

Original Paper

# An Explanation User Interface for Artificial Intelligence–Supported Mechanical Ventilation Optimization for Clinicians: User-Centered Design and Formative Usability Study

Ian-C Jung, MSc; Maria Zerlik, MSc; Katharina Schuler, MSc; Martin Sedlmayr, Prof Dr; Brita Sedlmayr, PhD

Institute for Medical Informatics and Biometry, Faculty of Medicine and University Hospital Carl Gustav Carus, TUD Dresden University of Technology, Dresden, Germany

**Corresponding Author:**

Ian-C Jung, MSc

Institute for Medical Informatics and Biometry

Faculty of Medicine and University Hospital Carl Gustav Carus

TUD Dresden University of Technology

Postfach 113

Fetscherstraße 74

Dresden, 01307

Germany

Phone: 49 351 458 87 7711

Email: [ian-christopher.jung@tu-dresden.de](mailto:ian-christopher.jung@tu-dresden.de)

## Abstract

**Background:** The integration of artificial intelligence (AI) into clinical decision support systems (CDSSs) for mechanical ventilation in intensive care units (ICUs) holds great potential. However, the lack of transparency and explainability hinders the adoption of opaque AI models in clinical practice. Explanation user interfaces (XUIs), incorporating explainable AI algorithms, are considered a key solution to enhance trust and usability. Despite growing research on explainable AI in health care, little is known about how clinicians perceive and interact with such explanation interfaces in high-stakes environments such as the ICU. Addressing this gap is essential to ensure that AI-supported CDSS are not only accurate but also trusted, interpretable, and seamlessly integrated into clinical workflows.

**Objective:** This study aimed to evaluate the first iteration of the design and evaluation phase of an XUI for an AI-based CDSS intended to optimize mechanical ventilation in the ICU. Specifically, it explores how different user groups—ICU nurses and physicians—perceive and prioritize explanation concepts, providing the empirical foundation for subsequent refinement iterations.

**Methods:** A midfidelity prototype was developed using the prototyping software Justinmind, based on existing guidelines, scientific literature, and insights from previous user-centered design (UCD) phases. The design process followed ISO (International Organization for Standardization) 9241-210 principles for UCD and combined qualitative and quantitative feedback to identify usability strengths, design challenges, and role-specific explanation needs. The prototype was evaluated formatively through 2 usability walkthroughs (walkthrough 1: 4 resident physicians and walkthrough 2: 4 ICU nurses), which included guided group discussions and Likert-scale assessments of explanation concepts in terms of understandability, suitability, and visual appeal.

**Results:** The XUI was structured into 2 levels: a first level displaying high-level explanations (outlier warning and output certainty) alongside the CDSS output, and a second level offering more detailed explanations (available input, feature importance, and rule-based explanation) for users seeking deeper insight. While both user groups appreciated the first level, physicians found the second level of the XUI useful, whereas ICU nurses found it overly detailed. Thus, the structure was able to address the differing needs for explanations. The layered design helped balance transparency and information overload by providing initially concise explanations and more detailed ones on demand. The evaluation further strengthened evidence for role-dependent explanation needs, suggesting that nurses prefer actionable, concise insights, whereas physicians benefit from more granular transparency information.

**Conclusions:** This study underscores the importance of UCD in designing XUIs for CDSS. It highlights the differing information needs of physicians and ICU nurses, emphasizing the value of involving users early in the development of suitable XUIs. The findings provide practical guidance for designing layered, role-sensitive explanation interfaces in critical care and form the basis for future iterative evaluations and experimental studies assessing their impact on decision-making and clinician trust.

## KEYWORDS

CDSS; clinical decision support system; XAI; explanation user interface; XUI; usability; formative evaluation; intensive care unit; ICU; explainable artificial intelligence

## Introduction

### Background

Artificial intelligence (AI) is expected to benefit the medical domain with improvements in patient care and reduced workload for medical personnel. AI models are increasingly proposed as a knowledge or inference base of clinical decision support systems (CDSSs) to achieve this. These AI-based CDSS are considered well-suited for use in the intensive care unit (ICU) due to the amount of available documented data in the ICU, but the transfer of these systems into clinical practice is still lacking [1]. AI models for which it is impossible to directly infer why they generate an output due to their complexity are called black-box models. The black-box nature of these models poses a barrier to adoption in the medical domain, as it makes it difficult for health care providers to understand, trust, or accept these models [2-4]. According to the EU (European Union) AI Act, AI-based CDSS are classified as high-risk AI systems [5]. For high-risk AI systems, the act requires the possibility of human oversight and the provision of appropriate interfaces for this supervision (EU AI Act, Chapter III, Article 13 and 14) [6]. Approaches from the field of explainable AI (XAI) are a solution to these barriers [7-11]. XAI provides additional information to the output of an AI model to make model reasoning more transparent through explanation user interfaces (XUIs). In the medical domain, XAI has been shown to influence users' trust [12,13], the perceived usefulness [14] or value [13] of a system, the medical professionals' confidence in decisions [14], and decision-making [15]. XAI research focuses on creating, optimizing, and evaluating XAI algorithms. Explanations are often designed to meet the needs of (X)AI developers while neglecting the needs of clinical end users [16]. Recently, calls for a user-centered design (UCD) of XUIs have been raised in the medical informatics domain [16-20].

The objective of this paper was to conduct the initial design and evaluation phases of the UCD process for an XUI of an AI-based CDSS intended for use in the ICU. As a use case, we selected the provision of recommendations for ventilation parameter settings in patients receiving mechanical ventilation in the ICU. Patients may undergo mechanical ventilation using ventilators—machines that assist breathing by delivering air into the lungs when they cannot breathe adequately. Another form of life support is extracorporeal membrane oxygenation, which bypasses the lungs—and sometimes the heart—by circulating the patient's blood through an external machine that oxygenates the blood and extracts carbon dioxide from it. To enable the continuous optimization of a patient's ventilation, a CDSS must be able to recommend a range of parameters. In mechanical ventilation, these include, for example, the ventilator mode, which defines the overall strategy and level of mechanical support; tidal volume, which specifies the amount of air

delivered per breath; and respiratory rate, which determines the number of breaths per minute.

The development of such AI-based CDSS is being carried out within the IntelliLung project, a European Union–funded initiative aimed at improving ventilatory care and reducing the workload of clinical staff [21]. The CDSS will generate recommendations for both mechanical ventilation and extracorporeal membrane oxygenation settings. The research presented in this paper used the IntelliLung CDSS as a real-world use case to anchor the design in clinical practice. However, the design is independent of any project-specific AI algorithm decisions and instead focuses conceptually on XUI design for such a CDSS. As Mohseni et al [22] proposed, this type of research can be conducted before selecting and developing the AI algorithms and corresponding XAI methods.

### State of the Art

Although much research on XAI focuses on the health care domain [23], to the authors' knowledge, no paper explicitly addresses the UCD of XUIs for AI-based CDSS aimed at optimizing continuous mechanical ventilation in the ICU. Several papers report on the use of AI algorithms and XAI methods from a technical perspective, for example, to predict the ventilation duration for patients with acute respiratory distress syndrome [24], to predict the disease progression of patients with acute respiratory distress syndrome [25], to predict the ventilator support for patients with COVID-19 [26], and to generate ventilator setting recommendations [27]. Hughes et al [28] report on the context and work process analysis for CDSS supporting the weaning process in the ICU to inform the overall CDSS design for weaning support, but do not focus on optimizing continuous ventilation or designing XUIs for the CDSS.

A limited number of papers address the UCD of XUIs in the ICU [14,29-31]. A preprint by Clark et al [29], published after this study concluded, investigates user requirements for using XAI in CDSS, without focusing on a specific ICU use case. Nagendran et al [30] demonstrated the feasibility of eye-tracking analysis to assess physicians' interactions with an XUI for an AI-based CDSS, although not focused on ventilation, in a simulated ICU setting. Barda et al [31] designed and evaluated an XUI for an AI-based CDSS that predicts mortality risk in a pediatric ICU. The acceptance of an XUI for a sepsis treatment CDSS by ICU clinicians was investigated by Sivaraman et al [14]. Outside the ICU setting, Schoonderwoerd et al [32] investigated pediatricians' perspectives on different explanation concepts for a diagnostic CDSS.

### Previous Research

The UCD process was conducted in accordance with DIN (Deutsches Institut für Normung) EN (Europäische Norm) ISO (International Organization for Standardization) 9241-210 [33]

to design the XUI for an AI-based CDSS that recommends ventilator settings. This paper focuses exclusively on the first iteration of the design and evaluation phases of the UCD process. These phases rely on the results of the context analysis and the specification of user requirements. The results of the first 2 phases of the UCD process are summarized to provide the necessary context.

During the context analysis, the intended use characteristics of the CDSS were defined as (1) a stand-alone CDSS, which should not replace the ventilation monitor, and (2) a CDSS that would be displayed on a stationary tablet next to the ventilation monitor of an individual patient. The identified main user groups of this CDSS were resident physicians and ICU nurses with the authority to change the patient's ventilator settings. An initial investigation into user requirements revealed conflicting opinions among future users about the importance of explanations [34]. These results led to a follow-up investigation into users' preferences for explanation, to identify which explanations to include in a first design iteration. Seventeen explanation concepts, ranging from less complex ones such as available input information to more sophisticated ones such as counterfactual explanations, were presented to the users as low-fidelity mockups in a questionnaire. This step was performed before this study to attain a manageable number of explanation concepts for the first design iteration and a formative evaluation. The selection of explanation concepts for this first iteration of the XUI was based solely on user preferences, rather than on commonly used explanation concepts. Users selected the explanation concepts, output certainty, and outlier warning to be displayed alongside the CDSS recommendations. The output certainty concept informs users about the CDSS's confidence level in its recommendation. The outlier warning alerts users if one or more input parameter values are highly unusual. Three additional explanation concepts were selected but were not designated to appear directly next to the CDSS output: available input, feature importance, and rule-based explanation. The available input explanation concept informs the user which parameter values were available when generating the recommendation. The feature importance explanation concept shows users how each parameter influenced the

recommendation. The rule-based explanation concept presents simplified rules in text or decision tree format, approximating how the recommendation was calculated. These findings served as the foundation for the XUI design phase.

## Study Scope and Research Question

### Overview

The underlying research aim of this paper is threefold to (1) design midfidelity mockups of the XUI, focusing on its overall structure and different design variants for the explanation concepts, (2) give users an impression of the XUI without fully designing each explanation concept in detail, and (3) gather early formative feedback from future users to support iterative refinement of the XUI in subsequent design phases. Therefore, this study does not aim to deliver a finalized XUI design, nor does it seek to conduct a comprehensive summative evaluation of the XUI, as this will be the focus of future iterations.

The following research questions, aligned with the design and evaluation phases, were developed to guide the investigation.

### Design Phase

The design phase focused on generating design options for the XUI, with the following research question:

- RQ1: How should the XUI of an AI-based CDSS for mechanical ventilation optimization in the ICU be designed?

