

BMJ Open Effects of the Visual Patient Avatar on vital sign deviations and audible alarm burden in perioperative care: a dual-centre, quasi-experimental pre-post large-scale observational study protocol (NewYork-Presbyterian/Weill Cornell and University Hospital Zurich)

Silis Y Jiang,¹ Tadzio R Roche,^{2,3} Krzysztof Cybulski,⁴ Gaspar Dugac,⁴ Lukas Meier ,⁵ Virginia E Tangel,^{1,6} Max Ebensperger,⁷ Andreas Maskos,^{2,3} Michael Tucci,^{2,3} Christoph B Nöthiger,^{2,3} Markus Kalisch,⁴ Zachary A Turnbull,¹ David W Tscholl ^{2,3}

To cite: Jiang SY, Roche TR, Cybulski K, *et al.* Effects of the Visual Patient Avatar on vital sign deviations and audible alarm burden in perioperative care: a dual-centre, quasi-experimental pre-post large-scale observational study protocol (NewYork-Presbyterian/Weill Cornell and University Hospital Zurich). *BMJ Open* 2026;**16**:e121099. doi:10.1136/bmjopen-2026-121099

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<https://doi.org/10.1136/bmjopen-2026-121099>).

Received 12 April 2026
Accepted 04 August 2026



© Author(s) (or their employer(s)) 2026. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ Group.

For numbered affiliations see end of article.

Correspondence to

Dr David W Tscholl;
david.tscholl@usz.ch

ABSTRACT

Introduction Perioperative monitoring requires clinicians to integrate multiple physiological data streams while managing interruptions and audible alarms. The Visual Patient Avatar (VPA) is an avatar-based monitor visualisation that integrates vital signs and sensor states into an animated virtual patient. This protocol evaluates whether VPA implementation is associated with objective changes in monitoring and alarm outcomes during routine perioperative care.

Methods and analysis This investigator-initiated, dual-centre, quasi-experimental pre-post observational study uses routine perioperative monitoring and alarm data from NewYork-Presbyterian/Weill Cornell, USA and University Hospital Zurich, Switzerland. It includes approximately 58 000 anaesthesia cases (16 000 and 42 000, respectively); adult and paediatric cases will be analysed separately. Data span pre-implementation, adaptation and post implementation phases; technical deployment intervals are excluded. Each anaesthesia case is one unit of analysis. The primary outcome is the duration of vital-sign deviation per 60 min of valid measurement. Secondary outcomes are a time-based area under the curve metric and audible alarm burden (count, time and reaction time). Multivariable models will adjust for patient, procedure and care-context covariates. Analyses will be performed separately by centre.

Ethics and dissemination The Weill Cornell Medicine Institutional Review Board approved the study and granted a waiver of informed consent (IRB No. 22-05024803). At University Hospital Zurich, only cases from patients with written general consent for research use of routinely collected health data were included. The Cantonal Ethics Committee Zurich issued a declaration of non-jurisdiction (BASEC Req-2021-00756; 6 July 2021), and the University Hospital Zurich Data Governance Board approved data

STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ This dual-centre study uses routinely collected perioperative monitoring and alarm data from two tertiary academic hospitals on different continents.
- ⇒ The large sample and heterogeneous care areas support robust evaluation of Visual Patient Avatar (VPA) implementation under real-world conditions.
- ⇒ The quasi-experimental pre-post design cannot establish causality and remains vulnerable to temporal confounding.
- ⇒ Centre-specific alarm configurations and VPA thresholds limit direct comparison of absolute alarm rates, so analyses will be performed separately by centre.
- ⇒ The outcomes are process-level monitoring and alarm metrics rather than patient-centred clinical outcomes.

export and anonymisation (DGR0000504; 31 January 2025). A data-sharing agreement was established on 11 April 2023. Data-collection windows are complete; extraction, linkage and quality control are ongoing, and no comparative outcome analyses have been performed. Findings will be disseminated through peer-reviewed publications and conferences.

INTRODUCTION

Continuous physiological patient monitoring is universally regarded as a cornerstone of safe anaesthesia and surgery, forming the backbone of perioperative patient safety.^{1–3}

Perioperative patient monitoring requires clinicians to continuously integrate multiple

physiological data streams, including heart rate, blood pressure, oxygen saturation and end-tidal carbon dioxide, while managing time pressure, interruptions and competing tasks. Conventional monitors typically present vital signs as separate numerical values and waveforms that clinicians must sequentially read and mentally integrate into an overall assessment of the patient's physiological state. This process can impose substantial cognitive demands, particularly in dynamic, high-acuity clinical environments.^{4–12}

Audible alarms are intended to alert clinicians to potentially hazardous situations, but are also a major contributor to alarm fatigue. High alarm rates and prolonged alarm sounding may distract clinicians, increase cognitive workload and reduce responsiveness to clinically relevant events.^{13–15} Beyond alarm burden, time spent outside predefined safe ranges for key physiological variables, such as intraoperative hypotension or hypothermia, has been associated with adverse outcomes in perioperative and critical care research.^{16–18} These findings motivate efforts to support earlier recognition of physiological deviations and timely corrective action without adding to cognitive load.

