup>4</sup>, Corinna Froehlich <sup>4</sup>, Donat R. Spahn <sup>1</sup>, Christoph B. Noethiger <sup>1</sup>, David W. Tscholl <sup>1</sup> and Tadzio R. Roche <sup>1</sup>

<sup>1</sup> Institute of Anesthesiology, University of Zurich and University Hospital Zurich, Zurich, Switzerland

<sup>2</sup> Departments of Epidemiology and Biostatistics, Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Zurich, Switzerland

<sup>3</sup> Department of Anesthesiology, Intensive Care Medicine, and Pain Therapy, University Hospital Frankfurt, Goethe University Frankfurt, Frankfurt, Germany.

<sup>4</sup> Department of Anesthesiology, Intensive Care, Emergency, and Pain Medicine, University Hospital Wuerzburg, Wuerzburg, Germany.

\* Correspondence: author: David Tscholl, david.tscholl@usz.ch

**Abstract:** **Background:** Machine learning can analyze vast amounts of data and make predictions for events in the future. Our group created machine learning models for vital sign predictions. To transport the information of these predictions without numbers and numerical values and make them easily usable for human caregivers, we aimed to integrate them into the Philips Visual-Patient-avatar, an avatar-based visualization of patient monitoring. **Methods:** We conducted a computer-based simulation study with 70 participants in three European university hospitals. We validated the vital sign prediction visualizations by testing their identification by anesthesiologists and intensivists. Each prediction visualization consisted of a condition (e.g., blood pressure low) and an urgency (a visual indication of the timespan in which the condition is expected to occur). To obtain qualitative user feedback, we also conducted standardised interviews and derived statements that participants later rated in an online survey. **Results:** The mixed logistic regression model showed 77.9% (95%CI 73.2-82.0%) correct identification of prediction visualizations (i.e., condition and urgency both correctly identified) and 93.8% (95%CI 93.7-93.8%) for conditions only (i.e., without considering urgencies). Forty-nine of 70 participants completed the online survey. The online survey participants agreed that the prediction visualizations were fun to use (32/49, 65.3%), and that they could imagine working with them in the future (30/49, 61.2%). They also agreed that identifying the urgencies was difficult (32/49, 65.3%). **Conclusions:** This study found that care providers correctly identified >90% of the conditions (i.e., without considering urgencies). The accuracy of identification decreased when considering urgencies in addition to conditions. Therefore, in future development of the technology, we will focus on either only displaying conditions (without urgencies) or improve the visualizations of urgency to enhance usability for human users.

**Keywords:** avatar; machine learning; monitoring; predictive models; visual patient; vital sign predictions

## 1. Introduction

Vast amounts of data are being generated daily within healthcare, especially in electronic anesthesia records, where among other data, continuous patient monitoring data is stored. The ever-increasing use of this data will fundamentally change and improve the way medical care will be practiced in the future [1-3]. A pressing challenge is to adequately process the data so that caregivers

can make evidence-based decisions for the benefit of patients [1]. Machine learning (ML) can curate and analyze large amounts of data, identify the underlying logic and generate models that can accurately recognize a situation or predict a future state [4,5]. Predictive ML models have already been developed for various fields of medicine [6,7]. However, a significant gap exists between the amount of developed models, clinically tested applications, and commercially available products [7].

There are several reasons why ML models do not deliver the expected performance in clinical trials [8,9]. One of which is a lack of trust of the users in the models [6,10,11]. To increase trust, clinically meaningful models should be developed with good unbiased data and not patronize the users but support them in their clinical work [9,12]. An integral part of such a clinically meaningful model is the presentation of information without imposing additional cognitive load on the user [13]. A decision support tool that uses a ML model should not lead to alarm fatigue or increased workloads but provide actionable advice that fits into existing workflows [11].

To make ML models we developed for vital sign predictions in surgical patients clinically meaningful and usable, we developed a user-centered, patient avatar-based graphical representation to visualize vital sign predictions. These visualizations are an extension to Visual Patient (VP), an avatar-based patient monitoring technology [14]. VP has been available in Europe since 2023 as Philips Visual-Patient-avatar. Studies reported that healthcare providers were able to retrieve more vital signs with higher diagnostic confidence and lower perceived workload when using VP rather than wave- and number-based monitoring, allowing them to get a comprehensive picture of the patient's condition more quickly [14,15]. Additionally, care providers positively reviewed the technology and found it intuitive and easy to learn and use [16].

The project's objective is to implement vital sign predictions into the VP (provisional name VP Predictive). To achieve this goal the project aims to integrate the front-end - i.e., the way predictions are presented to the users - with the back-end - i.e., the ML models calculating the predictions.

In the present study, we report the validation process of the front-end. Specifically, we aimed to determine how accurately users identify the different vital sign prediction visualizations after a short educational video. The development and validation process of the back-end ML models is the subject of a separate study.

## 2. Methods

A declaration of non-jurisdiction (BASEC Nr. Req-2022-00302) was issued by the Cantonal Ethics Committee, Zurich, Switzerland. Due to the study's exemption from the Human Research Act, ethical approval was not required for the German study centers. Participation was voluntary and without any financial compensation. All participants signed a consent for the use of their data. In reporting the study, we followed the Guidelines for Reporting Simulation Research in Health Care, an extension of the CONSORT and STROBE statements [17].

