



ISSN : 2350-0743



## RESEARCH ARTICLE

### BRIDGING THE FIDELITY GAP: A MULTI-PHASE USABILITY ENGINEERING APPROACH FOR VALIDATING AUTOMATED MONITORING SYSTEMS IN THE ICU

Ana Cristina Beitia Kraemer Moraes<sup>1,2</sup>; Larissa Oliveira Daneluz Vaz<sup>1,2</sup>; Inês Moraes Hirdes<sup>1,2</sup>; Tássia Helena Pozzobom Fellin<sup>1,2</sup>; Cintia Protzen Fonseca<sup>1</sup>; Ian Borges Quintana<sup>1</sup> and Rafael Guerra Lund<sup>2</sup>

<sup>1</sup>Catholic University of Pelotas, Pelotas, Brazil; <sup>2</sup>Federal University of Pelotas, Brazil

#### ARTICLE INFO

##### Article History:

Received 15<sup>th</sup> February, 2026  
Received in revised form  
24<sup>th</sup> March, 2026  
Accepted 19<sup>th</sup> April, 2026  
Published online 28<sup>th</sup> May, 2026

##### Keywords:

Urine Output Monitoring; Usability Testing; In-situ Simulation; Patient Safety; Medical Device Design.

##### \*Corresponding author:

Ana Cristina Beitia Kraemer Moraes

#### ABSTRACT

**Purpose:** Automated urine output monitoring enhances accuracy and traceability in critical care, yet its implementation is often hindered by ergonomic and environmental barriers. This study evaluated the safety and usability of an automated monitoring system through a Human Factors and User-Centered Design (UCD) approach to identify factors influencing clinical integration. **Methods:** A prospective, multi-phase study was conducted following IEC 62366-1/2 standards. First, a PRISMA-ScR scoping review was performed to establish design requirements. Subsequently, usability validation was conducted with 30 nursing professionals across a progressive fidelity framework: (I) formative laboratory evaluation; (II) summative testing; and (III) high-fidelity *in-situ* simulation within an active Intensive Care Unit (ICU). Usability was quantified via the System Usability Scale (SUS) and task performance analysis. **Results:** The scoping review identified a lack of structured human factors validation in existing technologies. In the validation phase, although no critical safety errors occurred, a significant "fidelity gap" was observed: perceived usability declined from "Good" in the laboratory (SUS = 77.5 +/- 14.9) to "Marginal" in the real ICU environment (SUS = 63.5 +/- 24.6). Physical ergonomic constraints, specifically device fixation and height-gravity dilemmas, were the primary barriers to optimal usability, rather than software interface logic. **Conclusion:** The evaluation of this automated monitoring technology highlights the critical role of multi-phase usability testing in identifying safety gaps that remain latent during controlled laboratory evaluations. The emergence of a 'fidelity gap' in the *in-situ* phase demonstrates that the clinical environment is a decisive variable in the performance of medical devices.

Copyright©2025, Ana Cristina Beitia Kraemer Moraes et al. 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.

**Citation:** Ana Cristina Beitia Kraemer Moraes, Larissa Oliveira Daneluz Vaz, Inês Moraes Hirdes et al. 2025. "Bridging the Fidelity Gap: A Multi-Phase Usability Engineering Approach for Validating Automated Monitoring Systems in the ICU". *International Journal of Recent Advances in Multidisciplinary Research*, 13,(05), 12471-12478.

## INTRODUCTION

Accurate urine output monitoring is a cornerstone in the management of critically ill patients, serving as a sensitive indicator of hemodynamic instability and a primary early marker for Acute Kidney Injury (AKI) (1). Despite its clinical significance, this measurement remains predominantly manual in global hospital practice—a method inherently susceptible to recording errors, parallax inaccuracy, and delayed recognition of oliguria (2). Such limitations can directly compromise the timing and efficacy of therapeutic interventions. To mitigate these gaps, automated systems utilizing capacitive sensors or weight-based scales (e.g., Serenno™, Clarity RMS™) have been developed, demonstrating superior precision over manual methods (3). Evidence suggests these technologies can detect oliguric episodes up to two hours earlier than standard nursing rounds, potentially revolutionizing AKI prevention (1,4). Nevertheless, widespread clinical adoption is frequently hindered not by sensor

performance, but by systemic barriers: high costs, complex training requirements, and, most critically, poor physical and ergonomic integration into the high-density ICU environment (4). Consequently, the usability of these devices emerges as a decisive determinant of patient safety. The interaction between clinicians and technology must transcend mere functional accuracy to comply with international usability engineering standards (IEC 62366-1) and risk management protocols (ISO 14971). A device that achieves technical precision but fails to account for the "divided attention tax" of the ICU introduces latent risks, such as alarm fatigue and use errors (5). Therefore, a robust validation framework must extend beyond the isolated laboratory setting to encompass the environmental complexities of realistic clinical contexts. In this context, the present study adopts a dual-objective approach: (I) to conduct a Scoping Review identifying functional requirements and prevalent design flaws in current automated technologies; and (II) to validate the safety and usability of an automated monitoring system through a progressive fidelity framework. This methodology encompasses formative evaluation,

summative testing, and high-fidelity *in-situ* simulation within an active ICU. By bridging the gap between engineering development and clinical reality, this research seeks to provide evidence-based insights for regulatory compliance and the safe implementation of monitoring innovations in high-complexity healthcare settings.