The Philips (Koninklijke Philips N.V., Amsterdam, The Netherlands) Visual Patient Avatar (VPA) is an avatar-based patient monitoring visualisation designed to support clinicians' situation awareness.^{8 11 19–21} The system displays multiple vital signs and sensor states simultaneously on a single animated virtual patient, while conventional numerical values and waveforms remain available. Changes in physiological status are conveyed through intuitive visual cues, such as skin colour for oxygen saturation and pulsation dynamics for pulse rate and blood pressure.

This approach builds on situation-awareness-oriented design and human visual perception by mapping vital signs and sensor states to integrated visual features such as colour, shape, motion and spatial location.²² Independent human-factors research on graphical and configural patient-monitoring displays has shown that integrated visual displays can improve the speed and accuracy of physiological data interpretation in simulated or controlled clinical contexts.^{23–27} A recent independent systematic review of VPA research similarly concluded that the approach is promising for supporting perception and comprehension of monitoring data, while emphasising that most available evidence remains laboratory- or simulation-based.²³

Most VPA-specific empirical studies to date have been conducted by groups involved in the development and clinical implementation of the technology. This partly reflects that VPA was developed by members of the author group and has only recently entered broader clinical use. Although these studies suggest improved information transfer, diagnostic confidence and perceived workload compared with conventional monitoring displays,^{9 10 12 28–34} independent VPA-specific evidence remains limited. As adoption expands beyond the development centres,

additional independent evaluations should become feasible. Meanwhile, the concentration of evidence and the lack of objective real-world outcome data reinforce the need for transparent, prespecified evaluation with explicit conflict-of-interest disclosure and analytical safeguards.

This protocol describes a dual-centre evaluation at NewYork-Presbyterian Hospital/Weill Cornell Medical Center (NYP/WCM; New York, USA) and University Hospital Zurich (USZ; Zurich, Switzerland). The study aims to assess whether routine clinical implementation of VPA is associated with (1) reduced time spent outside predefined vital-sign thresholds and (2) reduced audible alarm burden across diverse perioperative environments, including operating rooms, post-anaesthesia care units and procedural areas (figure 1).

METHODS AND ANALYSIS

Study design

This investigator-initiated, dual-centre, quasi-experimental, large-scale, pre-post observational study uses routinely collected perioperative monitoring and alarm data to evaluate the real-world implementation of VPA. At each centre, the study follows three prespecified phases: a pre-implementation phase, an adaptation phase during staged installation and training and a post implementation phase after implementation is complete. Short technical deployment intervals between formal phase windows are excluded from analyses.

The unit of analysis is the individual anaesthesia case. The primary comparison is between the pre-implementation and post implementation phases; the adaptation phase is excluded from primary analyses to reduce contamination from installation, training and early adoption effects. Although outcome data are analysed retrospectively, the implementation phases, outcome definitions and analysis strategy were defined prospectively before comparative outcome analyses.

Reporting guidelines

This protocol was revised in accordance with the Transparent Reporting of Evaluations with Nonrandomized Designs (TREND) checklist and the Reporting of Studies Conducted using Observational Routinely Collected Health Data (RECORD) extension of Strengthening the Reporting of Observational Studies in Epidemiology (STROBE), where applicable to a protocol using routinely collected health data.^{35 36} Completed TREND and RECORD checklists accompany the revised submission.

Rationale for the pre-post study design

A pre-post design was chosen because VPA is implemented as a software and workflow change at the level of patient-monitoring infrastructure and affects entire care areas rather than individually randomisable patients or clinicians. Randomisation at the patient or provider level would have been operationally infeasible and risked