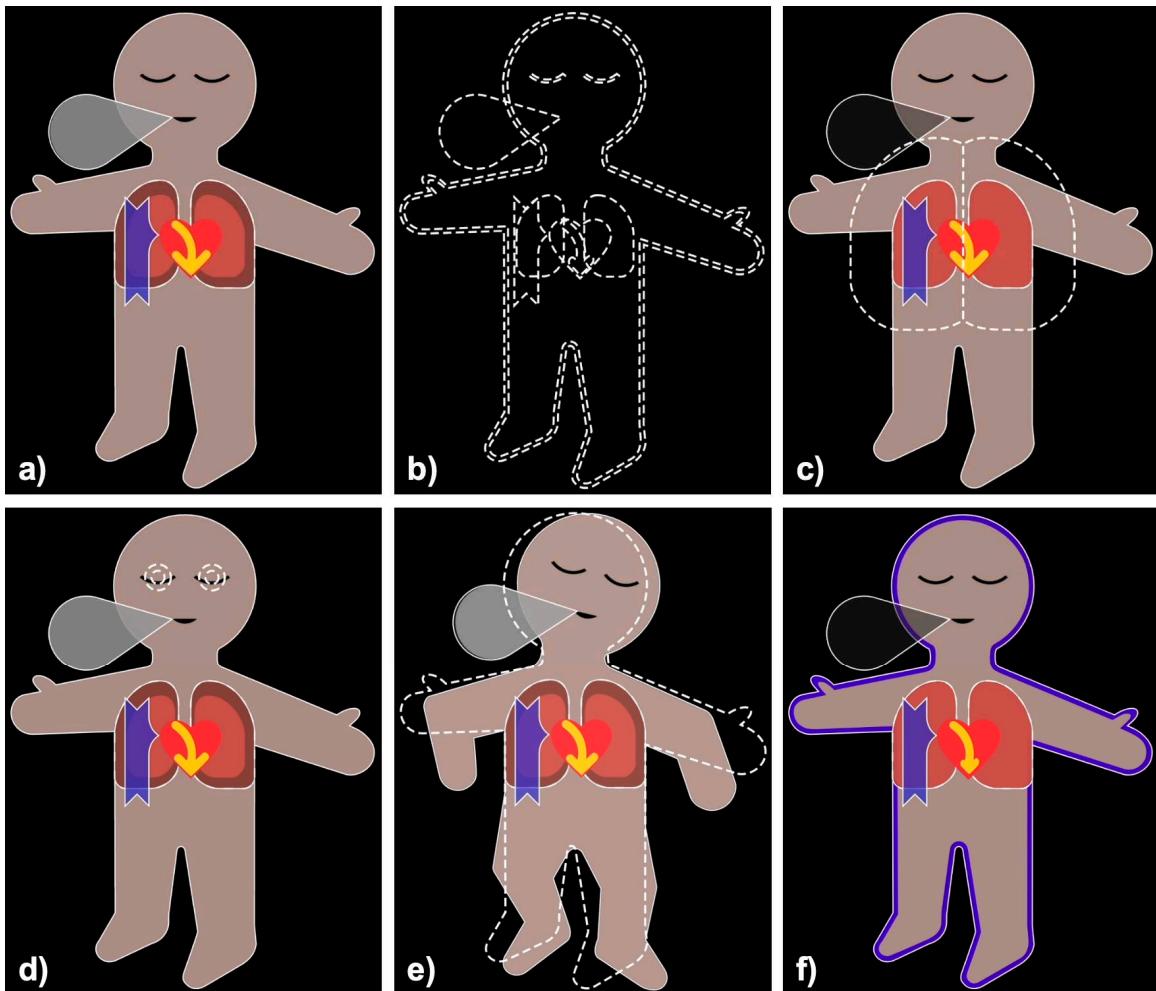
### 2.1. Study Design and Population

We conducted an investigator-initiated, prospective, multi-center, computer-based simulation study at the University Hospitals of Zurich, Frankfurt, and Wuerzburg. The study consisted of three parts. First, we validated the prediction visualizations by testing their identification by physicians. We included senior and resident physicians employed in the study centers' anesthesia or intensive care departments according to availability. Following this part, we invited participants from Frankfurt and Wuerzburg to take part in face-to-face, standardized interviews. From the interview transcripts, we identified key topics, and derived representative statements. In the third study part, the participants from all three centers rated these statements on Likert scales.

### 2.2. VP and VP Predictive

VP is a user-centered visualization technology specifically developed to improve situation awareness (**Video 1**). It creates an animated avatar of the patient to visually display various vital signs using real-time according to the real-time conventional monitoring data.

VP Predictive was developed as an add-on to VP, with the goal of integrating vital sign predictions into the standard VP. A prediction consists of a condition and an urgency. The condition signals which vital sign is predicted to change and in which direction (low/high), while the urgency gives the time horizon in which this change is expected to occur. The VP Predictive educational video ([Video 2](#)) and Figure 1 explain the technology.



**Figure 1.** Visual Patient and Visual Patient Predictive: (a) Visual Patient displays vital signs in the form of colored visualizations, (b) Visual Patient Predictive uses the same visualizations as blank figures with dashed borders. Images (c) to (f) show examples where tidal volume (c), Bispectral index (d) and Train of Four ratio (e) are predicted to become high, and oxygen saturation (f) is predicted to become low, respectively.

### 2.2.1. Condition

There are 22 condition visualizations, which are based on the original VP visualizations. These conditions are displayed as blank visualizations with white dashed borders and superimposed on the VP. The only exception to this display method is oxygen saturation, for which a “low” condition is shown by coloring in blue the blood pressure shadow of the original VP.

### 2.2.2. Urgency

There are three different urgencies: urgent, intermediate, and non-urgent. For an urgent prediction, the corresponding condition is shown for 3.5s every 7s and flashes during the display. An intermediate urgency prediction is shown for 3.5s every 14s and is not flashing. Finally, a non-urgent prediction is shown for 3.5s every 28s and is partially transparent. This way, a more urgent prediction is displayed more frequently than a less urgent one. The additional flashing (urgent) and

transparency (non-urgent) are designed to allow users to distinguish the different urgencies upon first viewing.

### 2.3. Study Procedure

We conducted a computer-based simulation study followed by standardized interviews and an online survey.

#### 2.3.1. Part I: Simulation Study

Participants were welcomed into a quiet room. After a short session briefing, and completing a sociodemographic survey, we showed the participants a video explaining VP (**Video 1**). Afterward, participants had the opportunity to practice on a Philips Visual-Patient-avatar simulator for up to 5 minutes. Afterward, an educational video explaining VP Predictive was shown (**Video 2**).