### Evaluation Phase

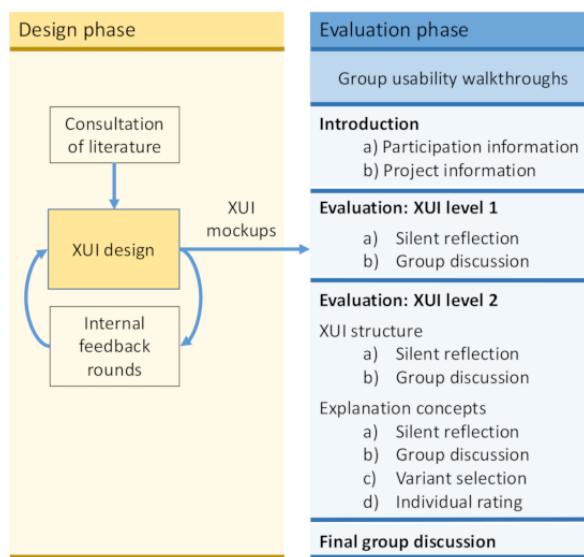
The evaluation phase focused on collecting early user feedback from clinicians, with the following research questions:

- RQ2: What are clinicians' perceptions of the proposed XUI of the AI-based CDSS for mechanical ventilation optimization?
- RQ3: What XUI design improvements should be addressed in subsequent design iterations?

## Methods

Figure 1 provides an overview of the methods used in each XUI development phase.

**Figure 1.** Overview of the methods used during the first iteration of the design and evaluation phases of the explanation user interface (XUI).



## XUI Design Phase

The design phase began with the development of a foundational knowledge base for the XUI. Relevant literature, such as norms and design guidelines, was reviewed. For this purpose, the ISO 9241 standard series on human-system interaction and human-centered design was reviewed. This foundation was expanded through a scoping review and the development of design recommendations for the user-centered XUI design for CDSS [35]. Searches for research papers related to the topic of XUIs for AI-based CDSS for the ventilation of patients in the ICU found no relevant results before the design of the XUI.

Midfidelity XUI mockups were developed using the Justinmind [36] prototyping software, based on the knowledge foundation. The design of the mockups was intended to direct users' attention to elements relevant to RQ2 and RQ3 and to prevent users from focusing on aesthetic aspects such as the color of the XUI. Therefore, mockups were designed with the following characteristics: (1) realistic variable names and content to support user familiarity; (2) realistic-looking data, for which synthetic data points were generated for each displayed input variable with a large language model (Claude Sonnet 3.5 [37]); (3) minimal use of color, used selectively to highlight system functionality; and (4) simple interactions (eg, navigation between interface levels).

A main screen for the CDSS was designed to present the XUI within its intended system context and to help users envision the future context of use for the XUI. The basic structure of the XUI and the design variants for the explanation concepts were developed based on the list of selected explanation concepts and their intended level within the XUI, which had been defined during the requirements specification phase of the UCD process. The mockups were designed iteratively: the first author created initial drafts, and members of the author team with usability expertise provided verbal feedback for improvements during multiple group meetings.

The design of the individual explanation concepts was oriented toward typical design elements used for these explanation concepts in the XAI literature. This decision was made to provide the users with a realistic representation of common explanation concepts and to keep the focus on the goal of this first evaluation iteration. We opted not to apply a more elaborate design strategy to the explanation concepts in this iteration; this will be addressed in later iterations, which typically focus on aesthetics and detailed design.

## XUI Evaluation Phase

In the evaluation phase, formative feedback was gathered from ICU physicians and nurses, focusing on the XUI's structure and explanation concepts. A formative evaluation was chosen to collect insights for improving the CDSS and understanding users' initial impressions of the CDSS. Given resource constraints—such as limited availability of clinical experts and their high workload—a pragmatic approach was taken. This evaluation is reported according to the STARE-HI (Statement on the Reporting of Evaluation studies in Health Informatics) guidelines to ensure consistency and standardization [38]. For the STARE-HI reporting checklist, refer to [Multimedia Appendix 1](#) and [39].

## Study Design

As a formative evaluation, 2 group usability walkthroughs were conducted—1 for each main user group (physicians and ICU nurses). A semistructured approach was followed, guided by a predefined protocol. Each walkthrough included (1) a presentation of the XUI, (2) individual reflection time, (3) guided group discussions, and (4) a quantitative assessment (QA) of the explanation concepts.

The walkthrough groups were separated by user group. This separation accounted for differing levels of project involvement and aimed to minimize hierarchical bias, particularly the risk of physicians influencing ICU nurses during discussions. The walkthroughs were moderated by a researcher with usability expertise, who introduced the mockup, guided the discussion

using predefined questions focused on information retrieval, overall interaction, and the structure of the XUI, and asked follow-up questions where necessary. A second researcher took notes to ensure thorough documentation.

### Sample Size and Recruiting Process

The target was to recruit 4 to 5 participants per user group, as small sample sizes are considered acceptable for formative evaluations in medical informatics [40], and sample size requirements can be adjusted based on system-specific characteristics [41].

Participants were recruited using convenience sampling. For the physician group, invitation emails were sent to 5 physicians in the same research group, which was involved in the IntelliLung project. Four physicians accepted the invitation. To increase the chances of successful recruitment of ICU nurses, 1 of the cooperating physicians identified an interested contact person among the nursing staff in 1 of the ICUs at the same hospital. The contact person informally inquired about his or her peers' interest in participating in the study. After the nurses were recruited, the contact person shared the contact details with the researchers. To be eligible to participate in the study, all participants were required to have at least 3 months of ventilation experience (demonstrated through working experience in the ICU or research activity in the field of mechanical ventilation). No compensation was paid for participation in the group usability walkthrough.

### Walkthrough Process

The details of the physician walkthrough are described in this section. Any predefined process deviations for the nurse walkthrough are reported in the section "Deviations for the Nurse Group Usability Walkthrough."

At the start, participants received in-depth information about their rights and the purpose of the usability walkthrough. Afterward, the moderator presented the CDSS use case and the main screen of the CDSS. The XUI mockup evaluation then began. The 2 XUI levels were evaluated sequentially, reflecting their hierarchical structure.

The evaluation began with the main CDSS screen, which included XUI level 1, displayed on a monitor. The moderator provided a verbal introduction. Participants were then given time to observe and reflect on XUI level 1 silently. The reflection was followed by a group discussion guided by the question: "Is the presentation of outlier warning and system certainty on the main screen sufficient?"

The XUI level 2, a subordinate screen providing additional explanation concepts, was evaluated in 2 steps. In the first step, the participants assessed the overall structure of the level. The XUI level 2 was shown, and the participants silently analyzed it. Afterward, the group briefly discussed and then collectively answered the following questions: "In which part of the interface do you find explanations for the recommendation?" and "How do you switch between the explanations?"

In a second step, the individual explanation concepts in XUI level 2 were assessed using the following steps:

- The initial version of the explanation concept was shown, and the participants silently observed and analyzed each explanation concept.
- The group discussed the explanation concept and then provided a collective answer to the questions (Table 1). Whether the group answered the questions correctly was documented.
- Alternative design variants of each explanation concept were presented. The group discussed these variants and selected their preferred variant.
- Each participant then individually assessed the selected version on Likert scales regarding the following 3 dimensions (ie, QA-1, QA-2, and QA-3) with the following anchors:
  - QA-1 (ease of understandability): "I find this explanation concept easy to understand."
  - QA-2 (suitability): "I find this explanation concept suitable for everyday clinical work."
  - QA-3 (appealingness): "I find this explanation concept appealingly visualized."
- Strongly Disagree, Disagree, Neutral, Agree, and Strongly Agree
- Additional options: no answer and free-text comments

**Table 1.** List of questions and items used to evaluate each explanation concept.

Explanation concept	Group discussion questions
Available input (AI)	<ul style="list-style-type: none"> <li>• AI-1: What information do you get from this view?</li> <li>• AI-2: In which sections were not all values available when the recommendation was generated?</li> <li>• AI-3: How would you get more detailed information on input values?</li> <li>• AI-4: How should the sections be sorted in the "available input" area?</li> <li>• Version A (static order): (1) in the presented order and (2) in a different order.</li> <li>• Version B (dynamic order): areas with missing values at the top?</li> <li>• If Version A was selected, would you prefer (1) or (2)?</li> </ul>
Feature importance (FI)	<ul style="list-style-type: none"> <li>• FI-1: What information do you get from this view?</li> <li>• FI-2: Which parameter had an importance of 80% for generating the recommendation?</li> <li>• FI-3: Where would you have to click to display the importance of more parameters?</li> </ul>
Rule-based explanation (R)	<ul style="list-style-type: none"> <li>• R-1: What information do you get from this view?</li> </ul>

The complete version of the instrument is provided in [Multimedia Appendix 2](#).

The group usability walkthrough concluded with a general discussion focusing on the following guiding questions: “Should there be more or fewer explanation concepts presented?” and “Are there any missing explanation concepts?” (The physicians had access to previous presentations of other explanation concepts before the group usability walkthrough).

### Deviations for the Nurse Group Usability Walkthrough

To accommodate the ICU nurses’ limited previous exposure to the project, the walkthrough session was adapted to include a more comprehensive introduction to CDSS and AI concepts. Insights from the physician group informed minor refinements to the mockups. Elements that had proven consistently irrelevant or confusing across previous sessions were excluded to maintain focus, while still allowing nurses to contribute unique, user-specific feedback.

Furthermore, in the final group discussion, the nurses were verbally introduced to the possibility of additional explanation concepts (since they had not previously been exposed to other types, such as counterfactual explanations). They were also asked whether the information presented in level 1 of the XUI was sufficient.

### Analysis Approach

#### Qualitative Analysis of Group Discussions

After each group usability walkthrough, the first author combined the notes from both researchers. The second researcher reviewed the merged version, and any disagreements were discussed and resolved jointly. The first author then categorized the notes by XUI level and discussion questions and organized them by content. No predefined coding scheme was used; categories emerged inductively based on the content of the group discussions. Summarizing bullet points were translated into English, as both group usability walkthroughs had been conducted in German, the participants’ native language. The second researcher reviewed the final summary. No verbatim quotes were recorded during documentation of the group usability walkthroughs.

#### Quantitative Data Analysis

The QAs (QA-1, QA-2, and QA-3) of each explanation concept were analyzed purely descriptively due to the limited number of responses per user group. For each explanation concept and user group combination of the QAs (QA-1, QA-2, and QA-3), the number of responses, the number of “No Answer,” the median, and the IQR are reported. In addition, all responses for the items are presented visually as horizontally oriented stacked bar plots.

### Ethical Considerations

This study was conducted in accordance with ethical research principles. It involved low-risk, formative group usability walkthroughs as a qualitative survey or interview method in which clinician participants were shown interface mockups and asked structured questions about usability, understandability, and suitability. No medical interventions were performed, no patient data were collected, and the study did not constitute epidemiological research. In line with the institutional guidance of the Ethics Committee of TU Dresden for such non-interventional survey or interview studies, no IRB or REB approval was sought [42]. All procedures were carried out in accordance with the principles of the Declaration of Helsinki. Informed consent was obtained from all participants after they received detailed information about the study purpose, the walkthrough procedure, data handling, and their rights. Participation was voluntary, and participants could withdraw at any time without consequences. Participants were not compensated for their participation. Privacy and confidentiality were safeguarded by collecting and storing data in pseudonymized form; only non-sensitive, profession-related participant information was collected, no identifying information is reported, and access to raw data was limited to the research team.