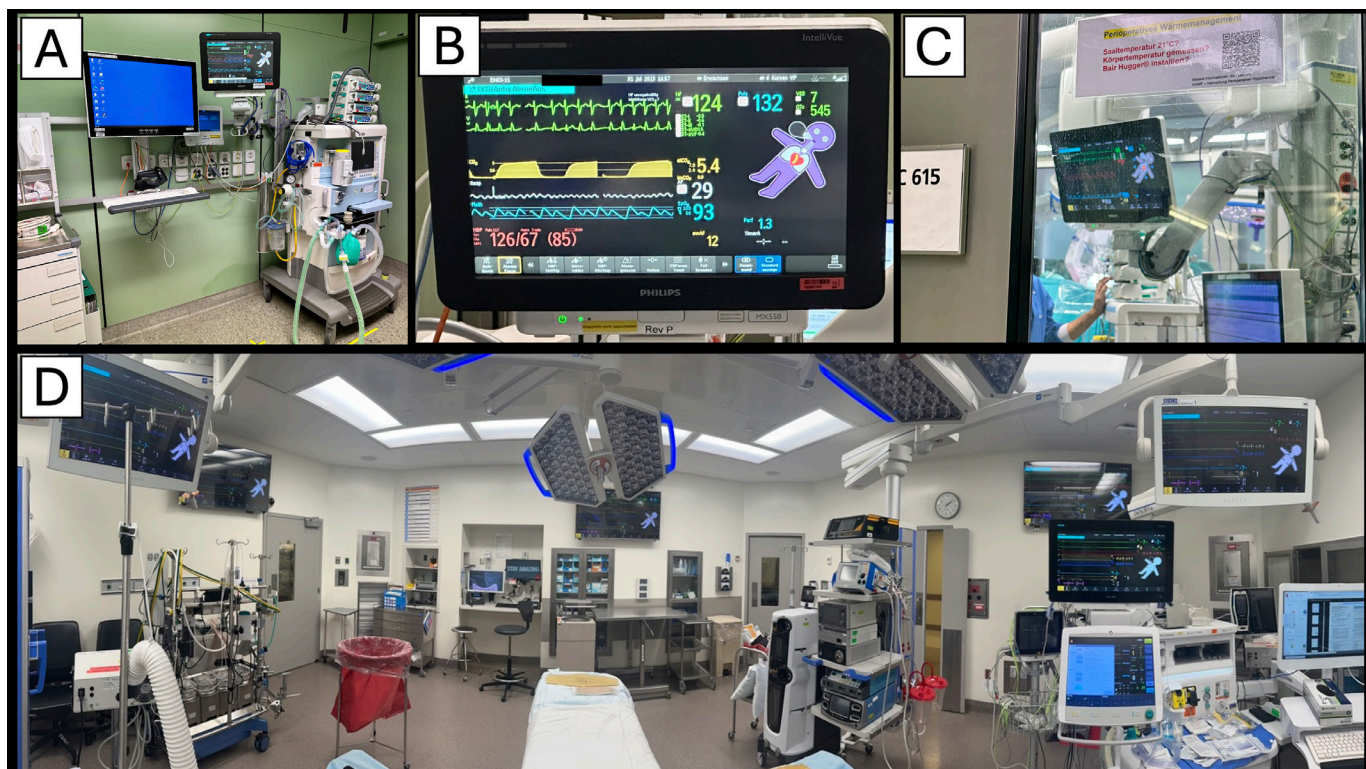


Figure 1 Clinical deployment of the Visual Patient Avatar (VPA) in routine perioperative monitoring. (A–D) Examples of VPA integration across operating room and post-anaesthesia care settings at University Hospital Zurich and NewYork-Presbyterian/Weill Cornell.

contamination, as clinicians move between rooms, beds and shifts.

Interrupted time-series analysis was considered but was not selected as the primary design. Deployment in large academic hospitals required sequential software updates to individual monitors over several days, followed by training and workflow adaptation; implementation therefore did not occur at one sharply defined time point. Routine clinical operations did not permit simultaneous activation across all devices or experimental on-off switching, and the exposure state could not be standardised across all monitors, rooms and time points. These conditions limit the interpretability of a conventional single-interruption analysis. We therefore treat the study as an implementation-level observational evaluation and will avoid causal claims. To address secular trends, we will describe patient and procedure characteristics by phase, adjust for prespecified covariates and conduct sensitivity analyses incorporating calendar time where feasible.

Patient and public involvement

Patients and the public were not involved in the design, conduct, reporting or dissemination plans of this study.

Study setting and participating centres

The study is conducted at two tertiary academic hospitals. NYP/WCM is a large academic medical centre in New York, USA, with over 4000 beds and comprehensive perioperative services; it was the first centre in North

America to clinically implement VPA. USZ is a tertiary academic hospital in Zurich, Switzerland, with approximately 900 beds and is the primary teaching hospital of the University of Zurich; it was the first centre worldwide to implement VPA in routine clinical practice.

Study phases, timeline and current project status

Phase timing differs between centres because regulatory approval and clinical deployment occurred at different times. The overall prespecified study window spans 31 August 2022 to 28 February 2026. Centre-specific phase dates and anticipated case numbers are provided in [table 1](#). Technical deployment intervals are not included in the primary pre-implementation or post implementation analyses. The adaptation phase lasted approximately 6 months at each centre. Its duration was prespecified based on the local rollout and training schedules to allow sequential installation, staff exposure and workflow uptake; it was not assumed to eliminate all learning effects. Because learning may continue after the formal adaptation phase, early-versus-late post implementation sensitivity analyses are prespecified.