**Study Design:** A prospective, multi-phase usability validation study was conducted, grounded in a user-centered design framework. The protocol strictly adhered to IEC 62366-1/2:2020 standards for the application of usability engineering to medical devices. The research was structured into three sequential phases: (I) a Scoping Review to establish design inputs; (II) Formative Evaluation within a controlled Simulation Laboratory; and (III) Summative Evaluation through High-Fidelity *In-situ* Simulation.

**Phase I: Design Input (Scoping Review):** To identify functional requirements, clinical needs, and prevalent design gaps in automated urine output monitoring, a scoping review was performed in accordance with the PRISMA-ScR guidelines (6). Systematic search was conducted between April and December 2024 across nine databases: PubMed, Web of Science, Scopus, Cochrane, PROSPERO, Scielo, Embase, LILACS, and the Virtual Health Library (BVS). The search strategy utilized MeSH, Emtree, and DeCS descriptors combined with Boolean operators ("AND", "OR"). Key terms included: "Urine Automatic Measurement," "Measuring Urine Flow Rate," and "Automated Diuresis Measurement," cross-referenced with "Device," "Equipment," "Randomized Controlled Trials," and "Clinical Study." Equivalent descriptors in Portuguese and Spanish were applied to regional databases to ensure comprehensive coverage.

**Inclusion and Exclusion Criteria:** Inclusion Criteria: Studies were eligible if they addressed automated urine output measurement in hospitalized patients, with a focus on Intensive Care Units (ICUs). Selected publications were required to report on at least one of the following domains: technical validation, clinical performance, comparative analysis with manual methods, practical applicability, or integration with Electronic Health Records (EHR) and sensor technologies. Randomized controlled trials, observational studies, prospective/retrospective cohorts, and systematic reviews were included. Exclusion Criteria: Studies were excluded if they focused exclusively on manual methods, veterinary applications, non-functional prototypes, standalone accessories without sensing capabilities, or theoretical descriptions lacking empirical validation. Additionally, duplicates, conference abstracts without full data, editorials, and case reports were removed.

**Selection and Extraction Process:** The selection process was conducted independently by two reviewers across five stages: identification, title screening, abstract screening, full-text analysis, and final inclusion. Rayyan software was utilized for bibliography management and deduplication. Any discrepancies between reviewers were resolved through consensus. The synthesis of these findings directly informed the device's interface architecture, ergonomic requirements, and alarm parameters.

**Phase II: Device Description and Clinical Workflow:** The system under evaluation is an automated urine output monitor designed for seamless integration into the ICU workflow, facilitating KDIGO-based clinical decision-making.

- **System Architecture:** The device comprises an electronic bedside console (user interface), a fluid handling unit featuring an intermediate precision reservoir, and a standard disposable collection bag.

- **Clinical Functionality:** The system performs automated gravimetric or volumetric measurements, calculates weight-adjusted output (mL/kg/h), and maintains a continuous log of hemodynamic data every 30 minutes for up to 72 hours.
- **Risk Mitigation (Safety Features):** The software architecture includes multi-modal (visual and auditory) alarms to signal catheter obstructions, bag capacity limits, and power/battery status, adhering to safety-critical design principles.

### Phase III: Usability & Safety Validation

**Setting and Participants:** The study was conducted at the University Hospital, utilizing both the Health Simulation Center (HSIM) and an operational Intensive Care Unit (ICU). A total of 30 nursing professionals (nurses and technicians) participated. The sample size followed FDA guidance for human factors validation, ensuring sufficient power to detect use errors.

#### Protocol Testing

**Following a progressive fidelity approach, the validation was divided into three stages:**

- **Formative Phase (Lab Simulation):** An initial task analysis was performed with five participants to identify preliminary design flaws and refine the training protocol (7, 8, 9).
- **Summative High-Fidelity In-situ Simulation (n=30):** Participants performed critical tasks in a real hospital room using a simulated patient (high-fidelity manikin). This setting replicated environmental stressors (noise, space constraints) without compromising patient safety, allowing for direct observation of task errors (e.g., device fixation, assembly) (9).
- **Real Clinical Experience (Unsupervised) (n=15):** Following the simulation, a subgroup of 15 participants utilized the device during standard clinical care shifts with real patients. To preserve patient privacy and prevent the Hawthorne effect, direct task observation was not performed by researchers. Instead, usability data was captured retrospectively via self-reported feedback immediately after shift (1, 10).

#### Data Collection and Metrics

Prior to testing, participants received a standardized training session (manual review and video). Tasks evaluated included: device setup, patient data entry, monitoring interpretation, and alarm resolution.

- **Performance Metrics (Observed in Simulation):** Tasks were coded as "Success," "Recoverable Error" (difficulty overcome), or "Use Error" (failure) (7).
- **Subjective Assessment:** The System Usability Scale (SUS) was administered after the Summative Simulation and after the Real Clinical Experience. The SUS yields a score from 0–100, with scores >68 indicating acceptable usability (11–13).
- **Qualitative Feedback:** Open-ended questions were applied to identify potential hazards and ergonomic barriers (7, 14, 15).