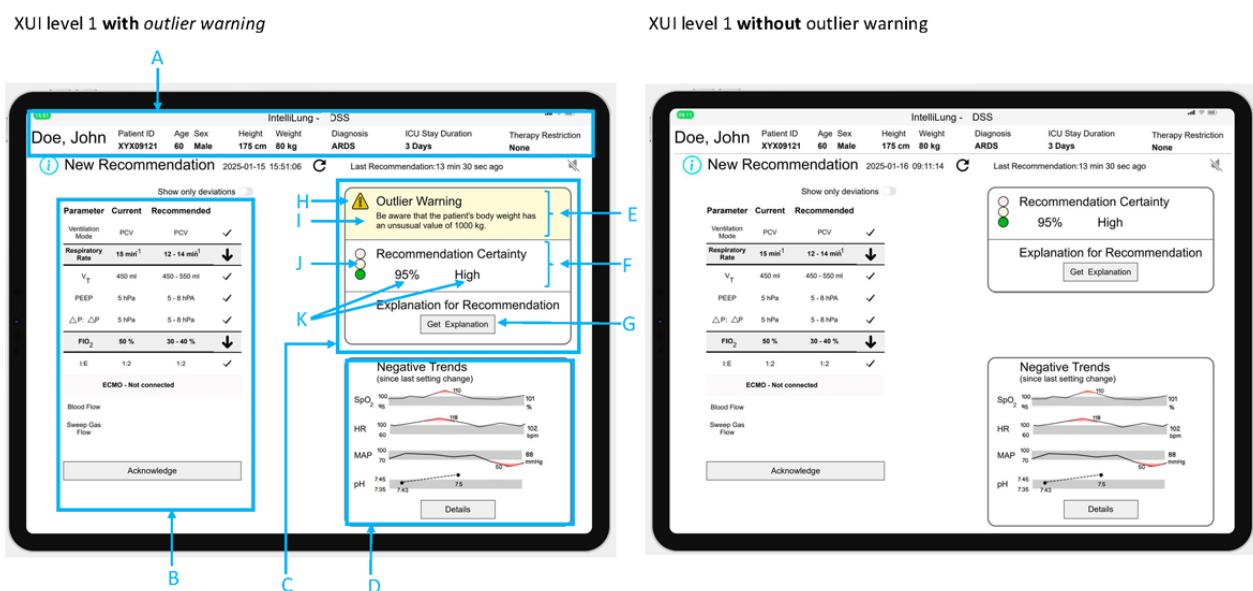
## Results

### XUI Design Phase: Structure of the XUI

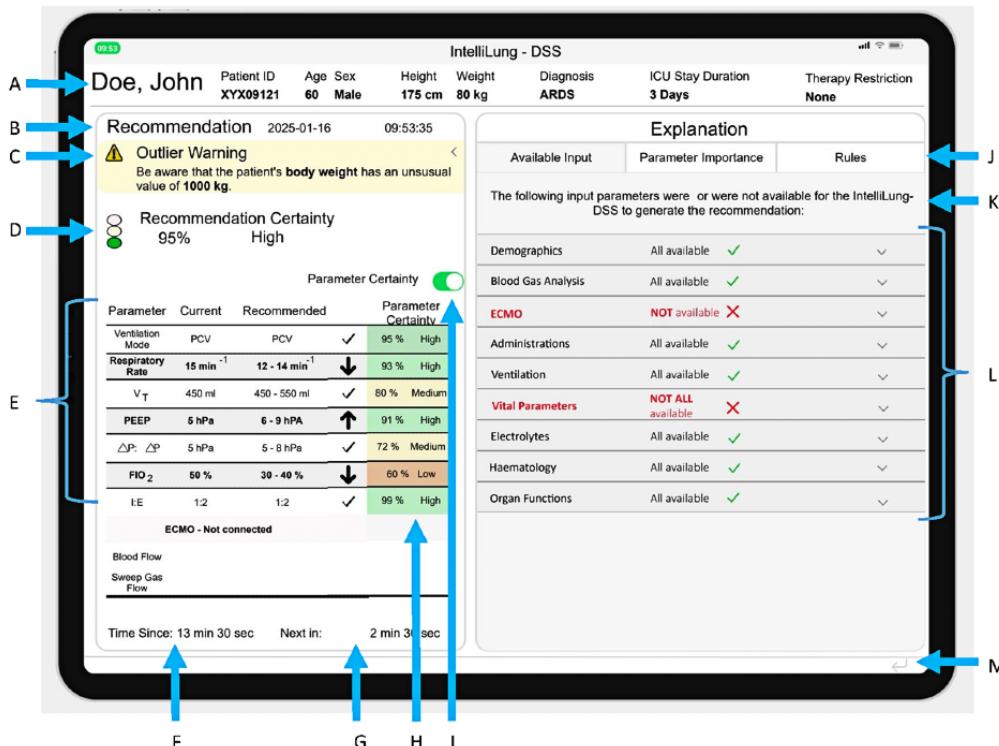
The design phase produced a 2-level XUI structure (XUI level 1 and XUI level 2). XUI level 1 is positioned on the main screen of the CDSS, adjacent to the system’s recommendations. [Figure 2](#) shows the mockup for the CDSS main screen, which contains a header with basic patient information (A in [Figure 2](#)), the recommendation (B in [Figure 2](#)), XUI level 1 (C in [Figure 2](#)), and a trend view of important parameters (D in [Figure 2](#)). XUI level 1 includes the explanation concept outlier warning (E in [Figure 2](#)), consisting of an icon with a header (H in [Figure 2](#)) and an information text (I in [Figure 2](#)); the explanation concept certainty information (F in [Figure 2](#)), for which the certainty level is presented as a traffic light (J in [Figure 2](#)), a number, and a category label (K in [Figure 2](#)); and a button (G in [Figure 2](#)) that links to XUI level 2.

[Figure 3](#) shows XUI level 2, which includes a header with basic patient information, the CDSS recommendation, a repetition of the explanation concepts from level 1, and an expanded panel on the right side for displaying explanation concepts in more detail, one explanation concept at a time. The following explanation concepts were used in level 2: available input, feature importance, and rule-based explanation. The users can switch between explanation concepts in the expanded section using tabs.

**Figure 2.** Main screen of the clinical decision support system (CDSS), showing level 1 of the explanation user interface (XUI) in the top-right section. Two versions are shown: the left screen includes the outlier warning, and the right screen does not.



**Figure 3.** Level 2 of the explanation user interface (XUI) with a description of its content. A: header with basic patient information; B: identification information for the recommendation; C and D: repetition of explainable artificial intelligence (XAI) concepts from the first level; E: repetition of the clinical decision support system (CDSS) recommendation for ventilator setting; F and G: information on when the last recommendation was generated and when the subsequent recommendation will be generated; H: certainty information for each recommended ventilator setting; I: toggle button to show or hide H; J: section for the presentation of explanation concepts, with tabs for selecting specific explanation concepts; K: short sentence introducing the explanation concept; L: space for the specific explanation concepts; M: link to the first level of the XUI. ARDS: acute respiratory distress syndrome; DSS: decision support system; ECMO: extracorporeal membrane oxygenation; FIO<sub>2</sub>: fraction of inspired oxygen; ICU: intensive care unit; PCV: pressure-controlled ventilation; PEEP: positive end-expiratory pressure; VT: tidal volume;  $\Delta P$ : driving pressure.



## Design of Explanation Concepts

### Outlier Warning

In XUI level 1, the outlier warning explanation is positioned at the top (E in **Figure 2**). If an outlier is detected in the input variables, a yellow warning triangle containing an exclamation

mark is presented next to the heading "outlier warning." The heading is followed by text naming the parameter and value that triggered the outlier warning. When an outlier is detected, the background of the outlier warning section is highlighted in yellow. Two options were created for the case that the outlier warning was not triggered: one that hides the outlier warning

section entirely, and another that displays the message “No Outlier Warning.” The outlier detection logic was intentionally left undefined at this stage, as the study focused on evaluating conceptual explanation designs within the XUI. The specific detection mechanism will be developed and validated in subsequent iterations.

### Output Certainty

Below the outlier warning section, the XUI presents the output certainty (F in Figure 2). On the left side of the output certainty

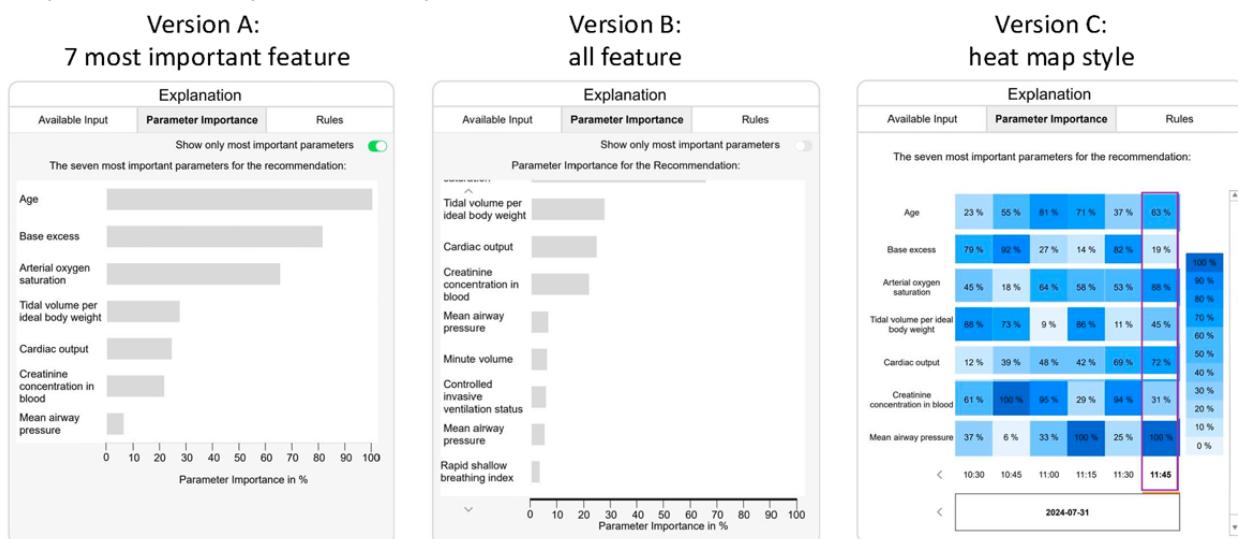
**Figure 4.** Three versions of the available input explanation concept. Version A: availability information organized in categories. Version B: availability information extended with the values of the parameters. Version C: availability information and parameter values extended with a time stamp for parameter value collection. CO2: carbon dioxide; DSS: decision support system; ECMO: extracorporeal membrane oxygenation; HCO3: bicarbonate; ICD: International Classification of Diseases; O2: oxygen; pH: potential of hydrogen.