At the time of revised protocol submission, the prespecified data-collection windows are complete at both centres. At USZ, data extraction has been completed, and dataset preparation is ongoing. At NYP/WCM, data extraction, linkage and quality control are ongoing. No prespecified comparative outcome analyses have been performed at either centre.

Table 1 Study phases, planned case numbers and Visual Patient Avatar (VPA) deployment by centre

	NewYork-Presbyterian/Weill Cornell	University Hospital Zurich
Pre-implementation phase timeframe	1 September 2024 to 14 February 2025 (167 days)	31 August 2022 to 31 January 2023 (154 days)
Technical deployment interval (excluded from primary phase analyses)	15 February 2025 to 28 February 2025	1 February 2023 to 17 February 2023
Adaptation phase timeframe	1 March 2025 to 31 August 2025 (184 days)	18 February 2023 to 15 August 2023 (179 days)
Post implementation phase timeframe	1 September 2025 to 28 February 2026 (181 days)	16 August 2023 to 12 February 2024 (181 days)
Approximate number of total cases	16 000	42 000
Total number of monitors with VPA in operating rooms and post-anaesthesia care units of the study centre	49	126
Technical deployment intervals are excluded from the primary pre-post comparison.		

Participating clinical care areas

The study includes perioperative care areas with differing workflows, professional groups and patient populations: operating rooms and post-anaesthesia care units at both centres and procedural areas at USZ only, including neuro-interventional radiology, interventional pulmonology and emergency trauma care.

Study participants: inclusion and exclusion criteria

All adult and paediatric anaesthesia cases conducted in participating care areas during the prespecified phase windows with available perioperative monitoring data will be screened for inclusion. The primary adult analysis will include cases in patients aged 18 years or older with a valid anaesthesia case identifier, usable case timestamps and at least one valid monitoring stream for the relevant outcome. At USZ, eligibility additionally requires documented written general consent for research use of routinely collected health-related data.

Cases will be excluded from the adult primary analysis if they involve patients aged 18 years or younger, occur outside the prespecified phase windows or during technical deployment intervals, lack linkable case meta-data, contain irreconcilable or invalid timestamps, have no valid monitoring data or, at USZ, lack documented written general consent. For outcome-specific analyses, cases with no valid measurement time for a given vital sign or alarm outcome will be excluded from that specific outcome only. Paediatric cases will be analysed and reported separately because paediatric physiological ranges, alarm thresholds and monitor configurations differ substantially from those in adult care.^{37 38}

Consent

At NYP/WCM, individual study-specific consent was not obtained because the Institutional Review Board granted a waiver of informed consent for this observational analysis of routinely collected clinical monitoring and alarm data. At USZ, only cases from patients with documented written general consent for research use

of routinely collected health-related data were included. The Cantonal Ethics Committee Zurich issued a declaration of non-jurisdiction for the analysis of anonymised routinely collected data, and the USZ Data Governance Board approved the anonymisation and export workflow.

Intervention: VPA implementation

The intervention is the introduction of VPA as a visual situation-awareness system via a CE- and FDA-certified software update to Philips IntelliVue patient monitors used in the participating centres (including MX450, MX550 and MX750 monitors). VPA is displayed alongside conventional numerical values and physiological waveforms in a split-screen view (figure 2). The avatar visualisation is enabled by default after implementation, and no routine non-VPA monitor layout is available in participating areas.

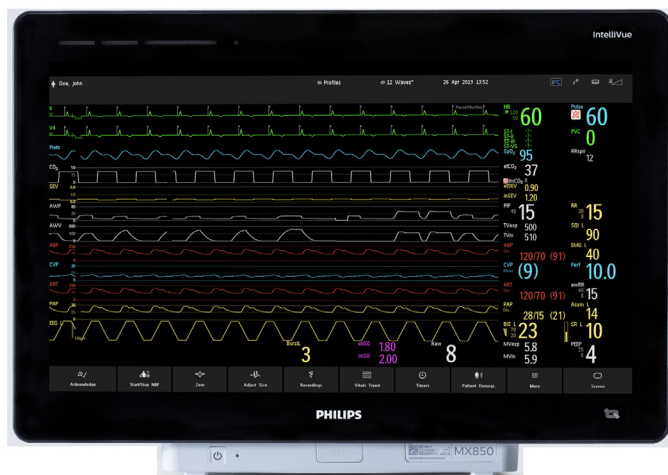
Clinicians can choose between predefined screen layouts that differ in the arrangement and relative prominence of the avatar and conventional monitoring elements. They can also select predefined care profiles tailored to clinical contexts such as adult or paediatric care. Profile selection adjusts both conventional and VPA thresholds. After profile selection, conventional alarm limits remain adjustable by clinicians according to clinical need, while VPA thresholds are fixed within each care profile to reflect the centre-specific standard configuration. Centre-specific alarm and VPA threshold settings are provided in online supplemental material 1.