**Statistical Analysis:** Quantitative data analysis was performed using PSpP Statistical Analysis Software, version 2.0 (GNU Project). Descriptive statistics (mean, standard deviation) summarized usability scores. The normality of data distribution was assessed using the Shapiro-Wilk test. To compare SUS scores between the simulated (summative) and real (*in-situ*) phases, the Mann-Whitney U test for independent samples was applied. Statistical significance was set at  $p < 0.05$  (9, 14, 16).

**Ethical Considerations:** The study complied with the Declaration of Helsinki and was approved by the Research Ethics Committee of the Catholic University of Pelotas under protocol number CAAE 7.092.467. All participants included in the study provided written informed consent.

## RESULTS

**Phase I: Design Inputs (Scoping Review):** Between April and December 2024, 553 publications were retrieved related to automated urine output measurement. After removing 17 duplicates, 536 titles and abstracts were screened. Of these, 474 were excluded for not meeting eligibility criteria. A total of 79 full-text articles were assessed, resulting in 15 studies included in the final qualitative synthesis. The analysis of these studies, published between 2009 and 2023, revealed a predominance of capacitive sensors and digital scale technologies. Crucially, while functional accuracy was frequently reported, none of the reviewed studies formally assessed usability according to IEC 62366-1 standards, and validated metrics (the System Usability Scale - SUS) were absent.

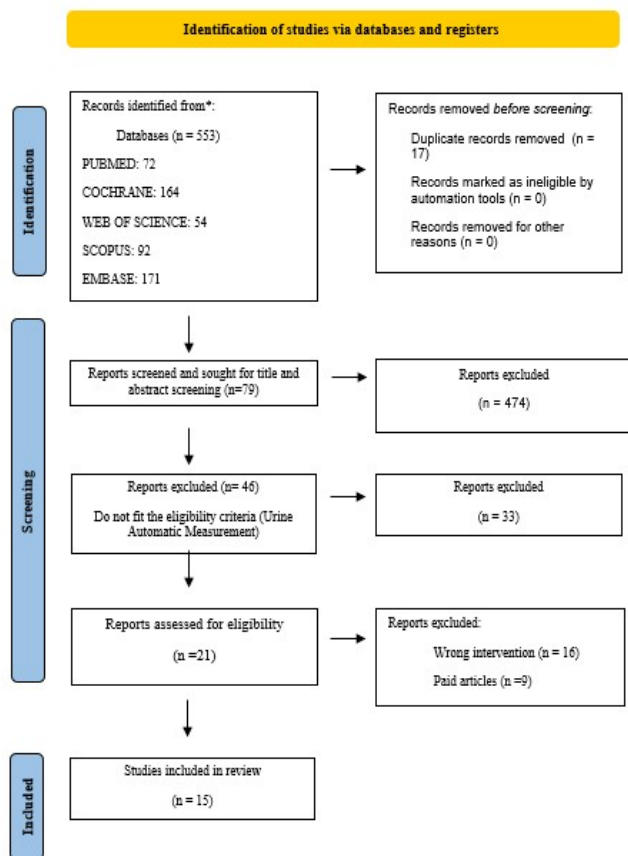


Figure 1. PRISMA flowchart of the included studies

This gap in the literature confirmed the need for a rigorous human factors evaluation for the new system. (Figure 1: PRISMA flowchart). Most studies were conducted in Israel (n=5) and the U.S. (n=4), with others in Sweden, Spain, and South Korea. Study designs included clinical trials (n=7), prospective studies (n=3), descriptive (n=2), retrospective (n=1), observational cohort (n=1), and one multicenter study, as summarized in Table 1. Technologies included capacitive sensors, digital scales with load cells, continuous flow sensors, and micro camera systems. Despite differing architectures, all aimed to automate urine output tracking, enhance precision, reduce manual error, and enable faster clinical responses.

**Trends and Clinical Validation:** Capacitive sensors emerged as promising in three studies. (3) validated a highly accurate system versus traditional urinometers;(5) demonstrated strong performance in pediatric ICU settings. Systems using digital scales (17) offered rapid response and alarm triggers for anomalies. Other innovations like micro cameras (2,18, 19) enabled real-time detection of critical events. Studies (10, 1) emphasized electronic integration and intelligent alerts. Devices were validated across adult/pediatric ICUs, operating rooms, and burn units. Main outcomes included urine volume accuracy, critical event timing, ease of use, and clinical impact.

To synthesize the main technological features, clinical application contexts, and operational functionalities of the automated urine output monitoring devices included in this systematic review, a comprehensive comparative table was developed (Table 2). This data systematization enabled the identification of development patterns, levels of technological maturity, digital integration strategies, and the potential clinical benefits of the analyzed devices. Despite technological progress, challenges remain calibration issues, training needs, clinical interference, and infrastructure constraints. Costs also limit adoption. High staff turnover, workload, and unstable clinical conditions affect device reliability (4,19). Devices demonstrated clinical viability highlighted intraoperative use (18) and showed system compatibility with hospital infrastructure(10, 20), but present lack FDA/CE/ANVISA certification. Low-cost prototypes (21) may fill gaps in under-resourced settings.