### Explanation concept: available input

Version A: availability			Version B: availability and raw values			Version C: availability, raw values and time stamp				
Available Input	Parameter Importance	Rules	Available Input	Parameter Importance	Rules	Available Input	Parameter Importance	Rules		
The following input parameters were or were not available for the IntelliLung-DSS to generate the recommendation:										
Demographics	All available	✓	Demographics	Parameter	Availability	Value	Demographics	All available	✓	
Blood Gas Analysis										
Arterial pH-Value	✓	Concentration of HCO3	✓	Central-venous oxygen saturation	✓	69 years	Age	✓	Up to date	✓
Partial pressure of CO2	✓	Concentration of standardised HCO3	✓	Mixed-venous oxygen saturation	✓	Caucasian	Ethnicity	✓	✓	✓
Partial pressure of O2	✓	Arterial oxygen saturation	✓	Lactate concentration in blood	✓	Male	Sex	✓	✓	✓
Base excess	✓	Venous oxygen saturation	✓			1000 kg	Bodyweight	✓	✓	✓
ECMO										
NOT available										
Rotational speed of ECMO-pump: X										
Administrations	All available	✓	Administrations	Parameter	Availability	Value	Administrations	All available	✓	✓
Ventilation	All available	✓	Ventilation				Ventilation	All available	✓	✓
Vital Parameters										
NOT ALL available										
Electrolytes										
Haematology										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										
All available										

**Figure 5.** Three versions of the “feature importance” explanation concept. Version A: vertical bar plot of the 7 most important features. Version B: vertical bar plot showing the feature importance of all features. Version C: feature importance displayed as a heat map–style history in table format.

### Explanation concept: feature importance



Version A (Figure 5) displays the top 7 features in a vertical bar plot, ranked by importance.

Version B (Figure 5) uses the same bar plot format but includes all available features. The user can scroll through the bar plot vertically and toggle between Versions A and B.

Version C (Figure 5) provides the 7 most important features of the current recommendation in a heat map history, which is designed as a table. The rightmost column contains the feature

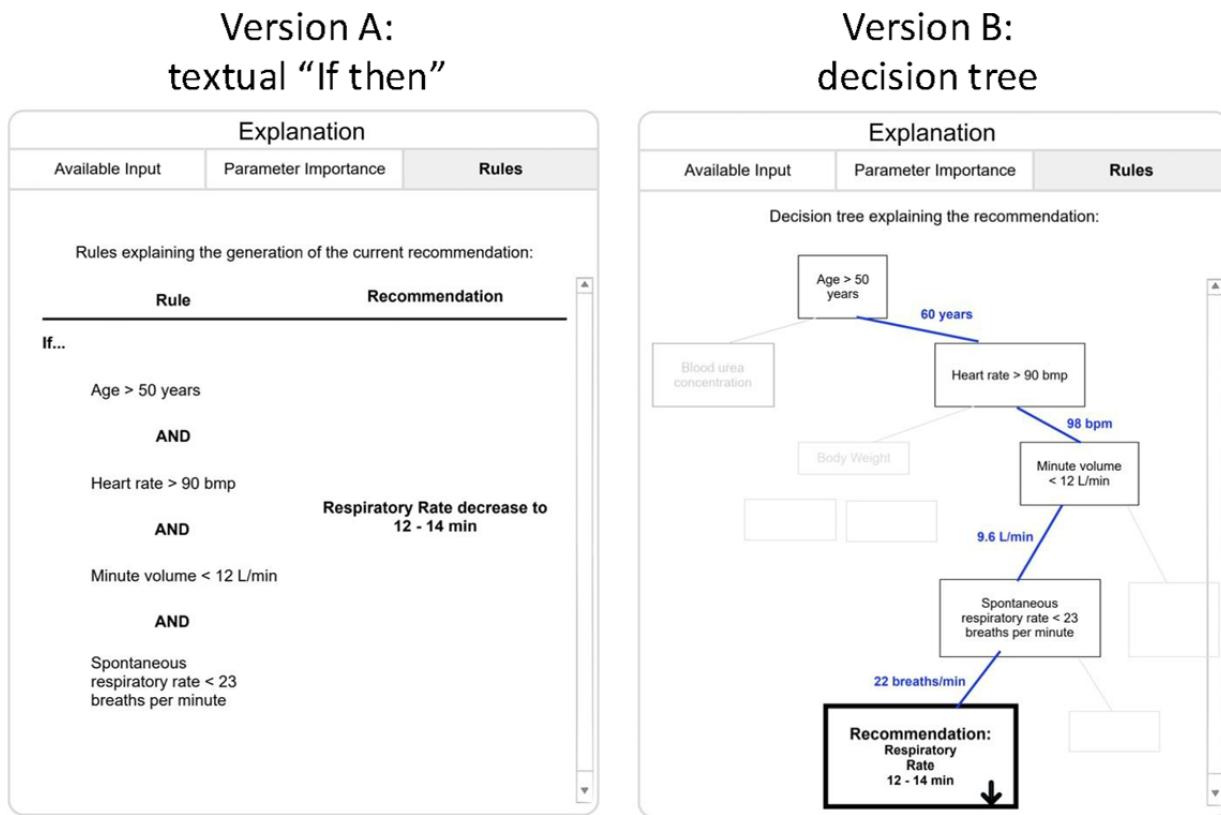
importance values for the current recommendation. The cell background color corresponds to the value of the feature importance presented in the legend. The preceding columns represent the feature importance for previous recommendations. Clicking on a column updates the recommendation displayed in the left-side panel of the XUI (B in Figure 3).

### Rule-Based Explanation

Two versions of the rule-based explanation concept were designed (Figure 6).

**Figure 6.** Different versions of the rule-based explanation concept. Version A: textual representation as an “if-then” phrase. Version B: graphical representation as a decision tree.

## Explanation concept: rule-based



Version A (Figure 6) provides the explanation concept as an “if-then” rule in text format. Version B (Figure 6) uses a visual decision tree to convey the same logic. The decision tree highlights the path to the recommendation with blue lines.

### XUI Evaluation Phase

#### Participants Description

The group usability walkthroughs were conducted with 8 health care professionals from a German university hospital, including 4 ICU physicians (4 male) and 4 ICU nurses (3 male and 1 female). All physicians participated in the preceding UCD process phases of the IntelliLung CDSS interface. In contrast, the recruited ICU nurses had no previous exposure to the IntelliLung CDSS. The recruited physicians were all anesthesiology residents with clinical and research experience in mechanical ventilation. The ICU nurses had at least 9 years

of experience working in the ICU. All nurses answered a question describing their experience selecting ventilator settings for patients as “a lot” and a question regarding their experience with AI-based CDSS as “non” (n=1) and “little” (n=3) on 4-point scales with the following answer options: “non,” “little,” “moderate,” “a lot.” The physicians were not formally asked about their AI knowledge, but some degree of knowledge can be assumed due to their participation as clinical experts in the development of the IntelliLung AI model.

#### Evaluation: XUI Level 1

Both user groups appreciated the content in XUI level 1 and favored a concise presentation format. However, the groups differed in their preferences regarding how the information should be presented. Table 2 provides a detailed summary of findings.

**Table 2.** Insights explanation user interface (XUI<sup>a</sup>) level 1 evaluation.

XUI part	Physician walkthrough	Nurse walkthrough
Overall assessment (level 1 XUI)	<ul style="list-style-type: none"> <li>Preferred a simple presentation with minimal content.</li> </ul>	<ul style="list-style-type: none"> <li>Found the information on level 1 important.</li> </ul>
Outlier warning	<ul style="list-style-type: none"> <li>Terminology improvement: suggested renaming the concept to “Input Warning” to avoid confusion with clinical outliers (ie, abnormal values).</li> <li>Navigation: clicking the warning should lead users to a section containing similar information to the available input explanation concept (Version C).</li> <li>In the absence of a warning, the section should contain the text “No Input Warning”; the section should not be empty.</li> <li>Presentation of the exact values of the parameter triggering the warning is not important.</li> </ul>	<ul style="list-style-type: none"> <li>Appreciated the explanation concept.</li> <li>Navigation: clicking on the warning should lead directly to a section where raw values are displayed.</li> <li>Requested a mechanism to acknowledge the warning and a clear distinction between new and acknowledged warnings.</li> <li>Presentation of the exact values of the parameter triggering the warning is important.</li> <li>Interaction: requested the ability to correct erroneous input values manually.</li> <li>In the absence of a warning, the section should be empty.</li> </ul>
Available input	<ul style="list-style-type: none"> <li>Structure: The explanation concept should be presented in an individual section of the interface and not at XUI level 2.</li> </ul>	—

<sup>a</sup>XUI: explanation user interface.

### Evaluation: XUI Level 2

#### Qualitative Analysis of Group Discussions

Overall, for the second level of the XUI, both the physician group and the nurse group had no problem identifying where the explanation concepts were presented and how to switch between them. However, 1 ICU nurse thought clicking on the recommendation would show an explanation text containing a detailed description of that specific recommended parameter.

The explanation concepts organize the following group discussion results.

#### Available Input

The nurses and physicians answered the questions AI-1, AI-2, and AI-3 correctly. Both groups correctly understood that the

explanation concept informs which input parameters were available to the CDSS when it generated its recommendation (AI-1), identified in which sections input parameters were missing (AI-2), and could describe how to access more detailed information about the input values (AI-3). The physicians generally preferred more detail, while nurses favored a more concise presentation, which was expressed in the group discussions, the free-text answers, and the selection of variants.

For question AI-4, which addressed parameter ordering, the physician group expressed a unanimous preference for the static ordering used in the mockup. To avoid redundancy and reduce cognitive load, this question was omitted from the nurse session—while still ensuring space for user-specific feedback. Detailed findings are provided in Table 3.

**Table 3.** Summary of usability walkthrough insights for the “available input” explanation concept.

Question	Physician walkthrough	Nurse walkthrough
AI <sup>a</sup> -4	<ul style="list-style-type: none"> <li>Preferred the static ordering of categories and parameters as presented in the mockup.</li> </ul>	<ul style="list-style-type: none"> <li>Not discussed</li> </ul>
Selected version	<ul style="list-style-type: none"> <li>Version C (most information).</li> </ul>	<ul style="list-style-type: none"> <li>Version A (most concise)</li> </ul>
Discussed insights	<ul style="list-style-type: none"> <li>Physicians should be deciding whether the values are outdated or not, and not the CDSS<sup>b</sup></li> <li>The value of the parameter triggering an “outlier warning” should be highlighted.</li> </ul>	<ul style="list-style-type: none"> <li>The other versions were seen as containing too much information.</li> <li>Preferred less information on the screens.</li> </ul>
Free-text answers	<ul style="list-style-type: none"> <li>None.</li> </ul>	<ul style="list-style-type: none"> <li>“Minimalize not important, not relevant data, or information.” [P5]</li> <li>“Deviations should be made visible.” [P5]</li> <li>“More is less.” [P5]</li> <li>“Too much information at once —&gt; just the current errors —&gt; one should be able to think that only the most recent blood gas analysis results would be considered [P6]</li> <li>“Red for highlighting is okay.” [P6]</li> <li>“Too much information.” [P7]</li> <li>“Not too much information! Keep it simple.” [P8]</li> </ul>

<sup>a</sup>AI-4: item 4 for group discussion of the available input explanation, as provided in [Table 1](#).