Implementation versus utilisation

The exposure definition is the implementation phase rather than individually measured utilisation. VPA is enabled by default, and no routine non-VPA layout is available, but the monitors do not routinely retain research-accessible, time-stamped logs of selected screen layouts or display-state changes. Consequently, the datasets cannot identify which VPA-containing layout was active on a specific monitor at a given time, its relative prominence, or clinicians' gaze and attention. The

A

Pre-Implementation:



B

Post-Implementation:

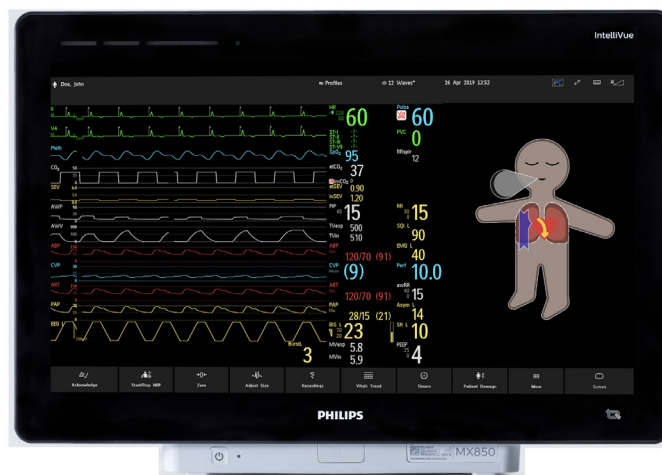


Figure 2 Conventional patient monitoring before and after implementation of the Visual Patient Avatar (VPA). (A) Conventional patient monitoring display before VPA implementation. (B) Identical monitoring configuration after VPA implementation, with avatar-based visualisation displayed alongside conventional numerical values and waveforms.

primary analysis, therefore, estimates the association between institutional VPA implementation and outcomes under real-world conditions, analogous to an intention-to-treat effect. Limited attention to the avatar, despite its availability, could lead to exposure misclassification and dilute observed associations.

Alarm and threshold stability

Centre-specific default alarm settings and VPA visualisation thresholds were configured before implementation in accordance with local standards of care and remained stable during the prespecified study windows. User-level alarm-limit changes during individual cases are part of routine clinical care and are reflected in alarm logs, but they are not uniformly available as analysable covariates across both centres. Vital-sign deviation outcomes are calculated using fixed, prespecified thresholds (table 2), whereas audible alarm outcomes are calculated from recorded alarm events generated by the clinical monitor configuration in use at the time of care.

Implementation process and training

Implementation was preceded by joint planning sessions involving clinical leadership, patient-safety and quality-management representatives, study authors and Philips representatives to adapt deployment to local workflows. The go-live phase included structured training in VPA and the concept of situation awareness through lectures, bedside teaching and walk-in training stations. Local change champions, who were practising anaesthetists at each centre, remained available after go-live for institutional teaching and bedside support.

Study outcomes

Primary outcome

Vital-sign deviation time is defined as the cumulative duration during which a given vital parameter exceeded prespecified upper or lower thresholds. To allow comparability across cases and monitoring durations, deviation time will be normalised to 60 min of valid measurement time for the respective vital sign: $60 \times (\text{deviation time} / \text{measurement time})$. Measurement time refers to the time during which valid measurements of the respective vital sign are available. Thresholds are centre-specific and reflect local configurations (table 2). Outcomes will be calculated separately for each vital sign displayed in VPA.

Secondary outcomes

VitalSign area under the curve (AUC) is a time-based AUC metric for deviations outside the admissible interval. For each minute i , the recorded value v_i is compared with the predefined interval $[L, U]$, and the absolute deviation d_i outside that interval is calculated. VitalSign AUC is defined as the sum of $d_i \times 1 \text{ min}$ across the case, capturing both magnitude and duration of deviations.

Alarm count per 60 min is the total number of audible alarms, aggregated across analysable monitored vital signs and normalised to 60 min of measurement time. Alarm time per 60 min is the cumulative duration of audible alarms, similarly aggregated and normalised. Alarm reaction time is the mean time from alarm onset to the point at which the alarm is acknowledged or paused via user input, as recorded in alarm logs. If multiple alarms occurred simultaneously, the reaction time will be calculated separately for each alarm event.