### Phase II: Formative Evaluation

The formative evaluation (n=5) conducted in the simulation lab identified distinct use errors related to the physical interface and software workflow. Key issues included:

- **Ergonomics:** Difficulty in assembling the reservoir components.
- **Interface:** Low interpretability of specific icons and poor screen readability under high-glare lighting.
- **Workflow:** Confusion regarding the "power on" sequence. These findings informed us of iterative design changes and the restructuring of the training protocol used in the subsequent summative phase.

### Phase III: Summative and In-situ Validation

**Quantitative Usability Metrics (SUS):** A total of 30 evaluations were conducted. The System Usability Scale (SUS) scores revealed a significant impact of environmental fidelity on perceived usability (Table 1). In the Summative Phase (Simulated ICU, n=15), the device achieved a mean SUS score of 77.5 (SD ± 12.0), placing it in the "Good" to "Excellent" range (Acceptability Range: High). However, in the In-situ Phase (Real ICU, n=15), the mean score decreased to 63.5, indicating that environmental constraints (space, lighting, workflow) introduced friction does not present in the laboratory setting(11-13).(Table 3)

### Task Performance and Error Analysis(Simulated Phase)

Task completion rates and error frequencies were analyzed during the Simulated and Training phases, where direct observation was feasible. Task completion was high (>95%), verifying the device's safety logic. However, observational data highlighted specific physical interaction difficulties (14, 15, 22). (Table 4)

Table 1. Summary of included studies according to study design, authorship, title, journal, country, and year

| Publication/Study Type                 | Author                     | Title   | Journal   | Country     | Year |
|--|----------------------------|---|---|-------------|------|
| Prospective Study                      | Hersch M <i>et al.</i>     | Accuracy and ease of use of a novel electronic urine output monitoring device compared with standard manual urinometer in the intensive care unit                       | Journal of Critical Care                                  | Israel      | 2009 |
| Descriptive Study                      | Otero A <i>et al.</i>      | A New Device to Automate the Monitoring of Critical Patients' Urine Output  | BioMed Research International                             | Spain       | 2014 |
| Prospective Study                      | Eklund A <i>et al.</i>     | Performance and user evaluation of a novel capacitance-based automatic urinometer compared with a manual standard urinometer after elective cardiac surgery.            | Critical Care   | Sweden      | 2015 |
| Clinical Trial Study                   | Goldman <i>et al.</i>      | A Novel Electronic Device for Measuring Urine Flow Rate: A Clinical Investigation   | Clinical Medicine Insights: Trauma and Intensive Medicine | Israel      | 2017 |
| Clinical Trial Study                   | Chang A.J <i>et al.</i>    | Validation of a real-time minute-to-minute urine output monitor and the feasibility of its clinical use for patients undergoing cardiac surgery.                        | Anesthesia & Analgesia                                    | USA         | 2017 |
| Retrospective and Observational Study  | Brotfain E <i>et al.</i>   | Urine flow rate monitoring in hypovolemic multiple trauma patients  | World Journal of Emergency Surgery                        | Israel      | 2017 |
| Clinical Trial Study                   | Yeh <i>et al.</i>          | Inaccuracy of Urine Output Measurements due to Urinary Retention in Catheterized Patients in the Burn ICU.  | Journal of Burn Care & Research                           | USA         | 2022 |
| Prospective observational cohort study | Slettengren <i>et al.</i>  | A Pilot Evaluation of a Capacitance-Based Automatic Urinometer in a Pediatric Intensive Care Setting  | Pediatric Critical Care Medicine                          | Sweden      | 2019 |
| Retrospective Review                   | Kushnir A <i>et al.</i>    | Improving fluid output monitoring in the intensive care unit  | Journal of Intensive Care Medicine                        | USA         | 2020 |
| Clinical Trial Study                   | Willner <i>et al.</i>      | Early identification of acute kidney injury in the ICU with real-time urine output monitoring: a clinical investigation.  | Nephrology  | Israel      | 2021 |
| Clinical Trial Study                   | Slettengren <i>et al.</i>  | Silicone oil decreases biofilm formation in a capacitance-based automatic urine measurement system.   | Sensors   | Sweden      | 2021 |
| Clinical Trial Study                   | Kushnir <i>et al.</i>      | Improving Fluid Output Monitoring in the Intensive Care Unit.   | <i>Journal of Intensive Care Medicine</i>                 | USA         | 2022 |
| Prospective Study                      | Oh S.K <i>et al.</i>       | Accuracy and availability of automated urine output monitoring in the operating room using a smart scale.   | Advances in Medical Sciences                              | South Korea | 2023 |
| Descriptive Study                      | Lafuente J.L <i>et al.</i> | Development of a urinometer for automatic measurement of urine flow in catheterized patients  | Plos One  | Spain       | 2023 |
| Prospective Observational Study        | Fishman G <i>et al.</i>    | Early detection of oliguric events in critically ill patients in the ICU with a novel continuous urine flow measurement device: results of an initial validation study. | Journal of Clinical Monitoring and Computing              | Israel      | 2023 |