<sup>b</sup>CDSS: clinical decision support system.

### Feature Importance

The nurse and physician groups correctly answered questions—FI-1, FI-2, and FI-3—corresponding to the feature importance explanation concept. They correctly noted that the explanation concept presented the most influential factors for the recommendation (FI-1), identified which parameter had an

importance of 80% (FI-2), and recognized where to click to display more parameters in the explanation (FI-3). Both groups selected Version A as their favorite. The nurses mentioned that the nursing staff would likely not use this explanation concept in clinical practice. Detailed insights for this explanation concept are provided in [Table 4](#).

**Table 4.** Insights from both group usability walkthroughs for the explanation concept: feature importance.

Question	Physician walkthrough	Nurse walkthrough
Selected version	<ul style="list-style-type: none"> <li>Version A.</li> </ul>	<ul style="list-style-type: none"> <li>Version A.</li> </ul>
Discussed insights	<ul style="list-style-type: none"> <li>Version C was viewed as beneficial for research, but not appropriate for daily clinical workflows.</li> <li>X-axis labels: What format would be most intuitive for clinicians?</li> <li>X-Axis labels: the group suggested empty labels; exact values were deemed less relevant than the relative differences between features.</li> <li>Participants questioned how the explanation would be interpreted or acted upon in clinical practice.</li> </ul>	<ul style="list-style-type: none"> <li>Version C would not be interesting for the nursing staff.</li> <li>Suspected feature importance would not be used in daily clinical practice.</li> </ul>
Free-text answers	<ul style="list-style-type: none"> <li>“Importance is not clear. Which consequences should be drawn?” [P3]</li> </ul>	<ul style="list-style-type: none"> <li>“Only if someone would like to see complete information, not to be shown as default information, only shown on demand.” [P5]</li> <li>“Not relevant for the daily clinical working process.” [P6]</li> </ul>

### Rule-Based

The physicians unanimously voiced strong concerns about this explanation concept. They feared the concept would lead to an oversimplified perception of the AI model and might mislead

nonexperts. They recommended excluding this explanation from the final XUI design. Due to the physicians’ strong objections, the explanation concept was not shown to the nurses. Notes from the discussion are reported in [Table 5](#).

**Table 5.** Insights from both group usability walkthroughs for the explanation concept: rule-based.

Question	Physician walkthrough
R-1	<ul style="list-style-type: none"> <li>• Correct answer.</li> </ul>
Selected version	<ul style="list-style-type: none"> <li>• Version A.</li> </ul>
Discussed insights	<ul style="list-style-type: none"> <li>• Group consensus: the explanation concept should not be part of the XUI<sup>a</sup> for this specific CDSS<sup>b</sup> for the following reasons: (1) would suggest to the user a too simple model of the system; (2) users with a low level of expertise could learn rules from the system that are not correct or clinically validated, without knowing this; and (3) an explanation that the rules, are a approximation of the system, would occupy additional resources.</li> </ul>
Free-text answers	<ul style="list-style-type: none"> <li>• “Leave [this concept] out.” [P1]</li> <li>• “False security.” [P2]</li> <li>• “Relationships from the menu cannot be established in this way. Possibly conveys false security, false or unproven connections.” [P3]</li> <li>• “Creates a ‘false’ sense of security or being able to understand the model, but only represents it incompletely.” [P4]</li> </ul>

<sup>a</sup>XUI: explanation user interface.

<sup>b</sup>CDSS: clinical decision support system.

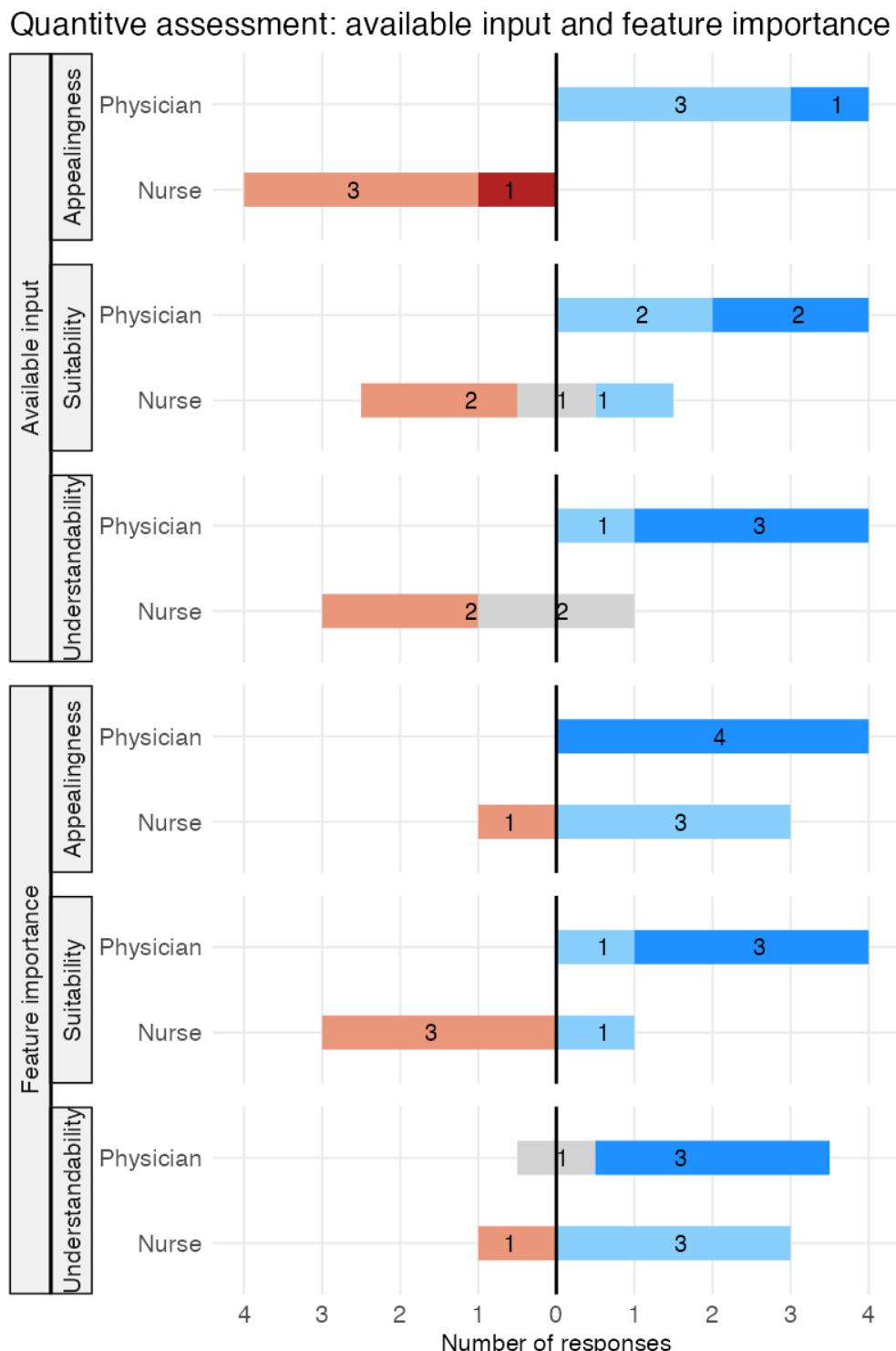
### Final Group Discussions of the XUI

In the final group discussion, both groups agreed that no explanation concepts should be added to the XUI. The nurse group reiterated their preference for using only the first level of the XUI, suggesting that users with greater interest could

optionally access Level 2. The nurses emphasized that the interface should minimize screen space usage.

### Quantitative Data Analysis

Figure 7 provides the QA results (QA-1 to QA-3) for the available input and feature importance concepts.

**Figure 7.** Results of quantitative assessments (QA-1 to QA-3) for the explanation concepts: available input and feature importance.

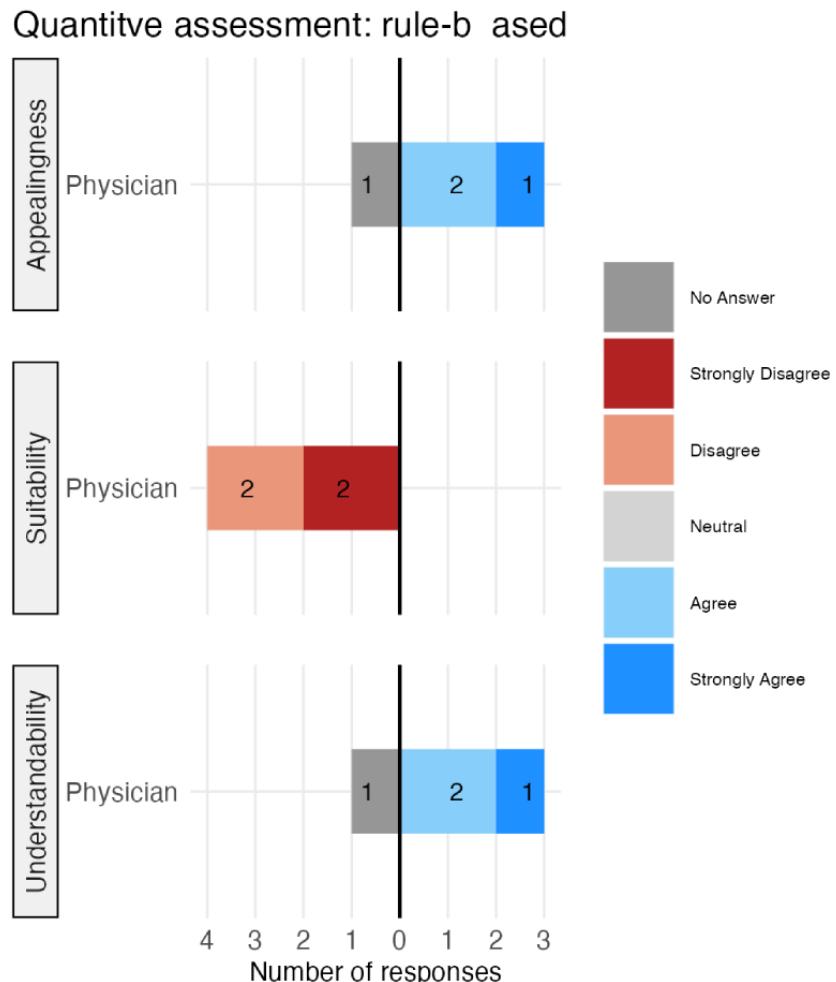
As shown in the top half of [Figure 7](#), the physicians agreed or strongly agreed that the available input concept was understandable, suitable for clinical work, and visually appealing. The nurses mostly disagreed or responded neutrally regarding the concept's suitability and understandability, and found the visualization aesthetically unappealing. The nurses' and physicians' ratings, therefore, differ across all assessed dimensions, with nurses being more critical overall.