Table 2 Adult threshold settings used at NewYork-Presbyterian/Weill Cornell Medicine and University Hospital Zurich

Vital sign	Low threshold	High threshold	Units
NewYork-Presbyterian/Weill Cornell			
Heart rate (ECG)	45	135	/min
Pulse	45	135	/min
Blood pressure (mean)	55	125	mm Hg
CVP (mean)	2	14	mm Hg
STE	None	None	mV
SpO ₂	94	100	%
Respiratory rate	10	25	/min
Tidal volume	200	1200	mL
etCO ₂	25	55	mm Hg
Body temperature	35.5	38	°C
BIS	<65	≥65	N/A
TOF ratio	<90	≥90	%
Cardiac index	4.5	8	L/min/m ²
Peak airway pressure	5	35	mbar
FiO ₂	21	80	%
University Hospital Zurich			
Heart rate (ECG)	55	115	/min
Pulse*	55	115	/min
Blood pressure (mean)	65	105	mm Hg
CVP (mean)	4	12	mm Hg
STE	<-0.2	>0.2	mV
SpO ₂	94	None	%
Respiratory rate	8	16	/min
Tidal volume†	250	750	mL
etCO ₂	4.2	5.5	kPa
Body temperature‡	36.1	37.5	°C
BIS	<55	none	–
TOF ratio*	<90	none	%
Peak airway pressure	10	30	mbar

Vital-sign thresholds are centre-specific and reflect local configurations used in routine clinical monitoring. These thresholds will be used to calculate vital-sign deviation time and VitalSign area under the curve (AUC). Vital signs with corresponding analysable audible alarms will also be used for alarm-related outcomes.

*Pulse and TOF ratio have no corresponding acoustical alarms and will therefore not be included in alarm-related analyses.

†No direct audible alarm is defined for tidal volume. Where minute ventilation alarm data and corresponding measurement-time denominators are available, these alarms will be analysed as a respiratory-volume-related alarm outcome; otherwise tidal volume will be restricted to vital-sign deviation and AUC outcomes.

‡Body temperature data with low temporal resolution will not be considered for alarm-related analyses.

BIS, bispectral index; CVP, central venous pressure; etCO₂, end-tidal carbon dioxide; FiO₂, fraction of inspired oxygen; SpO₂, peripheral oxygen saturation; STE, ST-segment elevation; TOF, train-of-four.

Not all vital signs have corresponding audible alarms or sufficient temporal resolution for alarm-related analyses. Details are provided in [table 2](#). Respiratory volume related alarms will be analysed only where corresponding alarm data and measurement-time denominators are available; otherwise, the relevant vital sign will be restricted to vital-sign deviation and AUC analyses.

Rationale for the chosen outcome measures

The primary outcome was selected because VPA is designed to improve clinicians' perceptions and comprehension of deviations in vital signs. If improved situation awareness translates into earlier corrective action, this should be reflected by less time outside clinically accepted ranges before it necessarily affects rare patient-centred

endpoints. Vital-sign deviation time is also objective, available from routine monitoring data and interpretable at the case level across varied care areas.

Audible alarms were included because alarm frequency, duration and reaction time reflect important aspects of clinical workload and alarm management. Because users can pause or adjust alarms while fixed vital-sign deviation thresholds remain available for analysis, evaluating both vital sign based and alarm-related outcomes provides a more complete assessment of implementation effects.

Data sources and data collection

At NYP/WCM, monitoring and alarm data are recorded through the institutional data warehouse infrastructure, providing high-resolution vital-sign and alarm data linked to the anaesthesia record. Data extraction and linkage are performed by the NYP/WCM study and data warehouse teams.

At USZ, vital signs are captured via CapsuleTech nodes at a 1-minute sampling rate, and case-level metadata are obtained from the institutional data lake (IMDSoft MetaVision). Audible alarm logs are downloaded from Philips PICiX central stations at a resolution of 1 Hz. Planned variables and source systems are summarised in [table 3](#).

Data management, security and confidentiality

Data management is conducted by trained and authorised personnel at each centre in accordance with local policies, data governance requirements and regulatory approvals. Access is restricted to approved study personnel, and all data transfers use secure institutional mechanisms. The datasets used for analysis are anonymised or de-identified in accordance with local regulatory requirements before transfer to statisticians.