Table 2 . Overview of Included Studies on Automated Urine Output Monitoring

| Study   | Year | Study Design                           | Population   | Intervention  | Comparator  | Outcomes   |
|---|------|--|--|---|---|--|
| Accuracy and ease of use of a novel electronic urine output monitoring device compared with standard manual urinometer in the intensive care unit                       | 2009 | Prospective Study                      | ICU patient (Intensive Care Unit patient)                    | Automated electronic meter                                  | Manual urinometer(manual urine output meter)              | Ease of use; superior accuracy                         |
| A New Device to Automate the Monitoring of Critical Patients' Urine Output  | 2014 | Descriptive Study                      | Critically ill patients                                      | Automated urine output device                               | Standard technique  | Reduced reading errors; automation                     |
| Performance and user evaluation of a novel capacitance-based automatic urinometer compared with a manual standard urinometer after elective cardiac surgery.            | 2015 | Prospective Study                      | Postoperative period of cardiac surgery                      | Automatic capacitive sensor (orAutomated capacitive sensor) | Manual urine output meter                                 | High precision and user acceptance                     |
| A Novel Electronic Device for Measuring Urine Flow Rate: A Clinical Investigation   | 2017 | Clinical Trial Study                   | Adult ICU (Adult Intensive Care Unit)                        | Electronic flow sensor                                      | Conventional reading (orStandard reading)                 | High accuracy; usability                               |
| Validation of a real-time minute-to-minute urine output monitor and the feasibility of its clinical use for patients undergoing cardiac surgery.                        | 2017 | Clinical Trial Study                   | Patients undergoing cardiac surgery                          | Real-time monitor   | Manual recording  | Minute-by-minute monitoring; early detection           |
| Urine flow rate monitoring in hypovolemic multiple trauma patients  | 2017 | Retrospective and Observational Study  | Hypovolemic polytrauma                                       | Continuous flow sensor                                      | No  | Effective monitoring in critical contexts              |
| Inaccuracy of Urine Output Measurements due to Urinary Retention in Catheterized Patients in the Burn ICU.  | 2022 | Clinical Trial Study                   | Patients in the burn ICU (Burn Intensive Care Unit patients) | Conventional system   | No  | Inaccuracies due to urinary retention                  |
| A Pilot Evaluation of a Capacitance-Based Automatic Urinometer in a Pediatric Intensive Care Setting  | 2019 | Prospective observational cohort study | Pediatric ICU (Pediatric Intensive Care Unit)                | Capacitive sensor   | No  | High sensitivity to small volumes                      |
| Improving fluid output monitoring in the intensive care unit  | 2020 | Retrospective Review                   | Adult ICU (Adult Intensive Care Unit)                        | Automated digital system                                    | Traditional medical record (orPaper-based medical record) | Continuous tracking; failure reduction                 |
| Early identification of acute kidney injury in the ICU with real-time urine output monitoring: a clinical investigation.  | 2021 | Clinical Trial Study                   | Adult ICU(Adult Intensive Care Unit)                         | Real-time automated system                                  | Serum creatinine  | Early detection of acute kidney injury (AKI)           |
| Silicone oil decreases biofilm formation in a capacitance-based automatic urine measurement system  | 2021 | Clinical Trial Study                   | Experimental system  | Silicone sensor   | Silicone-free sensor                                      | Biofilm reduction; biosafety                           |
| Improving Fluid Output Monitoring in the Intensive Care Unit  | 2022 | Clinical Trial Study                   | Adult ICU (Adult Intensive Care Unit)                        | Electronic system with alerts                               | Without a comparator                                      | Enhanced precision; clinical support                   |
| Accuracy and availability of automated urine output monitoring in the operating room using a smart scale.   | 2023 | Prospective Study                      | Surgical center  | Smart digital scale   | Standard reading  | High sensitivity and immediate response                |
| Development of a urinometer for automatic measurement of urine flow in catheterized patients  | 2023 | Descriptive Study                      | Laboratory / prototype                                       | Non-invasive continuous flow system                         | Without a comparator (or No comparator)                   | Low cost; large-scale application                      |
| Early detection of oliguric events in critically ill patients in the ICU with a novel continuous urine flow measurement device: results of an initial validation study. | 2023 | Prospective Observational Study        | Adult ICU (Adult Intensive Care Unit)                        | Continuous digital sensor                                   | Clinical team   | Early detection of oliguria (up to 2 hours in advance) |

**Table 3. System Usability Scale (SUS) Results by Environment**

| Environment               | N  | Mean Score | SD ( $\pm$ ) | Adjectival Rating |
|---------------------------|----|------------|--------------|-------------------|
| Simulated ICU (Summative) | 15 | 77.5       | 14,9         | Good / Excellent  |
| Real ICU (In-situ)        | 15 | 63.5       | 24,6         | Marginal / OK     |

**Table 4. Frequency of Use Difficulties (Observed during Simulation/Training)**

| Task                      | Difficulty Rate (%) | Description of Issue  |
|---------------------------|---------------------|---|
| Device Fixation & Height  | 40.0% (12/30)       | Participants reported difficulty fixing the equipment at the ideal height on the bed or support stand.                |
| Buoy & Reservoir Assembly | 23.3% (7/30)        | Users struggled specifically with verifying the buoy position or snapping the intermediate reservoir into the module. |
| Interface Configuration   | 20.0% (6/30)        | Challenges reported in configuring preferences (brightness/sound) or navigating specific sub-menus.                   |
| Transport Mode Activation | 16.7% (5/30)        | Difficulty in pausing measurement and correctly activating the transport mode workflow                                |

**Statistical Analysis Results:** The Shapiro-Wilk test indicated a non-normal distribution of SUS scores, particularly in the in-situ phase, validating the use of non-parametric analysis. Comparing the Simulated and Real phases, the Mann-Whitney U test showed that despite the numerical decrease in mean usability (from 77.5 to 63.5), the difference was not statistically significant ( $U = 84.5$ ,  $p = 0.091$ ). This lack of statistical significance, coupled with the high standard deviation in the real environment ( $SD \pm 24.6$ ), suggests that clinical environmental factors introduced high variability in user experience, rather than a uniform failure of the device interface.