The QA results for the feature importance explanation concept are shown in the bottom half of [Figure 7](#). Most physicians

strongly agreed that the feature importance concept was appealing, suitable, and easy to understand. The nurses mostly agreed with its understandability and appeal, but tended to disagree with its suitability for clinical work. Both user groups, therefore, rate the understandability and visual appeal of the explanation concepts positively. However, the user groups differ in their assessment of the suitability for their clinical work, with a positive assessment by the physicians and a negative assessment by the nurses.

Figure 8 provides the QA results for the rule-based explanation concept. Three physicians strongly agreed or agreed with its appeal and understandability; one chose the “no answer” option. All 4 physicians disagreed or strongly disagreed with its suitability for clinical practice.

**Figure 8.** Quantitative assessment results of the rule-based explanation concept.



## Discussion

### Principal Findings

This paper provides findings from the first design and formative evaluation phase of an XUI for an AI-based CDSS aimed at optimizing mechanical ventilation in the ICU.

The design led to an XUI with 2 levels. Level 1 of the XUI, positioned next to the CDSS recommendations on the main screen, includes the output certainty and outlier warning concepts. Level 2 of the XUI provides on-demand access to additional explanation concepts: available input, feature importance, and rule-based explanation. Multiple versions of the explanation concepts on the second level have been designed and evaluated.

The initial evaluation consisted of 2 group usability walkthroughs with ICU physicians and nurses. Both user groups liked the first level of the XUI. The physicians were satisfied with XUI level 2 but provided valuable feedback for improving

The calculated descriptive statistics for each QA for all 3 explanation concepts on the second level are provided in [Multimedia Appendix 3](#). The descriptive statistics mirror the results visible in [Figures 7](#) and [8](#).

it. The nurses were skeptical about whether they would use level 2 of the XUI in clinical practice.

### XUI Design Phase

For the design phase, the following research question was investigated: RQ1: “How should the XUI of an AI-based CDSS for mechanical ventilation optimization in the ICU be designed?”

To address this research question, the first focus was on the XUI structure. The 2-level structure of the XUI allowed the provision of easy-to-grasp explanation concepts next to the CDSS recommendation and more complex and more detailed explanation concepts on the second level, on demand. This structure was based on the input of the previous UCD design phases and current design recommendations [35]. This design choice aligns with new recommendations published as a preprint after the work reported in this paper was concluded. The preprint recommends multiple levels with increasing detail for XUIs for the ICU [29].

After the structure of the XUI was created, the investigation addressed the presentation of the explanation concepts themselves. The individual visualizations for the explanation concepts are comparable to typical visualizations found for XUIs in the medical domain [43,44]. The explanation concepts were kept in line with the design of XUIs to provide clinicians with an impression of the possibility of explanations at the current state of the art of research in the field.

The different versions of the explanation concepts in level 2 of the XUI varied in the level of detail presented to determine users' preferred level of detail for this CDSS. Based on the literature, this is considered an important design decision [29,35]. Minor adaptations were made to incorporate design recommendations for the UCD of XUIs; for example, the explanation concept of feature importance was presented to the user under the term "parameter importance," as the term feature importance was assumed not to be self-explanatory for non-AI experts. In addition, to add some variety to the different presented explanation concepts, the typical bar plot visualization for the feature importance explanation concept was extended by a version inspired by the works of Bienefeld et al [45], which created a heat map-style feature importance history in an XUI.

### XUI Evaluation Phase

During the evaluation phase, the following research question, RQ2, was investigated: "What are clinicians' perceptions of the proposed XUI of the AI-based CDSS for mechanical ventilation optimization?"

The users' first impression of the 2-level structure of the XUI suggests that this structure is the right approach for this CDSS. Both groups were content with level 1 of the XUI, and the physician group was satisfied with level 2 of the XUI, while the nurses were more skeptical of their intentions to use level 2. No concerns were raised by either group about the general structure. The structure allows both user groups to access the information they prefer while preventing information clutter on the main screen of the CDSS.

The combined results from group discussions, concept selection, QAs, and free-text responses indicate that ICU nurses and physicians had noticeably different preferences regarding the level of detail and complexity in the XUI. The nurses preferred less complex explanation concepts, such as output certainty and outlier warning, explicitly requested a lower level of detail for the explained concepts, and rated the suitability of the explanation concepts on level 2 lower than the physicians. The literature supports this, as it has been reported that nurses wished for clear and concise explanations, outputs, and actionable information when dealing with AI [31,46]. In contrast, the physicians were satisfied with both XUI levels, which allowed them to consult low-detail and less complex information on level 1 and more detailed and complex explanation concepts on level 2 of the XUI. This implies that the explanation needs and preferences between the nurses and physicians differ, with physicians having a higher need for more complex and detailed explanations than the ICU nurses. The work from Barda et al [31] found a similar insight. Current research supports this by showing that different health care provider groups can have different explanation preferences [15] or that different

stakeholders have different explanation needs [47]. The observed differences may reflect role-specific workflows in ventilator care. ICU nurses typically operate in fast-paced bedside routines with frequent monitoring and rapid adjustments, which may favor concise, immediately actionable explanations. Physicians, in contrast, often integrate XAI outputs into broader diagnostic reasoning and decision validation, which may increase the perceived value of more detailed, mechanism-oriented explanations provided on level 2.

Notably, divergence between groups was most consistent for clinical suitability rather than for understandability or visual appeal (Figure 7). This suggests that acceptance may not only be a matter of whether clinicians can comprehend an explanation and perceive it as visually appealing, but also whether it fits their real-world task context and decision-making responsibilities. Therefore, suitability should be treated as a primary design target in XUI development, alongside clarity and aesthetics.

These results emphasize the need for a rigorous application of the UCD process when an XUI is developed for different health care providers. In the XUI developed and evaluated in this study, the 2-level solution appears to meet the differing explanation needs and preferences of the user groups. Future iterations should evaluate whether this approach remains sufficient or if separate XUIs are needed. It should be kept in mind that ICU nurses will probably not consult level 2 of the XUI in future iterations. Therefore, presenting potentially necessary and valuable information only on level 2 should be avoided. Clinical role alone may not fully explain the differences; AI literacy and ventilator-care experience likely act as moderators that shape how much explanation depth is perceived as useful versus burdensome. Future work should disentangle professional role effects from training- and experience-related factors, for example, by stratifying participants accordingly.

Interestingly, both groups were skeptical about their peers' adaptations of the explanation concepts in the clinical daily routine, which was voiced during the group discussions and in the free-text answers. The nurse group voiced more skepticism than the physician group, and their skepticism increased with the level of detail and complexity of the explanation concepts. The physicians were mainly satisfied with the proposed explanation concepts, apart from the rule-based explanation concept. However, they stressed that the complexity of the explanation concepts should not exceed that used in the XUI. These results explain why both groups had no interest in more complex explanation concepts, such as a counterfactual explanation. This result contrasts with the results of a study by Jung et al [48], in which clinicians had a positive perception of counterfactual explanations.

The physician group positively assessed the explanation concepts on level 2 of the XUI, apart from the suitability of the rule-based explanations. The understandability ratings for the explanation concepts are similar to previous research, showing that physicians can understand the selected concepts [32]. The positive assessment of the feature importance explanation concept for the physician group is in line with other research efforts [14,15]. Due to the overall positive assessment by the

physicians of both levels of the XUI and nurses' positive perception of level 1 of the XUI, it seems that the results of the first design phase were a step in the right direction. However, the evaluation also revealed some improvement opportunities.

Second, the evaluation phase investigated research question RQ-3: "What XUI design improvements should be addressed in subsequent design iterations?"

Based on the results of the evaluation, all the explanation concepts should be revised, keeping the nurses' preferences for a low level of detail and complexity for the explanation concept, and their perception of the explanation concept as having too much detail in mind. In the following design iteration, all the explanation concept versions should be provided with a lower level of detail or complexity as a default version. From these default versions, the more interested user should be able to access a more detailed version on demand. This approach has to be balanced so that the more interested users are not burdened while accessing the explanation concepts with the current level of detail and complexity for which they provided positive feedback in the evaluation.

For the feature importance explanation concept, the evaluation revealed that further research is needed for the design of the x-axis. The x-axis should be redesigned to be understandable for physicians and nurses and to transport the appropriate amount of information without overwhelming them. As raised in the physician group walkthrough, one option could be an empty x-axis; the suitability of such a version should be tested in comparison to other typical variants of the x-axis for this explanation concept. As the question about the actionability of the information provided in the feature importance explanation concept arose in the physician group, it should be investigated whether incorporating information about how to act on the explanation concept would be valuable. For the outlier warning explanation concept, the evaluation revealed different preferences for the case in which no outlier would be detected. It should be investigated how the user groups would react to the version they did not prefer when interacting with a version of the XUI.

An increased effort should be put into the design of level 1 of the XUI. No alternative versions had been designed for the explanation concepts on level 1 for this initial prototype to reduce complexity. These should be a primary focus for the next design iteration, as both user groups saw them as valuable. While refining these explanation concepts, an emphasis should be placed on concise information visualizations and effortless information retention from both user groups.

## Limitations

This first iteration of the design and evaluation phase of the UCD process for the XUI of a CDSS has some limitations that influence the generalizability of the results and should be carefully considered when interpreting the results of this study. A pragmatic group usability walkthrough was selected for the evaluation phase due to the limited number of available and suitable participants per user group. Nonetheless, this pragmatic approach proved valuable as it allowed for the receipt of early feedback from subject matter experts on mechanical ventilation

with real-world and research experience. Their experience might have biased their assessment of the XUI, as they might not have had to rely on the CDSS in the first place due to their extensive experience. However, their knowledge allowed them to focus on the XUI design and incorporate extensive practical knowledge and expertise in their feedback and assessment of the XUI. The small group size of only 4 participants per group usability walkthrough might be seen as a limitation, but this group size allowed for a lively discussion between participants, and smaller sample sizes are in line with early evaluation iterations [40,41]. In addition, the group walkthrough allowed for the collection of feedback in a condensed amount of time, which was necessary due to resource restrictions. Nonetheless, these points limit the generalizability to the clinical population that should be addressed in the following design iterations.

Furthermore, the changes for the nurse walkthrough limit the comparability between the 2 walkthrough groups. The walkthrough had to be adjusted to avoid overwhelming the participants. Despite the minor adjustments, the collected insights allow for a preliminary assessment of the difference between the user groups. No actual data or patient vignettes were used to promote deeper engagement or trigger demand for an explanation from the CDSS, which limits the comparability of the assessment of XUI in real-world situations. The current state of the XUI mockup was sufficient to spark extensive discussions while preventing the participants from focusing on the medical validity of the situation or mockup content. In this initial iteration of the UCD, only subjective assessments of the XUIs were elicited. This may not provide a complete picture, as it has been shown for XUIs in the ICU that subjective assessments might not transfer to the use or influence of the XUI during the medical decision-making process [30,49]. Nevertheless, the amount of valuable insights generated through subjective assessments of the XUI in this study shows that this type of feedback is helpful during such an early design phase. These limitations should be addressed in future evaluation iterations.