Table 3 Planned data acquisition per case and data sources by centre

Data	Measurement	NYP/WCM	USZ
Demographic and case-context data			
Study phase	Categorical	DWC	Data lake
Case date/calendar period (sensitivity analyses)	Date/month	DWC	Data lake
Care area/location (adjustment/stratification)	Categorical	DWC	Data lake
End of anaesthesia	Time in seconds	DWC	Data lake
End of surgery	Time in seconds	DWC	Data lake
Sex/gender (adjustment)	Categorical	DWC	Data lake
Age (adjustment)	Continuous or categorical	DWC	Data lake
Body mass index (adjustment)	Continuous or categorical	DWC	Data lake
ASA classification (adjustment)	Categorical	DWC	Data lake
Surgical risk classification (adjustment)	Categorical; 1–6 scale with higher values indicating higher complexity/risk	DWC	Data lake
Urgency of surgery (adjustment)	Categorical	DWC	Data lake
Type of surgery/procedure	Categorical	DWC	Data lake
Surgical specialty	Categorical	DWC	Data lake
Type of anaesthesia (adjustment)	Categorical	DWC	Data lake
Type of airway device	Categorical	DWC	Data lake
Type of regional anaesthesia	Categorical	DWC	Data lake
Vital-sign data			
Vital-sign deviation	Time in minutes	DWC	Data lake/CapsuleTech
Vital-sign time-based AUC	AUC metric	DWC	Data lake/CapsuleTech
Alarm data			
Alarm count	Numeric	DWC	PICiX
Alarm time	Time in seconds	DWC	PICiX
Alarm reaction time	Time in seconds	DWC	PICiX
ASA, American Society of Anesthesiologists; AUC, area under the curve; DWC, Philips Data Warehouse Connect; NYP/WCM, NewYork-Presbyterian/Weill Cornell Medicine; PICiX, Philips central station; UHZ, University Hospital Zurich.			

Measures to ensure data quality

Data extraction relies on established routine clinical documentation and monitoring pipelines at both centres. Quality checks are performed during data export and before statistical analysis and include completeness checks, plausibility range checks, timestamp consistency checks and verification of successful linkage between case metadata, vital-sign streams and alarm logs. Data-cleaning decisions will be documented in analysis scripts.

Statistical analysis

The unit of analysis is the individual anaesthesia case. Analyses will be performed separately for each centre and for primary adult analyses, separately from paediatric analyses. The primary comparison is post implementation versus pre-implementation; the adaptation phase is excluded from primary analyses.

Case characteristics will be summarised by centre and phase using means with SD or medians with IQR for continuous variables and counts with percentages for categorical variables. Descriptive comparisons between phases may use t-tests or Wilcoxon tests for continuous variables and χ^2 or Fisher's exact tests for categorical variables.

For each outcome, multivariable regression models will estimate the association between the implementation phase and outcome while adjusting for prespecified covariates. Covariates were selected a priori using a conceptual framework in which patient characteristics, procedure characteristics, care context and calendar time may influence both monitoring outcomes and the distribution of cases across implementation phases. Planned adjustment variables include age, sex, body mass index, American Society of Anesthesiologists (ASA) physical status, urgency of surgery, surgical risk classification, surgical specialty or procedure type, type of anaesthesia, airway device, regional anaesthesia and care area, subject to availability and missingness.

Model choice will be guided by outcome distribution and diagnostics. Linear regression will be used as the default for normalised outcomes. If residual diagnostics suggest marked non-normality or heteroskedasticity, outcomes may be log-transformed with a small constant where needed, or analysed using generalised linear models, such as Gamma regression with a log link for right-skewed positive outcomes. If an outcome has a large point mass at zero, two-part or zero-inflated models may be considered. Heteroskedasticity-robust standard errors will be used where appropriate.

Clustering by provider, operating room, unit or time period was considered. Provider-level identifiers are not available in the anonymised analysis datasets and would not fully represent team-based anaesthesia care; therefore, provider-level mixed-effects models or provider-level generalised estimating equations are not planned. Care-area identifiers will be used for adjustment or stratification, and location-level fixed effects will be explored where sufficiently complete identifiers are available.

Residual correlation within providers, rooms or time periods remains a limitation.

To address multiple testing, the Bonferroni correction will be applied within each outcome family and centre. The five outcome families are vital-sign deviation time, VitalSign AUC, alarm count, alarm time and alarm reaction time. Within a family, each analysable vital sign contributes one primary phase-effect p value. The maximum number of comparisons is 15 vital signs at NYP/WCM and 13 at USZ for vital-sign deviation and AUC outcomes; alarm-related families will include fewer comparisons because not all vital signs have analysable audible alarms.

Missingness will be summarised by centre, phase, outcome and covariate. We expect missing covariate values to arise mainly from routine documentation or linkage errors. If fewer than 5% of values are missing for a covariate, complete-case analysis for that covariate will be used; if 5%–40% are missing and the missingness pattern is compatible with missing completely at random or missing at random based on observed data, multiple imputation by chained equations will be used, and estimates will be pooled using Rubin's rules.³⁹ Covariates with more than 40% missingness will not be included in primary adjusted models but may be described or used in sensitivity analyses if clinically important.