#### Qualitative User Experience (Thematic Analysis)

To complement the quantitative data, open-ended feedback collected from the "User Experience" forms (Real Context) and post-training interviews was analyzed. While direct observation of errors was not performed during real patient care, the thematic analysis of user comments confirmed the barriers identified in the simulation. Three dominant themes emerged:

- Theme 1: Physical Ergonomics - The "Height" Problem - The most frequent complaint in the real ICU setting concerned the vertical positioning. Users reported that the low support stand necessary for gravity drainage forced them into uncomfortable postures to read the display.

*"Height for use, support is too low... difficult to visualize due to height."* (Participant 29 - Real Context)

- Theme 2: Interface Responsiveness Participants noted that while the menu logic was sound, the touchscreen hardware required firm pressure, which was occasionally difficult when wearing gloves during rapid procedures.

*"The touch sometimes freezes... touchscreen has a bit of resistance."* (Participant 10)

- Theme 3: Clinical Value vs. Usability A distinct pattern emerged where users distinguished "clinical utility" from "ease of use". Despite the ergonomic complaints, the automated accuracy was highly valued for patient safety.

*"Precision of urinary output, more reliable control... great utility in monitoring."* (Participant 8)

## DISCUSSION

This study employed a multi-phase usability engineering framework to validate a novel automated urine output

monitoring system (23). Our findings demonstrate that while the device interface is intuitive, its integration into the high-complexity ICU environment reveals a "Fidelity Gap" where physical and environmental constraints significantly degrade perceived usability.

**The Fidelity Gap and Environmental Sensitivity:** A pivotal finding was the significant divergence in System Usability Scale (SUS) scores between the simulated environment (77.5 \ 14.9) and the *in-situ* clinical phase (63.5 \ 24.6). This drop and the accompanying spike in Standard Deviation validates the "fidelity gap" phenomenon. In the laboratory, users interacted with the device in isolation; in the real ICU, the device competed for physical space and cognitive resources against infusion pumps, ventilators, and clinical alarms. This high variability suggests that usability is not an intrinsic property of the device, but a dynamic emergent property of the interaction between user, task, and environment. While the SUS confirmed functional stability, its unidimensional nature likely masked specific stressors. The inclusion of a multidimensional tool, such as the NASA-TLX, could have further deconstructed whether this workload was primarily perceptual (visibility/audibility) or cognitive (information processing under stress) (25-27). The discrepancy indicates that a system may be 'technically functional' in a lab while approaching a safety-critical threshold in practice, where the "divided attention tax" increases the risk of omission errors or delayed responses.

**Ergonomics vs. Physics: The Design Conflict:** The most frequent obstacle, affecting 40.0% of participants, was "Device Fixation and Height." This reveals a fundamental design conflict: while gravity drainage necessitates a low reservoir position, this placement forces the digital display into a suboptimal ergonomic zone. Unlike previous studies focused solely on sensor accuracy (1, 10), this *in-situ* analysis proves that physical integration is as critical as electronic precision. To mitigate this, future iterations should decouple the display unit from the reservoir, allowing for optimal sensor placement and ergonomic data visualization.

**Safety, Efficiency, and Regulatory Alignment:** Despite the ergonomic friction, no critical safety errors occurred, aligning with ISO 14971 risk management principles. The issues identified such as small icons for "Transport Mode" resulted in operational delays (efficiency loss) rather than hazardous situations (effectiveness failure). Furthermore, this study fills a methodological gap, as most devices in our scoping review lacked structured validation per IEC 62366-1.

Data also suggests that mechanical tasks, such as "buoy positioning," are highly sensitive to instruction quality (28). The 23.3% difficulty rate in this task suggests that modular video training may be superior to static manuals for hardware-heavy medical technologies.

### Limitations

This study's safety focused design introduced certain trade-offs. The reliance on retrospective self-reporting (*in-situ*) was necessary to avoid the Hawthorne effect and respect patient privacy, yet it may result in the underreporting of minor use errors compared to the observed simulated phase. Additionally, while the *in-situ* setup replicated physical constraints, the use of a simulated patient cannot fully capture the emotional urgency and peak cognitive load of managing a hemodynamically unstable patient. Finally, while our sample (n=30) is sufficient to detect the majority of usability problems, the findings may not fully generalize across nursing populations with diverse levels of digital literacy (8).