During the simulation, each participant was shown 33 videos. Each video displayed a standard VP with all vital signs in the normal range, along with an overlaid prediction visualization containing a single condition and urgency. To provide each participant with a randomized set of 33 videos and to ensure that each video was equally represented, we first created randomized sets of 66 videos (3 urgencies x 22 conditions). Then, each set was split in two (videos 1 to 33 and 34 to 66) and watched in sequence by the participants. During the videos, the participants were asked to select the shown condition (22 possible answers) and urgency (3 possible answers). We stopped the video as soon as the participant had answered, but at the latest after one minute. After the participant had completed all questions, we played the next video in the set. All data were collected on an Apple iPad (Apple Inc., Cupertino, CA, USA) using the app iSurvey (Harvestyourdata.org, Wellington, New Zealand) [18].

#### 2.3.2. Part II: Standardized Interviews

After a short break, we conducted a standardized interview with participants from Frankfurt and Wuerzburg. The question was: "What do you think about the VP Predictive visualizations?". The answers were recorded using an Apple iPhone and later automatically transcribed using Trint (Trint Limited, London, UK). The transcripts were then manually checked for accuracy and translated into English using DeepL (DeepL SE, Cologne, Germany). After manually checking the translation, we divided the text into individual statements for analysis. Using the template approach, we developed a coding tree [19]. Two study authors independently coded each statement. Differences in coding were discussed, and a joint coding per statement was agreed upon.

#### 2.3.3. Part III: Online Survey

Based on the interview results, we created six statements on recurring topics to be rated using Likert scales in an online survey. This survey was designed using Google Forms (Google LLC, Mountain View, CA, USA) and sent by email to all participants of study part I. The survey remained active for three weeks in July-August 2022. Halfway through this period, a single reminder email was sent.

### 2.4. Outcomes

#### 2.4.1. Part I: Simulation Study

We defined *correct prediction identification* as the primary outcome. If participants correctly identified both condition and urgency, we counted this as correctly identifying the prediction. As secondary outcomes, we chose *correct condition identification* and *correct urgency identification*, defined as the correctly identified condition and urgency, respectively. In addition, we analyzed the 22 conditions and the 3 urgencies individually.

#### 2.4.2. Part II and III: Standardized Interviews and Online Survey

For the standardized interviews, we analyzed the distribution of individual statements within the topics of the coding tree. For the online survey, we analyzed the distribution of the answers on the 5-point Likert scale for each statement (from "strongly disagree" to "strongly agree").

#### 2.5. Statistical Analysis

For descriptive statistics, we show medians and interquartile ranges for continuous data and numbers and percentages for categorical data.

##### 2.5.1. Part I: Simulation Study

We used mixed logistic regression models with just an intercept to estimate the correct prediction, condition and urgency identification while considering that we had repeated, non-independent measurements from each study participant. The estimates are given as percentages with 95% confidence intervals (95%CI). For estimates by condition, we added the condition information to the aforementioned model. We used a mixed logistic regression model to see if there was a learning effect by including the number of the respective question (between 1 and 33). Estimates of this model are given as odds ratios (OR).

##### 2.5.2. Part II and III: Standardized Interviews and Online Survey

In study part II, we assessed the agreement of the two coders prior to consensus by calculating the interrater reliability using Cohen's Kappa. In study part III, we used the Wilcoxon matched-pairs signed-rank test to evaluate whether the answers significantly deviated from neutral. We used Microsoft Word, Microsoft Excel (Microsoft Corporation, Redmond, WA, USA), and R version 4.2.0 (R Foundation for Statistical Computing, Vienna, Austria) to manage and analyze our data. We used GraphPad Prism version 9.4.1 (GraphPad Software Inc., San Diego, CA, USA) to generate the figures. We considered a p-value <0.05 to be statistically significant.

##### 2.5.3. Sample Size Calculation

To assess the appropriate sample size for the simulation study, we conducted a pilot study with six participants at the University Hospital Zurich. Correct prediction identification was 94.4%. Considering that these participants were already familiar with VP (but did not know VP Predictive), we calculated the sample size based on a true proportion of 90%. In this case, 70 participants are needed to construct a 95%CI for an estimated proportion that extends no more than 10% in either direction.

### 3. Results

We recruited 70 anesthesiologists and intensive care physicians in April-May 2022. All participants completed the simulation study. Twenty-one of the 70 participants (30.0%) gave an interview and 49 participants (70.0%) completed the online survey. Table 1 shows the study and participants characteristics.

**Table 1.** Participants and study characteristics. USZ, University Hospital Zurich; UKW, University Hospital Wuerzburg; KGU, University Hospital Frankfurt; IQR, interquartile range.

	Part I (Simulation Study)	Part II (Standardized Interviews)	Part III (Online Survey)
<b>Participants characteristics</b>			
Participants, n	70	21	49
Participants from USZ, n (%)	35 (50)	0 (0)	
Participants from UKW, n (%)	18 (26)	15 (71)	