## Future Research

Based on this research, the XUI mockup should be improved. The improvements should address the identified opportunities for enhancement in the XUI, making it more suitable for both user groups without drastically altering it, as physicians were already satisfied with the current design. A higher focus should be placed on level 1 of the XUI, as this is likely the most used part. In addition, feedback should be sought from the AI and XAI algorithm developer and incorporated into the mockup to enhance the realism of the XUI for the next iteration. An improved and extended XUI mockup would be a suitable basis for a more elaborate evaluation. The enhanced mockup could be accompanied by different vignettes of patients and situations that might trigger the need for an explanation from the user. Future evaluation rounds should have a higher experimental character to avoid the potential shortcomings of subjective assessments of an XUI in the ICU and allow the users to interact with the XUI directly. The proposed evaluation iterations should aim for a higher sample size and a diversified participant pool concerning their ventilation experience, AI literacy, and clinical center to foster generalizability and reflectiveness of the

clinician population. To this end, a remote experimental study could be conducted in which the user directly interacts with the CDSS based on presented patient vignettes. During such an evaluation, participants reflecting a broad base of clinical users should be recruited from multiple institutions.

## Conclusion

To the best of our knowledge, this paper is the first to report on the UCD and formative evaluation of an XUI for an AI-based CDSS aimed at optimizing continuous mechanical ventilation in the ICU. The findings indicate that a 2-level XUI structure can effectively address the distinct explanation needs of ICU

clinicians. Nurses' information needs were largely met through concise, high-level explanations such as outlier warning and output certainty, presented alongside CDSS recommendations, whereas physicians preferred access to more detailed explanations—such as available input and feature importance—on a secondary level. These results provide empirical evidence that explanation needs differ by professional role and underscore the necessity of user-centered, iterative design processes in the development of XAI systems for critical care. Future work should refine these explanation concepts and evaluate their impact on clinical decision-making, user trust, and system adoption.

---

## Acknowledgments

The authors are grateful to the IntelliLung project team for their productive cooperation and to the participants of the evaluation phase for contributing their time and feedback. The authors declare that generative AI (GAI) tools were used in the research and writing process. According to the GAIDeT taxonomy (2025), tasks delegated to GAI tools under full human supervision included proofreading, editing, and translation. The GAI tools used in this work were ChatGPT-4o (OpenAI) [50], ChatGPT-5 (OpenAI), and Grammarly (Grammarly, Inc). Responsibility for the final manuscript remains entirely with the authors. GAI tools were not listed as authors and do not bear responsibility for the content or outcomes of this work. In addition, Claude Sonnet 3.5 (Anthropic) [37] was used to generate syntactic data points for the parameters displayed in the mockups, as described in the “Methods” section. The declaration was submitted by ICJ.

---

## Funding

The research work was carried out as part of the junior research group “CDS2USE-prospective user-centered design of clinical decision support systems in the context of personalized medicine,” funded by the German Federal Ministry of Research, Technology, and Space (BMFTR; grant number 01ZZ2002).

---

## Authors' Contributions

ICJ contributed to the conceptualization, methodology, and investigation, and was responsible for writing the original draft, reviewing and editing the manuscript, and creating the visualization. MZ and KS contributed to the Investigation and assisted with manuscript review and editing. MS and BS contributed to the manuscript review and editing, and were responsible for supervision and funding acquisition.

---

## Conflicts of Interest

None declared.

---

## Multimedia Appendix 1

STARE-HI checklist.

[[PDF File \(Adobe PDF File, 77 KB\)](#)-Multimedia Appendix 1]

---

## Multimedia Appendix 2

Quantitative assessment sheet.

[[PDF File \(Adobe PDF File, 87 KB\)](#)-Multimedia Appendix 2]

---

## Multimedia Appendix 3

Descriptive statistics for quantitative assessments of each explanation per user group.

[[DOCX File, 18 KB](#)-Multimedia Appendix 3]

---

## References

1. Barea Mendoza JA, Valiente Fernandez M, Pardo Fernandez A, Gómez Álvarez J. Current perspectives on the use of artificial intelligence in critical patient safety. *Med Intensiva (Engl Ed)*. 2025;49(3):154-164. [doi: [10.1016/j.medine.2024.04.002](https://doi.org/10.1016/j.medine.2024.04.002)] [Medline: [38677902](https://pubmed.ncbi.nlm.nih.gov/38677902/)]

2. Chen M, Zhang B, Cai Z, Seery S, Gonzalez MJ, Ali NM, et al. Acceptance of clinical artificial intelligence among physicians and medical students: a systematic review with cross-sectional survey. *Front Med (Lausanne)*. 2022;9:990604. [FREE Full text] [doi: [10.3389/fmed.2022.990604](https://doi.org/10.3389/fmed.2022.990604)] [Medline: [36117979](#)]
3. Hua D, Petrina N, Young N, Cho J, Poon SK. Understanding the factors influencing acceptability of AI in medical imaging domains among healthcare professionals: a scoping review. *Artif Intell Med*. 2024;147:102698. [FREE Full text] [doi: [10.1016/j.artmed.2023.102698](https://doi.org/10.1016/j.artmed.2023.102698)] [Medline: [38184343](#)]
4. Asan O, Bayrak AE, Choudhury A. Artificial intelligence and human trust in healthcare: focus on clinicians. *J Med Internet Res*. 2020;22(6):e15154. [FREE Full text] [doi: [10.2196/15154](https://doi.org/10.2196/15154)] [Medline: [32558657](#)]
5. van Kolfschooten H, van Oirschot J. The EU artificial intelligence act (2024): implications for healthcare. *Health Policy*. 2024;149:105152. [FREE Full text] [doi: [10.1016/j.healthpol.2024.105152](https://doi.org/10.1016/j.healthpol.2024.105152)] [Medline: [39244818](#)]
6. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). EUR-Lex. 2024. URL: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj> [accessed 2025-12-25]
7. Meskó B, Görög M. A short guide for medical professionals in the era of artificial intelligence. *NPJ Digit Med*. 2020;3:126. [FREE Full text] [doi: [10.1038/s41746-020-00333-z](https://doi.org/10.1038/s41746-020-00333-z)] [Medline: [33043150](#)]
8. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. *Nat Med*. 2022;28(1):31-38. [doi: [10.1038/s41591-021-01614-0](https://doi.org/10.1038/s41591-021-01614-0)] [Medline: [35058619](#)]
9. Amann J, Blasimme A, Vayena E, Frey D, Madai VI, Precise4Q consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med Inform Decis Mak*. 2020;20(1):310. [FREE Full text] [doi: [10.1186/s12911-020-01332-6](https://doi.org/10.1186/s12911-020-01332-6)] [Medline: [33256715](#)]
10. Mohanty A, Mishra S. A comprehensive study of explainable artificial intelligence in healthcare. In: Mishra S, Tripathy H, Mallick P, Shaalan K, editors. *Augmented Intelligence in Healthcare: A Pragmatic and Integrated Analysis*. Singapore. Springer; 2022.
11. Allgaier J, Mulansky L, Draelos RL, Pryss R. How does the model make predictions? A systematic literature review on the explainability power of machine learning in healthcare. *Artif Intell Med*. 2023;143:102616. [FREE Full text] [doi: [10.1016/j.artmed.2023.102616](https://doi.org/10.1016/j.artmed.2023.102616)] [Medline: [37673561](#)]
12. Rosenbacke R, Melhus Å, McKee M, Stuckler D. How explainable artificial intelligence can increase or decrease clinicians' trust in AI applications in health care: systematic review. *JMIR AI*. 2024;3:e53207. [FREE Full text] [doi: [10.2196/53207](https://doi.org/10.2196/53207)] [Medline: [39476365](#)]
13. Liu CF, Chen ZC, Kuo SC, Lin TC. Does AI explainability affect physicians' intention to use AI? *Int J Med Inform*. 2022;168:104884. [doi: [10.1016/j.ijmedinf.2022.104884](https://doi.org/10.1016/j.ijmedinf.2022.104884)] [Medline: [36228415](#)]
14. Sivaraman V, Bukowski LA, Levin J, Kahn JM, Perer A. Ignore, trust, or negotiate: understanding clinician acceptance of AI-based treatment recommendations in health care. 2023. Presented at: CHI '23: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems; April 23-28, 2023:1-18; Hamburg, Germany. [doi: [10.1145/3544548.3581075](https://doi.org/10.1145/3544548.3581075)]
15. Du Y, Antoniadi AM, McNestry C, McAuliffe FM, Mooney C. The role of XAI in advice-taking from a clinical decision support system: a comparative user study of feature contribution-based and example-based explanations. *Appl Sci*. Oct 13, 2022;12(20):10323. [doi: [10.3390/app122010323](https://doi.org/10.3390/app122010323)]
16. Nasarian E, Alizadehsani R, Acharya UR, Tsui KL. Designing interpretable ML system to enhance trust in healthcare: a systematic review to proposed responsible clinician-AI-collaboration framework. *Inf Fusion*. Aug 2024;108:102412. [doi: [10.1016/j.inffus.2024.102412](https://doi.org/10.1016/j.inffus.2024.102412)]
17. Lambert SI, Madi M, Sopka S, Lenes A, Stange H, Buszello C, et al. An integrative review on the acceptance of artificial intelligence among healthcare professionals in hospitals. *NPJ Digit Med*. 2023;6(1):111. [FREE Full text] [doi: [10.1038/s41746-023-00852-5](https://doi.org/10.1038/s41746-023-00852-5)] [Medline: [37301946](#)]
18. Combi C, Amico B, Bellazzi R, Holzinger A, Moore JH, Zitnik M, et al. A manifesto on explainability for artificial intelligence in medicine. *Artif Intell Med*. 2022;133:102423. [FREE Full text] [doi: [10.1016/j.artmed.2022.102423](https://doi.org/10.1016/j.artmed.2022.102423)] [Medline: [36328669](#)]
19. Rubinger L, Gazendam A, Ekhtiari S, Bhandari M. Machine learning and artificial intelligence in research and healthcare. *Injury*. 2023;54 Suppl 3:S69-S73. [doi: [10.1016/j.injury.2022.01.046](https://doi.org/10.1016/j.injury.2022.01.046)] [Medline: [35135685](#)]
20. Amann J, Vetter D, Blomberg SN, Christensen HC, Coffee M, Gerke S, et al. Z-Inspection initiative. To explain or not to explain? - Artificial intelligence explainability in clinical decision support systems. *PLOS Digit Health*. 2022;1(2):e0000016. [FREE Full text] [doi: [10.1371/journal.pdig.0000016](https://doi.org/10.1371/journal.pdig.0000016)] [Medline: [36812545](#)]
21. Artificial intelligence to help optimize ventilation for intensive care patients. IntelliLung. URL: <https://intellilung-project.eu/> [accessed 2024-02-14]
22. Mohseni S, Zarei N, Ragan ED. A multidisciplinary survey and framework for design and evaluation of explainable AI systems. *ACM Trans Interact Intell Syst*. Sep 03, 2021;11(3-4):1-45. [doi: [10.1145/3387166](https://doi.org/10.1145/3387166)]
23. Saarela M, Podgorelec V. Recent applications of explainable AI (XAI): a systematic literature review. *Appl Sci*. 2024;14(19):8884. [doi: [10.3390/app14198884](https://doi.org/10.3390/app14198884)]