## CONCLUSION

The evaluation of this automated monitoring technology highlights the critical role of multi-phase usability testing in identifying safety gaps that remain latent during controlled laboratory evaluations. The emergence of a 'fidelity gap' quantified by the significant decline in usability scores in the *in-situ* phase demonstrates that the clinical environment is a decisive variable in the performance of medical devices. Our findings reveal that even when interface logic is technically sound, physical and ergonomic integration remains a primary failure point in high-complexity settings. From a human factors perspective, the study identifies a fundamental design conflict between mechanical requirements (gravity-dependent drainage) and ergonomic visualization. Resolving this dilemma through hardware decoupling is essential to mitigate the 'divided attention tax' imposed on nursing staff. Furthermore, the robust safety profile observed, despite ergonomic friction, underscores the importance of risk-oriented design in preventing critical use errors. In conclusion, this study validates a methodological approach for the rigorous assessment of medical innovations. It emphasizes that sustainable adoption of monitoring technologies hinges on prioritizing human factors and environmental compatibility over mere electronic precision. These insights provide a scalable framework for future clinical validations, aiming to bridge the gap between technical functionality and safe, efficient bedside practice.

**Industry Involvement and Independence Statement:** This study was conducted as part of a public innovation project in partnership with industry. All phases of the usability testing including study design, participant recruitment, data collection, and analysis were conducted independently by the academic research team. The study complied with all applicable standards for research integrity and medical device evaluation.

### Author Contributions

A. C. B. K. M: Conceptualization, Methodology, Writing – Original Draft, Review, Supervision. L. O. D. Data Curation, Investigation, Writing, Formal Analysis, Visualization. I. M.

H.: Investigation, Data Analysis, Methodology. T. P. F: Data Interpretation, Validation, Software, Data Management. C. P. F.; L.B.: Writing – Review & Editing. L.B. C.D.N; E.G.S; R.G.L: Formal Analysis, Review, Visualization. All authors have read and approved the final manuscript.

### Statements and Declarations

**Competing Interests:** The authors declare that the assessment described in this study was part of a technology evaluation project supported by (UCPEL/LIFEMED). However, this manuscript was independently conceived as a methodological study to contribute to the field of usability science and human factors. The authors declare no financial or personal relationships that could have influenced the objectivity of the scientific findings presented herein.

### Funding

This work was partially supported by the *Secretaria de Inovação, Ciência e Tecnologia do Governo do Estado do Rio Grande do Sul* (SICT-RS) through the "Inova RS" program (EDITAL SICT n° 02/2022), under research grant no. 22/2500-0000232-0 for the technical evaluation of the technology. However, the funding source had no role in the study design, data analysis, interpretation of usability results, or the decision to submit this methodological manuscript for publication.

**AI tools:** During the preparation of this work the author(s) used Google Gemini pro in order to assist with language refinement and structuring. After using this tool/service, the author(s) reviewed and edited the content as needed and took(s) full responsibility for the content of the published article.

**Clinical trial number:** Not applicable.

**Consent to Participate:** Informed consent was obtained from all individual participants included in the study.

## REFERENCES

- Goldman, M., *et al.* (2009). Accuracy and ease of use of a novel electronic urine output monitoring device compared with standard manual urinometer in the intensive care unit. *Journal of Critical Care*, 24(4), 489–496.
- Chang, D., Holman, S. D., Bear, M., & Lin, H. M. (2014). A new device to automate the monitoring of critical patients' urine output. *Journal of Medical Devices*.
- Eklund, M., Johansson, P. I., & Stenqvist, O. (2015). Performance and user evaluation of a novel capacitance-based automatic urinometer compared with a manual standard urinometer after elective cardiac surgery. *Acta Anaesthesiologica Scandinavica*, 59(2), 255–262.
- Yeh, H., *et al.* (2022). Inaccuracy of urine output measurements due to urinary retention in catheterized patients in the burn ICU. *Burns*, 48(4), 947–954.
- Slettengren, M., Andersson, P. O., & Jonsson, L. (2019). A pilot evaluation of a capacitance-based automatic urinometer in a pediatric intensive care setting. *Pediatric Critical Care Medicine*, 20(5), e221–e225.