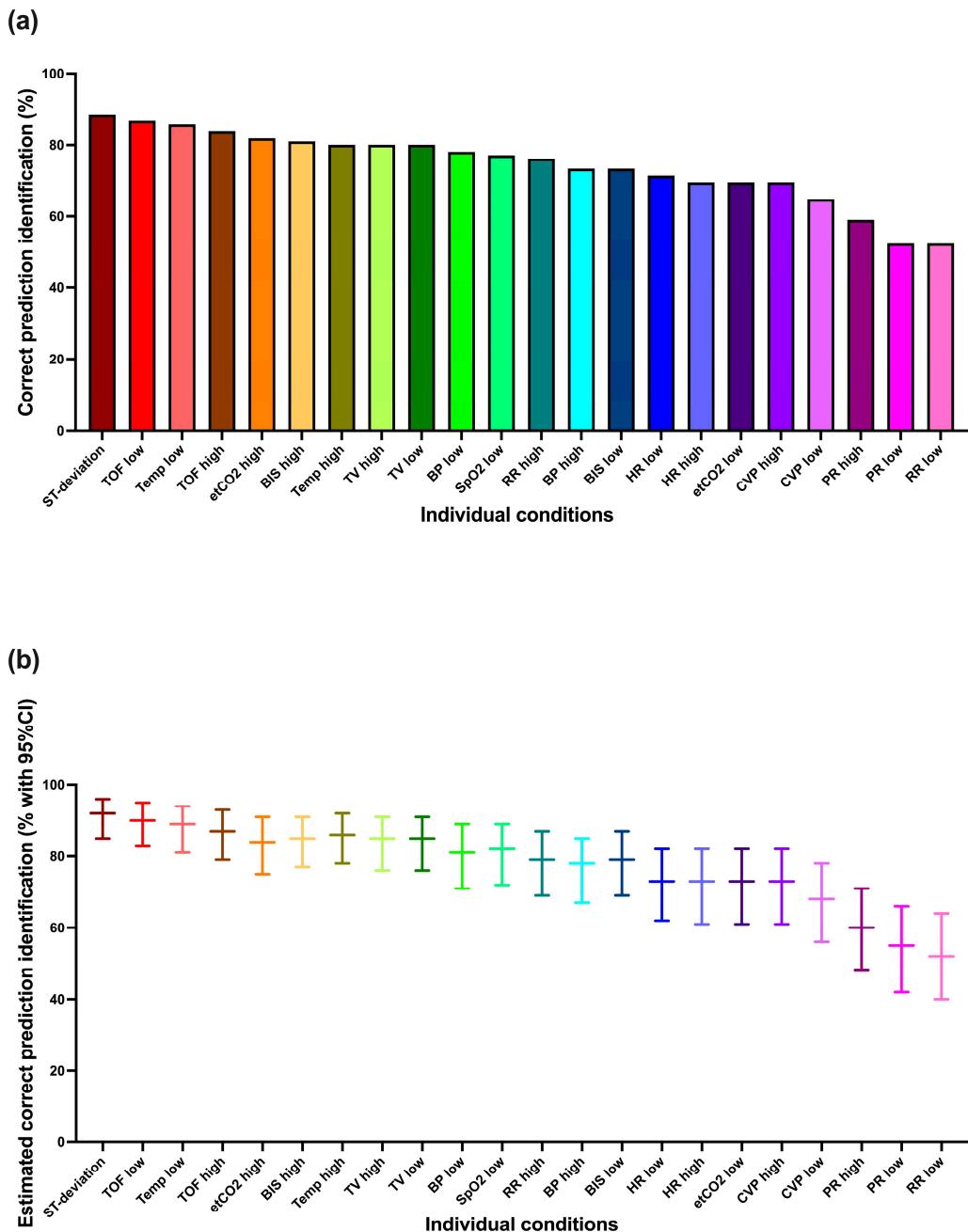
Participants from KGU, n (%)	17 (24)	6 (29)	
Gender female, n (%)	42 (60)	15 (71)	
Resident physicians, n (%)	56 (80)	17 (81)	34 (69)
Staff physicians, n (%)	14 (20)	4 (19)	15 (31)
Age (years), median (IQR)	31 (28-35)	33 (27.5-35.5)	34 (28-37)
Work experience (years), median (IQR)	3.5 (1-6)	3 (1.5-8)	4 (2-7)
Previous experience with Visual Patient, n (%)	19 (27)	4 (19)	
<b>Study characteristics</b>			
Different conditions studied, n	22		
Different urgencies studied, n	3		
Different predictions studied, n	66		
Randomly selected predictions per participant, n	33		

### 3.1. Part I: Simulation Study

#### 3.1.1. Correct Prediction Identification

In total 1716/2310 (74.3%) prediction visualizations (condition and urgency) were correctly identified. The mixed logistic regression model showed a slightly higher percentage (77.9%, 95%CI 73.2-82.0%).

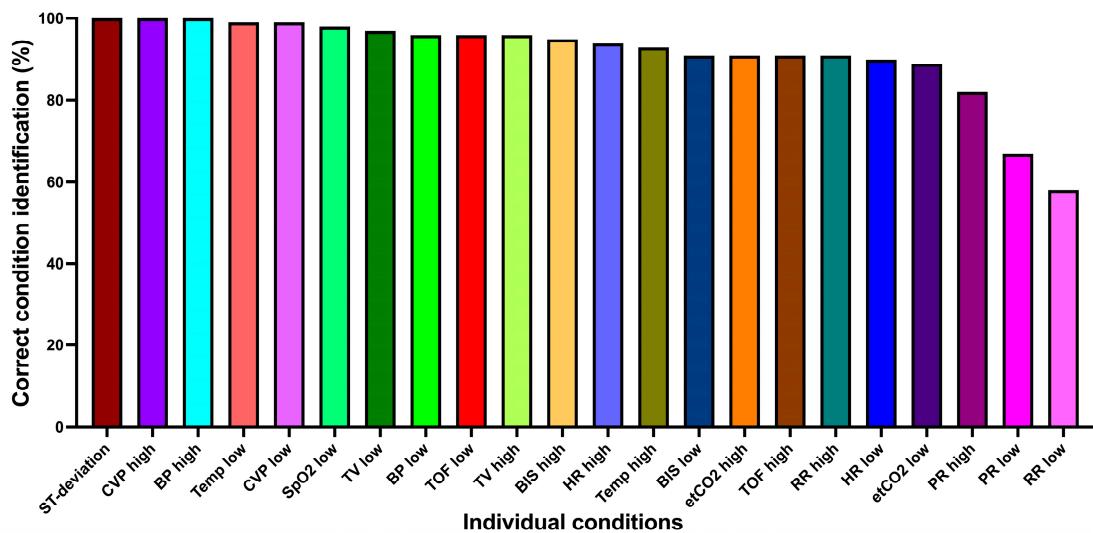
Figure 2 shows these results for each condition individually. It is apparent that not all conditions were identified equally well. The best-identified conditions showed close to 90% correct prediction identification, whereas a few showed less than 60% correct prediction identification. The mixed logistic regression model-based estimations tended to be a few percentage points higher.