24. Wang Z, Zhang L, Huang T, Yang R, Cheng H, Wang H, et al. Developing an explainable machine learning model to predict the mechanical ventilation duration of patients with ARDS in intensive care units. *Heart Lung*. 2023;58:74-81. [FREE Full text] [doi: [10.1016/j.hrtlng.2022.11.005](https://doi.org/10.1016/j.hrtlng.2022.11.005)] [Medline: [36423504](https://pubmed.ncbi.nlm.nih.gov/36423504/)]

25. Lin S, Yang M, Wang Y, Wang Z. An Explainable AI predictor to improve clinical prognosis for acute respiratory distress syndrome. 2023. Presented at: Computing in Cardiology (CinC); October 1-4, 2023; Atlanta, GA. [doi: [10.22489/cinc.2023.003](https://doi.org/10.22489/cinc.2023.003)]

26. Aslam N. Explainable artificial intelligence approach for the early prediction of ventilator support and mortality in COVID-19 patients. *Computation*. 2022;10(3):36. [doi: [10.3390/computation10030036](https://doi.org/10.3390/computation10030036)]

27. Safaei F, Nenadović M, Liessner R, Raphael T, Jakob W, Jens L, et al. X-Vent: ICU ventilation with explainable model-based reinforcement learning. 2024. Presented at: 27th European Conference on Artificial Intelligence (ECAI 2024); October 19-24, 2024:4719-4726; Santiago de Compostela, Spain. [doi: [10.3233/faia241069](https://doi.org/10.3233/faia241069)]

28. Hughes N, Jia Y, Sujan M, Lawton T, Habli I, McDermid J. Contextual design requirements for decision-support tools involved in weaning patients from mechanical ventilation in intensive care units. *Appl Ergon*. 2024;118:104275. [FREE Full text] [doi: [10.1016/j.apergo.2024.104275](https://doi.org/10.1016/j.apergo.2024.104275)] [Medline: [38574594](https://pubmed.ncbi.nlm.nih.gov/38574594/)]

29. Clark J, Wragg M, Nielsen E. Exploring the requirements of clinicians for explainable AI decision support systems in intensive care. *arXiv*. Nov 18, 2024. [FREE Full text]

30. Nagendran M, Festor P, Komorowski M, Gordon AC, Faisal AA. Eye tracking insights into physician behaviour with safe and unsafe explainable AI recommendations. *NPJ Digit Med*. 2024;7(1):202. [FREE Full text] [doi: [10.1038/s41746-024-01200-x](https://doi.org/10.1038/s41746-024-01200-x)] [Medline: [39095449](https://pubmed.ncbi.nlm.nih.gov/39095449/)]

31. Barda AJ, Horvat CM, Hochheiser H. A qualitative research framework for the design of user-centered displays of explanations for machine learning model predictions in healthcare. *BMC Med Inform Decis Mak*. 2020;20(1):257. [FREE Full text] [doi: [10.1186/s12911-020-01276-x](https://doi.org/10.1186/s12911-020-01276-x)] [Medline: [33032582](https://pubmed.ncbi.nlm.nih.gov/33032582/)]

32. Schoonderwoerd TAJ, Jorritsma W, Neerincx MA, van den Bosch K. Human-centered XAI: developing design patterns for explanations of clinical decision support systems. *Int J Hum Comput Stud*. 2021;154:102684. [doi: [10.1016/j.ijhcs.2021.102684](https://doi.org/10.1016/j.ijhcs.2021.102684)]

33. Ergonomics of human-system interaction—part 210: human-centred design for interactive systems. International Organization for Standardization (ISO). 2019. URL: <https://www.iso.org/standard/77520.html> [accessed 2025-02-05]

34. Jung IC, Zerlik M, Schuler K, Sedlmayr M, Sedlmayr B. User-centered development of explanation user interfaces for AI-based CDSS: lessons learned from early phases. *Stud Health Technol Inform*. 2024;316:570-574. [doi: [10.3233/SHTI240478](https://doi.org/10.3233/SHTI240478)] [Medline: [39176806](https://pubmed.ncbi.nlm.nih.gov/39176806/)]

35. Jung IC, Schuler K, Zerlik M, Grummt S, Sedlmayr M, Sedlmayr B. Overview of basic design recommendations for user-centered explanation interfaces for AI-based clinical decision support systems: a scoping review. *Digit Health*. 2025;11:20552076241308298. [FREE Full text] [doi: [10.1177/20552076241308298](https://doi.org/10.1177/20552076241308298)] [Medline: [39866885](https://pubmed.ncbi.nlm.nih.gov/39866885/)]

36. Justinmind version 9.9.0. Justinmind. URL: <https://www.justinmind.com/release-notes> [accessed 2025-12-25]

37. Claude 3.5 Sonnet. Anthropic. URL: <https://www.anthropic.com/news/clause-3-5-sonnet> [accessed 2025-12-25]

38. Talmon J, Ammenwerth E, Brender J, de Keizer N, Nykänen P, Rigby M. STARE-HI—statement on reporting of evaluation studies in health informatics. *Int J Med Inform*. 2009;78(1):1-9. [FREE Full text]

39. Brender J, Talmon J, de Keizer N, Nykänen P, Rigby M, Ammenwerth E. STARE-HI - statement on reporting of evaluation studies in health informatics: explanation and elaboration. *Appl Clin Inform*. 2013;4(3):331-358. [FREE Full text] [doi: [10.4338/ACI-2013-04-RA-0024](https://doi.org/10.4338/ACI-2013-04-RA-0024)] [Medline: [24155788](https://pubmed.ncbi.nlm.nih.gov/24155788/)]

40. Human factors engineering—design of medical devices. ANSI Webstore. 2009. URL: <https://webstore.ansi.org/standards/aami/ansiaamihe752009> [accessed 2025-12-25]

41. Applying human factors and usability engineering to medical devices: guidance for industry and food and drug administration staff. Food and Drug Administration (FDA). 2016. URL: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/applying-human-factors-and-usability-engineering-medical-devices> [accessed 2025-12-25]

42. Ethics Committee of TU Dresden. FAQ: When is an ethics application required? TU Dresden website. URL: [https://tu-dresden.de/tu-dresden/organisation/gremien-und-beauftragte/kommissionen/ethikkommission/faq#ck\\_Wann%20ist%20ein%20Antrag%20erforderlich](https://tu-dresden.de/tu-dresden/organisation/gremien-und-beauftragte/kommissionen/ethikkommission/faq#ck_Wann%20ist%20ein%20Antrag%20erforderlich)

43. Cálem J, Moreira C, Jorge J. Intelligent systems in healthcare: a systematic survey of explainable user interfaces. *Comput Biol Med*. 2024;180:108908. [FREE Full text] [doi: [10.1016/j.combiomed.2024.108908](https://doi.org/10.1016/j.combiomed.2024.108908)] [Medline: [39067152](https://pubmed.ncbi.nlm.nih.gov/39067152/)]

44. Subramanian HV, Canfield C, Shank DB. Designing explainable AI to improve human-AI team performance: a medical stakeholder-driven scoping review. *Artif Intell Med*. 2024;149:102780. [FREE Full text] [doi: [10.1016/j.artmed.2024.102780](https://doi.org/10.1016/j.artmed.2024.102780)] [Medline: [38462282](https://pubmed.ncbi.nlm.nih.gov/38462282/)]

45. Bienefeld N, Boss JM, Lüthy R, Brodbeck D, Azzati J, Blaser M, et al. Solving the explainable AI conundrum by bridging clinicians' needs and developers' goals. *NPJ Digit Med*. 2023;6(1):94. [FREE Full text] [doi: [10.1038/s41746-023-00837-4](https://doi.org/10.1038/s41746-023-00837-4)] [Medline: [37217779](https://pubmed.ncbi.nlm.nih.gov/37217779/)]

46. Hassan EA, El-Ashry AM. Leading with AI in critical care nursing: challenges, opportunities, and the human factor. *BMC Nurs*. 2024;23(1):752. [FREE Full text] [doi: [10.1186/s12912-024-02363-4](https://doi.org/10.1186/s12912-024-02363-4)] [Medline: [39402609](https://pubmed.ncbi.nlm.nih.gov/39402609/)]

47. Kim M, Kim S, Kim J, Song T, Kim Y. Do stakeholder needs differ? - Designing stakeholder-tailored explainable artificial intelligence (XAI) interfaces. *Int J Hum Comput Stud*. 2024;181:103160. [doi: [10.1016/j.ijhcs.2023.103160](https://doi.org/10.1016/j.ijhcs.2023.103160)]
48. Jung J, Kang S, Choi J, El-Kareh R, Lee H, Kim H. Evaluating the impact of explainable AI on clinicians' decision-making: a study on ICU length of stay prediction. *Int J Med Inform*. 2025;201:105943. [FREE Full text] [doi: [10.1016/j.ijmedinf.2025.105943](https://doi.org/10.1016/j.ijmedinf.2025.105943)] [Medline: [40318498](https://pubmed.ncbi.nlm.nih.gov/40318498/)]
49. Narendran M, Festor P, Komorowski M, Gordon AC, Faisal AA. Quantifying the impact of AI recommendations with explanations on prescription decision making. *NPJ Digit Med*. 2023;6(1):206. [FREE Full text] [doi: [10.1038/s41746-023-00955-z](https://doi.org/10.1038/s41746-023-00955-z)] [Medline: [37935953](https://pubmed.ncbi.nlm.nih.gov/37935953/)]
50. GPT-4o. OpenAI. 2024. URL: <https://openai.com/index/hello-gpt-4o> [accessed 2025-12-25]

## Abbreviations

**AI:** artificial intelligence

**CDSS:** clinical decision support system

**DIN:** Deutsches Institut für Normung

**EN:** Europäische Norm

**EU:** European Union

**FI:** feature importance

**ICU:** intensive care unit

**ISO:** International Organization for Standardization

**QA:** quantitative assessment

**STARE-HI:** Statement on the Reporting of Evaluation studies in Health Informatics

**UCD:** user-centered design

**XAI:** explainable artificial intelligence

**XUI:** explanation user interface

*Edited by A Schwartz, M Balcarras; submitted 15.May.2025; peer-reviewed by U Johansson, V Jawa; comments to author 16.Oct.2025; revised version received 25.Nov.2025; accepted 25.Nov.2025; published 03.Feb.2026*

*Please cite as:*

*Jung I-C, Zerlik M, Schuler K, Sedlmayr M, Sedlmayr B*

*An Explanation User Interface for Artificial Intelligence-Supported Mechanical Ventilation Optimization for Clinicians: User-Centered Design and Formative Usability Study*

*JMIR Form Res* 2026;10:e77481

*URL:* <https://formative.jmir.org/2026/1/e77481>

*doi:* [10.2196/77481](https://doi.org/10.2196/77481)

*PMID:*

©Ian-C Jung, Maria Zerlik, Katharina Schuler, Martin Sedlmayr, Brita Sedlmayr. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 03.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.