- PRISMA. (2015). *Preferred Reporting Items for Systematic Reviews and Meta-Analyses*. <http://www.prisma-statement.org>
- Lageat, C., Lafontaine, A., & Perrot, S. (2021). Formative and validation human factors studies of a new disposable autoinjector. *Expert Opinion on Drug Delivery*, 18(11), 1761–1775. <https://doi.org/10.1080/17425247.2021.1954906>
- Faulkner, L. (2003). Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments, & Computers*, 35(3), 379–383. <https://doi.org/10.3758/BF03195514>
- Aiyegbusi, O. L. (2020). Key methodological considerations for usability testing of electronic patient-reported outcome (ePRO) systems. *Quality of Life Research*, 29, 325–333. <https://doi.org/10.1007/s11136-019-02372-0>
- Kushnir, M., Lavon, A., & Fishman, R. (2022). Improving fluid output monitoring in the intensive care unit. *Critical Care Medicine*, 50(1), e75–e82.
- Brooke, J. (1996). SUS: A quick and dirty usability scale. In P. W. Jordan et al. (Eds.), *Usability evaluation in industry* (pp. 189–194). Taylor & Francis.
- Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean. *Journal of Usability Studies*, 4(3), 114–123.
- Cunningham, S. (2025). SUS Calculator. <https://stuart-cunningham.github.io/sus/>
- Aune A, Valberg M, Iversen P, Thoresen M, Borresen A. Iterative development, validation, and certification of a smartphone system to assess neonatal jaundice: development and usability study. *JMIR Pediatr Parent*. 2023;6:e40463. doi:10.2196/40463.
- Candidori, S., Lopes, L. A. F., Lima, P. M., & Borsari, G. (2024). Usability and performance assessment of a new medical device for postpartum hemorrhage. *Applied Ergonomics*, 117, 104223. <https://doi.org/10.1016/j.apergo.2023.104223>
- Santos-Costa P, Alves M, Sousa C, et al. Nurses' Involvement in the Development and Usability Assessment of an Innovative Peripheral Intravenous Catheterisation Pack: A Mix-Method Study. *Int J Environ Res Public Health*. 2022;19(17):11130. Published 2022 Sep 5. doi:10.3390/ijerph191711130.
- Oh SK, Kim D, Kim J, You B, Oh HS, Kwon C, Lee J, Oh SH. Accuracy and availability of automated urine output monitoring in the operating room using a smart scale. *Adv Med Sci*. 2023 Sep;68(2):265-269. doi: 10.1016/j.advms.2023.07.003. Epub 2023 Aug 22. PMID: 37619439.
- Chang AJ, Nomura Y, Barodka VM, Hori D, Magruder JT, Katz NM, Berkowitz DE, Hogue CW. Validation of a Real-Time Minute-to-Minute Urine Output Monitor and the Feasibility of Its Clinical Use for Patients Undergoing Cardiac Surgery. *Anesth Analg*. 2017 Dec;125(6):1883-1886. doi: 10.1213/ANE.0000000000002217. PMID: 29190218; PMCID: PMC5726301.
- Fishman G, Ram E, Gorfil D, Kassif Y, David R, Hershko T, Malbrain MLNG, Singer P, Sessler DI. Early detection of oliguric events in critically ill patients in the ICU with a novel continuous urine flow measurement device: results of an initial validation study. *J Clin Monit Comput*. 2023 Oct;37(5):1341-1349. doi: 10.1007/s10877-023-00991-w. Epub 2023 Apr 7. PMID: 37027058; PMCID: PMC10080177.
- Willner D, Goldman A, Azran H, Stern T, Kirshenbom D, Rosenthal G. Early identification of acute kidney injury in the ICU with real-time urine output monitoring: a clinical investigation. *BMC Nephrol*. 2021 Aug 26;22(1):293. doi: 10.1186/s12882-021-02485-w. PMID: 34445954; PMCID: PMC8394570.
- Lafuente JL, González S, Puertas E, Gómez-Tello V, Avilés E, Albo N, Mateo C, Beunza JJ. Development of a urinometer for automatic measurement of urine flow in catheterized patients. *PLoS One*. 2023 Aug 31;18(8):e0290319. doi: 10.1371/journal.pone.0290319. PMID: 37651353; PMCID: PMC10470914.
- Su, J. M., Yu, T. T., Hsu, C. H., & Tsai, Y. H. (2023). A mobile-based airway clearance care system using deep learning-based vision technology. *Digital Health*, 9, 20552076231207206. <https://doi.org/10.1177/20552076231207206>
- Beltzer A, Kölle J, Gil Miró M, et al. Benefits of usability evaluation in the development process of diabetes technologies using the example of a continuous glucose monitoring system prototype. *J Diabetes Sci Technol*. 2024. doi:10.1177/19322968241238146
- Santos-Costa P, Alves M, Sousa C, et al. Nurses' Involvement in the Development and Usability Assessment of an Innovative Peripheral Intravenous Catheterisation Pack: A Mix-Method Study. *Int J Environ Res Public Health*. 2022;19(17):11130. Published 2022 Sep 5. doi:10.3390/ijerph191711130.
- Liu K, Chen J, Smith B, Hansen M, O'Brien R. Heuristic evaluation and simulated use testing of infusion pumps to inform pump selection. *Int J Med Inform*. 2019; 131:103932. doi:10.1016/j.ijmedinf.2019.103932.
- Black CL, Zhou J, Miller D, Crook S, Greaves S. Heuristic evaluation of portable pulse oximeters for domiciliary use: implications for its use in assessing medical device usability. *Smart Health*. 2023; 27:100357. doi:10.1016/j.smhl.2022.100357.
- Beitia Kraemer Moraes AC, Yusef Bakri SF, Karlinski Scherer C, Daneluz Vaz LO, Yonara da Silva Galhardo J, Granemann Souza E, Nascimento CDDD, Lund RG. Systematic review of usability evaluations in medical devices: Methodological choices, heuristic application, and confounding factors. *Appl Ergon*. 2026 Jan 2;134:104715. doi: 10.1016/j.apergo.2025.104715. Epub ahead of print. PMID: 41483527.
- Rätz, R., Mayer, A., & Müller, S. (2024). Development and evaluation of a haptic hand and forearm trainer. *Frontiers in Neurobotics*, 18, 1351700. <https://doi.org/10.3389/fnbot.2024.1351700>

\*\*\*\*\*