**Figure 2.** Correct prediction identification (correctly identified condition and urgency) for each condition individually: (a) the percentages of correct prediction identification, (b) the estimates based on the mixed logistic regression model. ST-deviation, ST-segment deviation; TOF, Train of Four ratio; Temp, body temperature; etcO2, end-expiratory carbon dioxide concentration; BIS, Bispectral index; TV, tidal volume; BP, blood pressure; SpO2, oxygen saturation; RR, respiratory rate; HR, heart rate; CVP, central venous pressure; PR, pulse rate.

### 3.1.2. Correct Condition Identification

Considering conditions alone (without urgencies), 2117/2310 (91.7%) were correctly identified. Mixed logistic regression model: 93.8% (95%CI 93.7-93.8%). Figure 3 shows the correct condition identification for each condition individually. Most conditions were very well identified, with two exceptions: pulse rate low (68.6%), respiratory rate low (58.1%).



**Figure 3.** Correct condition identification for each condition individually. ST-deviation, ST-segment deviation; CVP, central venous pressure; BP, blood pressure; Temp, body temperature; SpO<sub>2</sub>, oxygen saturation; TV, tidal volume; TOF, Train of Four ratio; BIS, Bispectral index; HR, heart rate; etCO<sub>2</sub>, end-expiratory carbon dioxide concentration; RR, respiratory rate; PR, pulse rate.

### 3.1.3. Correct Urgency Identification

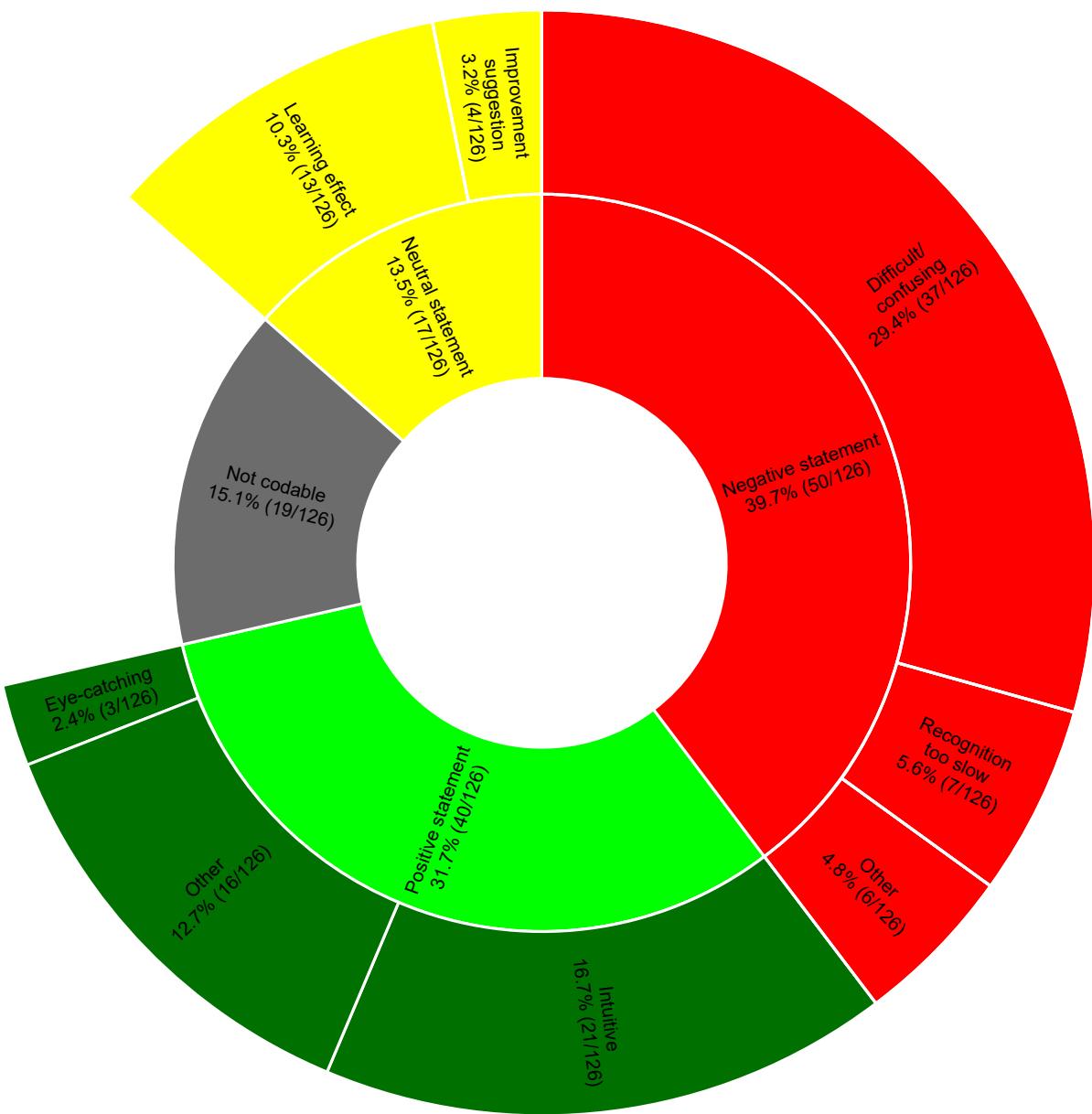
Urgency (without condition) was correctly identified in 1855/2310 (80.3%) cases. Mixed logistic regression model: 84.0% (95%CI 80.2-87.1%). Considering each urgency individually, the urgent one was correctly identified 629/770 (81.7%) times, the intermediate one 577/770 (74.9%) times, and the non-urgent one 649/770 (84.3%) times.

### 3.1.4. Learning Effect

The mixed logistic regression model showed a significant learning effect on correct prediction identification, with the odds of correctly identifying the predictions increasing by 3% for each additional prediction shown (OR 1.03, 95%CI 1.02-1.04, p<0.001).

## 3.2. Part II: Standardized Interviews

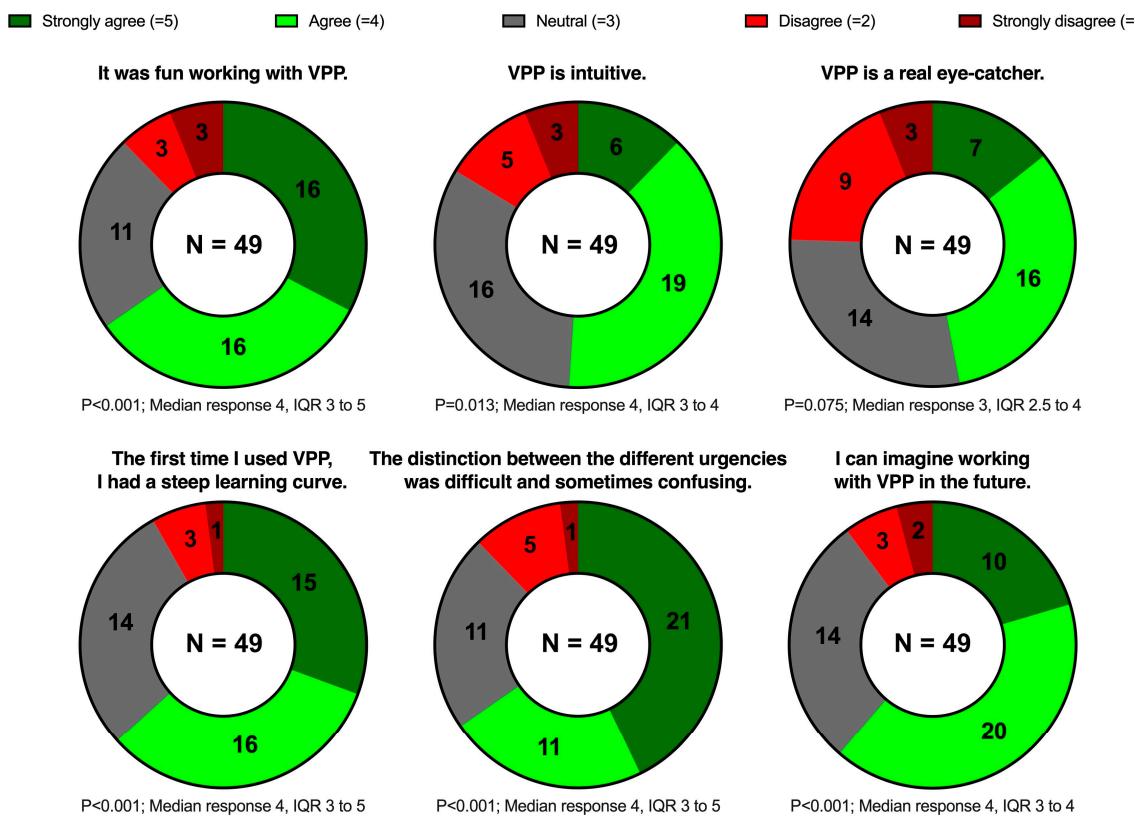
From the transcripts of the interviews, we identified 126 different statements. At first coding, the two independent raters agreed on the classification of 83.3% of the statements (105/126), with a Cohen's Kappa of 0.8. Most of the positive comments considered VP Predictive intuitive. Negative comments mainly concerned identification difficulties, especially of the different urgencies. Several participants noted a learning effect during the session or believed an additional learning effect could be achieved by using VP Predictive more frequently. Figure 4 shows the coding tree in detail. Note that 15.1% of the statements were not codable; these primarily represented statements not relevant to the posed question.



**Figure 4.** Distribution of the statements within the topics of the coding tree. We show percentages and numbers.

### 3.3. Part III: Online Survey

The questionnaire was completed by 70.0% of the invited participants (49/70). Most of the participants agreed or strongly agreed that VP Predictive was fun to use (32/49, 65.3%) and intuitive (25/49, 51.0%); many of them also agreed or strongly agreed that it was eye-catching (23/49, 46.9%). Almost two-thirds (32/49, 65.3%) agreed or strongly agreed that the urgency identification was difficult. Nevertheless, most participants (31/49, 63.3%) agreed or strongly agreed that they had a steep learning curve during the study session, and only very few (5/49, 10.2%) disagreed or strongly disagreed that they could imagine working with VP Predictive in the future. Figure 5 shows these results in detail.



**Figure 5.** Doughnut charts showing the statements and the distribution of the answers on the 5-point Likert scale. The results are shown as numbers. We calculated p-values using the Wilcoxon signed-rank test to determine whether the responses significantly deviated from neutral. VPP, Visual Patient Predictive; IQR, interquartile range.

#### 4. Discussion

We sought to investigate VP Predictive. This technology is an extension of the original VP designed to easily represent vital sign predictions with little cognitive load. Participants correctly identified both condition and urgency in the prediction visualizations in almost three quarters of the cases (74.3%). The majority found VP Predictive to be enjoyable to use, with 65.3% rating it as fun and only 16.3% considering it not intuitive.

In this study, correct condition identification was high (91.7%). Correct prediction identification (i.e., correct identification of both condition and urgency) was not equally high percentage (74.3%). This finding is also in line with the participants' subjectively perceived difficulty in identifying the different urgencies, expressed during the interviews and in the survey. The different urgencies aimed to provide vital sign predictions with an expected occurrence time. For example, the prediction for low blood pressure could be displayed with three different urgencies (e.g., 1, 5, or 20 minutes). The differences in the percentages of correct identification become understandable when considering that the identification of conditions alone involved the interpretation of less visual information than when additional urgencies also needed to be identified.

Interestingly, the primary outcome result in the pilot study differed significantly from the one in the actual study (pilot 94.4% vs. study 74.3% correct prediction identification). One possible explanation for this difference is that the pilot study cohort was already familiar with the original VP visualizations (although not with the prediction visualizations) and thus had fewer new things to learn before the study. In comparison, the majority of the actual study participants were encountering VP for the first time. This raises the question of whether a longer familiarization period could have improved the percentage of correct urgency identification, and thus also that of correct prediction identification.

This hypothesis is supported by the learning effect that we confirmed quantitatively and from the participants' feedback. Intuitiveness and learning ease are essential for accepting new technologies and crucial for their successful clinical introduction [16]. In our case, these requirements seem to be achieved, as the majority of the survey participants could imagine working with VP Predictive in the future.

Considering our study results, we believe that - with some modifications - VP Predictive may have the potential to display vital sign predictions generated by ML models in a way healthcare professionals can understand and translate into direct actions. VP Predictive is intended to guide users attention. When alerted by a prediction, caregivers should ultimately consider all available information and decide on an appropriate response (e.g., fluids or vasopressors in case of a low blood pressure prediction).

#### 4.1. Strengths and Limitations

Like all computer-based studies, this study has particular strengths and limitations. First, the conditions under which it took place differ from the clinical reality, in which many more factors are present [20]. In addition, participants evaluated only videos in which the VP was shown in a physiological state and in which exactly one prediction was shown at a time. Such scenarios differ from the more complex clinical reality, so studies in more realistic settings will be needed to evaluate the true clinical value (e.g., a high-fidelity simulation study) [21].

At the same time, a computer-based study also has advantages over a real-life study. First, it allows completely new technologies to be tested without patient risks [22]. It also standardizes the study conditions, an essential prerequisite for minimizing possible bias due to external disturbances.

Another strength of our study is that it was multicenter and multinational, allowing the results to be generalized to a certain extent. Based on the pilot study, the trial was adequately powered, however, the participants' selection was based on availability during working hours and, therefore, not random.

### 5. Conclusion

Despite promising results and feedback, the current Visual Patient Predictive visualizations need some modifications followed by further high-fidelity simulation studies to test its suitability for the intended task of displaying vital sign predictions to healthcare providers in an easily understandable way. In this study care providers correctly identified >90% of the conditions (i.e., without considering urgencies). The percentage of correct identification decreased when considering urgencies in addition to conditions. Therefore, in future development of the technology, we will focus on either only displaying conditions (without urgencies) or improve the visualizations of urgency to enhance usability for human users.

**Authors' contributions** **Conception and design:** AM, CBN, DRS, DWT, TRR **Data collection:** AM, DH, GS, FJR, FP, SH, CF, TRR **Data analysis and interpretation:** AM, DH, JB, DWT, TRR **Writing the article or critical revision of the manuscript:** AM, DH, GS, JB, KZ, FJR, PM, CBN, DRS, DWT, TRR **Approval of the final version:** AM, DH, GS, JB, KZ, FJR, FP, PM, SH, CF, CBN, DRS, DWT, TRR.

**Acknowledgments:** We would like to thank all the participants and all those who made this study possible for their time and support.

**Conflicts of interest:** DWT, CBN, AM and TRR are inventors of Visual Patient Predictive technology, for which the University of Zurich and Koninklijke Philips N.V. hold patent applications and design protections. Joint-development and licensing agreements exist with Philips Medizin Systeme Böblingen GmbH, Böblingen, Germany; Koninklijke Philips N.V., Amsterdam, The Netherlands; Philips Research/Philips Electronics Nederland BV, Eindhoven, The Netherlands; and Philips USA, Cambridge, MA, USA. Within the framework of these agreements, DWT, CBN and TRR receive travel support, lecturing and consulting honoraria, and DWT, CBN, AM and TRR may potentially receive royalties in the event of successful commercialization. DWT and CBN are designated inventors of Visual-Patient-avatar, for which the University of Zurich and Koninklijke Philips N.V. hold various patents, design protections, and trademarks. There are cooperation and licensing agreements with Philips Medizin Systeme Böblingen GmbH, Böblingen, Germany; Koninklijke Philips N.V.,

Amsterdam, The Netherlands; Philips Research/Philips Electronics Nederland BV, Eindhoven, The Netherlands; Philips USA, Cambridge, MA, USA. Under these agreements, DWT receives research funding and DWT and CBN receive travel support, lecturing honoraria, and may receive royalties in the event of successful commercialization. CBN and DWT received travel support, lecturing, and consulting honoraria from Instrumentation Laboratory – Werfen, Bedford, MA, USA. Additionally, DWT received travel support, lecturing, and consulting honoraria from the Swiss Foundation for Anaesthesia Research in Zurich, Switzerland, and the International Symposium on Intensive Care and Emergency Medicine in Brussels, Belgium. DWT also holds a position on the Philips Patient Safety Advisory Board. DRS academic department is receiving grant support from the Swiss National Science Foundation, Berne, Switzerland, the Swiss Society of Anesthesiology and Perioperative Medicine (SSAPM), Berne, Switzerland, the Swiss Foundation for Anesthesia Research, Zurich, Switzerland, Vifor SA, Villars-sur-Glâne, Switzerland and Vifor (International) AG, St. Gallen, Switzerland. DRS is co-chair of the ABC-Trauma Faculty, sponsored by unrestricted educational grants from Novo Nordisk Health Care AG, Zurich, Switzerland, CSL Behring GmbH, Marburg, Germany, LFB Biomédicaments, Courtaboeuf Cedex, France and Octapharma AG, Lachen, Switzerland. DRS received honoraria / travel support for consulting or lecturing from: Danube University of Krems, Austria, US Department of Defense, Washington, USA, European Society of Anesthesiology and Intensive Care, Brussels, BE, Korean Society for Patient Blood Management, Seoul, Korea, Korean Society of Anesthesiologists, Seoul, Korea, Network for the Advancement of Patient Blood Management, Haemostasis and Thrombosis, Paris, France, Alexion Pharmaceuticals Inc., Boston, MA, Bayer AG, Zürich, Switzerland, B. Braun Melsungen AG, Melsungen, Germany, CSL Behring GmbH, Hattersheim am Main, Germany and Berne, Switzerland, Celgene International II Sàrl, Couvet, Switzerland, Daiichi Sankyo AG, Thalwil, Switzerland, Haemonetics, Braintree, MA, USA, Instrumentation Laboratory (Werfen), Bedford, MA, USA, LFB Biomédicaments, Courtaboeuf Cedex, France, Merck Sharp & Dohme, Kenilworth, New Jersey, USA, Novo Nordisk Health Care AG, Zurich, Switzerland, PAION Deutschland GmbH, Aachen, Germany, Pharmacosmos A/S, Holbaek, Denmark, Pfizer AG, Zürich, Switzerland, Pierre Fabre Pharma, Alschwil, Switzerland, Portola Schweiz GmbH, Aarau, Switzerland, Roche Diagnostics International Ltd., Reinach, Switzerland, Sarstedt AG & Co., Sevelen, Switzerland and Nümbrecht, Germany, Shire Switzerland GmbH, Zug, Switzerland, Tem International GmbH, Munich, Germany, Vifor Pharma, Munich, Germany, Neuilly sur Seine, France and Villars-sur-Glâne, Switzerland, Vifor (International) AG, St. Gallen, Switzerland, Zuellig Pharma Holdings, Singapore, Singapore. KZ academic department received support from B. Braun Melsungen, CSL Behring, Fresenius Kabi, and Vifor Pharma for the implementation of Frankfurt's Patient Blood Management program. KZ has received honoraria for participation in advisory board meetings for Haemonetics and Vifor and received speaker fees from CSL Behring, Masimo, Pharmacosmos, Boston Scientific, Salus, iSEP, Edwards and GE Healthcare. He is the Principal Investigator of the EU-Horizon 2020 project ENVISION (Intelligent plug-and-play digital tool for real-time surveillance of COVID-19 patients and smart decision-making in Intensive Care Units) and Horizon Europe 2021 project COVend (Biomarker and AI-supported FX06 therapy to prevent progression from mild and moderate to severe stages of COVID-19). FJR received speaker fees from CSL Behring, Helios Germany, university hospital Würzburg and Keller Medical GmbH. FJR received financial support by HemoSonics LLC, pharma-consult Petersohn, LifeSystems and Boehringer Ingelheim. FP received honoraria from Pharmacosmos for scientific lectures. The other authors report no conflicts of interest regarding this paper.

## References

1. Rajkomar, A., J. Dean, and I. Kohane, *Machine Learning in Medicine*. N Engl J Med, 2019. **380**(14): p. 1347-1358.
2. Shilo, S., H. Rossman, and E. Segal, *Axes of a revolution: challenges and promises of big data in healthcare*. Nat Med, 2020. **26**(1): p. 29-38.
3. Pastorino, R., et al., *Benefits and challenges of Big Data in healthcare: an overview of the European initiatives*. Eur J Public Health, 2019. **29**(Supplement\_3): p. 23-27.
4. Shah, N.H., A. Milstein, and D.S. Bagley Ph, *Making Machine Learning Models Clinically Useful*. JAMA, 2019. **322**(14): p. 1351-1352.
5. Gambus, P.L. and S. Jaramillo, *Machine learning in anaesthesia: reactive, proactive... predictive!* Br J Anaesth, 2019. **123**(4): p. 401-403.
6. Topol, E.J., *High-performance medicine: the convergence of human and artificial intelligence*. Nat Med, 2019. **25**(1): p. 44-56.
7. Ben-Israel, D., et al., *The impact of machine learning on patient care: A systematic review*. Artif Intell Med, 2020. **103**: p. 101785.

8. Emanuel, E.J. and R.M. Wachter, *Artificial Intelligence in Health Care: Will the Value Match the Hype?* JAMA, 2019. **321**(23): p. 2281-2282.
9. Gianfrancesco, M.A., S. Tamang, J. Yazdany, and G. Schmajuk, *Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data.* JAMA Intern Med, 2018. **178**(11): p. 1544-1547.
10. Hashimoto, D.A., et al., *Artificial Intelligence in Anesthesiology: Current Techniques, Clinical Applications, and Limitations.* Anesthesiology, 2020. **132**(2): p. 379-394.
11. *Towards trustable machine learning.* Nat Biomed Eng, 2018. **2**(10): p. 709-710.
12. Andaur Navarro, C.L., et al., *Risk of bias in studies on prediction models developed using supervised machine learning techniques: systematic review.* BMJ, 2021. **375**: p. n2281.
13. Ehrmann, D.E., et al., *Evaluating and reducing cognitive load should be a priority for machine learning in healthcare.* Nat Med, 2022. **28**(7): p. 1331-1333.
14. Tscholl, D.W., et al., *Situation Awareness-Oriented Patient Monitoring with Visual Patient Technology: A Qualitative Review of the Primary Research.* Sensors (Basel), 2020. **20**(7).
15. Roche, T.R., et al., *Avatar-based patient monitoring in critical anaesthesia events: a randomised high-fidelity simulation study.* Br J Anaesth, 2021. **126**(5): p. 1046-1054.
16. Tscholl, D.W., et al., *User perceptions of avatar-based patient monitoring: a mixed qualitative and quantitative study.* BMC Anesthesiol, 2018. **18**(1): p. 188.
17. Cheng, A., et al., *Reporting Guidelines for Health Care Simulation Research: Extensions to the CONSORT and STROBE Statements.* Simul Healthc, 2016. **11**(4): p. 238-48.
18. Tscholl, D.W., M. Weiss, D.R. Spahn, and C.B. Noethiger, *How to Conduct Multimethod Field Studies in the Operating Room: The iPad Combined With a Survey App as a Valid and Reliable Data Collection Tool.* JMIR Res Protoc, 2016. **5**(1): p. e4.
19. Brooks, J., S. McCluskey, E. Turley, and N. King, *The Utility of Template Analysis in Qualitative Psychology Research.* Qual Res Psychol, 2015. **12**(2): p. 202-222.
20. Kurup, V., V. Matei, and J. Ray, *Role of in-situ simulation for training in healthcare: opportunities and challenges.* Curr Opin Anaesthesiol, 2017. **30**(6): p. 755-760.
21. Merry, A.F., et al., *Retesting the Hypothesis of a Clinical Randomized Controlled Trial in a Simulation Environment to Validate Anesthesia Simulation in Error Research (the VASER Study).* Anesthesiology, 2017. **126**(3): p. 472-481.
22. Moorthy, K., C. Vincent, and A. Darzi, *Simulation based training.* BMJ, 2005. **330**(7490): p. 493-4.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.