Prespecified sensitivity analyses will include stratification and interaction analyses by care area, surgical urgency, ASA physical status, surgical risk classification, type of anaesthesia and early versus late post implementation periods. Additional sensitivity analyses may include alternative outcome distributions, exclusion of cases with substantial missingness, adjustment for calendar month or season and location-level fixed effects where data permit. Sensitivity analyses are intended to assess robustness and heterogeneity of the implementation association, not to establish causality.

Sample size considerations

This study leverages routinely collected data from large perioperative case volumes during prespecified implementation windows. Formal power calculations are not performed because the sample size is determined by clinical caseloads and deployment timing. Approximately 16000 cases at NYP/WCM and 42000 cases at USZ are expected, supporting precise estimation of associations and stratified analyses by centre and care area.

ETHICS AND DISSEMINATION

Ethics and data governance approvals

The study was approved by the Institutional Review Board of Weill Cornell Medicine (IRB No. 22-05024803), which granted a waiver of informed consent. At USZ, only cases from patients with documented written general consent for research use of routinely collected health-related data were included. The Cantonal Ethics Committee Zurich issued a declaration of non-jurisdiction in accordance with

Swiss regulations for the analysis of anonymised routinely collected data (BASEC Req-2021-00756; 6 July 2021). At USZ, the institutional Data Governance Board approved the project, including the data export and anonymisation workflow (DGR0000504; 31 January 2025). A collaboration and data-sharing agreement between Weill Cornell Medicine and the University of Zurich was established on 11 April 2023.

Data deposition and curation

Data supporting the findings of this study are derived from routinely collected perioperative monitoring and alarm data recorded at the participating institutions during the prespecified study periods. The datasets are not publicly available due to data protection regulations and institutional restrictions. Anonymised data and statistical code may be made available from the corresponding author on reasonable request, subject to approval by the participating institutions and relevant data governance bodies.

Analytical independence and bias mitigation

Several authors have financial or intellectual interests related to VPA, which are fully disclosed. To reduce investigator bias, outcomes, thresholds and primary analyses were prespecified before comparative outcome analyses; statistical modelling will be implemented by statistics coauthors who were not involved in the invention or licensing of VPA; data extraction and quality-control steps will be documented; and Philips will have no role in data analysis, interpretation, manuscript preparation or publication decisions.

Dissemination

Study findings will be submitted to peer-reviewed scientific journals and presented at national and international scientific meetings. A primary joint publication including data from both centres is anticipated. Secondary analyses using centre-specific datasets may be published separately after completion of the primary analysis and in accordance with institutional policies and the collaboration agreement.

DISCUSSION

This protocol describes a pragmatic dual-centre evaluation of an avatar-based situation-awareness display introduced into routine perioperative care. By focusing on objective monitoring metrics, the study will test whether human-factors benefits observed in controlled settings translate into measurable changes in vital-sign deviations and audible alarm burden in real-world practice.

The study's main strength is the use of large volumes of routinely collected data from two tertiary academic centres with different healthcare systems, workflows and alarm configurations. This heterogeneity improves external validity but also means that absolute alarm rates and threshold-based outcomes should not be directly

pooled or interpreted as interchangeable between centres. The primary interpretation will therefore focus on centre-specific estimates and consistency of direction across settings.

The most important limitation is the uncontrolled pre-post design. Temporal changes unrelated to VPA, including staffing changes, trainee rotations, workflow modifications, institutional quality-improvement activities, seasonal case-mix changes and secular trends in perioperative practice, may influence outcomes. Adjustment for observed patient and procedure characteristics, along with sensitivity analyses, can reduce but not eliminate these concerns. The study will therefore report associations rather than causal effects.

Implementation and utilisation are also distinct. VPA is enabled by default in participating areas after implementation, but monitors do not routinely retain research-accessible, time-stamped logs of screen-layout or display-state changes. The datasets, therefore, cannot establish which VPA-containing layout was active on a particular monitor at a given time, its relative prominence, or the actual clinician attention it received. This may introduce exposure misclassification. Because such misclassification would likely reduce differences between phases, null findings should not be interpreted as definitive evidence that VPA was not used or could not be effective under more controlled utilisation conditions.

The analysis further relies on routinely collected monitoring and alarm data that were not created primarily for research. Data quality, missingness, changes in documentation practices and incomplete capture of provider-, room- or team-level clustering may affect precision and internal validity. These limitations will be addressed through prespecified data-quality checks, transparent reporting of missingness and cautious interpretation.

Finally, the chosen outcomes are process-level metrics rather than direct patient-centred clinical endpoints. This is appropriate for an implementation study of a monitoring visualisation whose first plausible effects should occur through perception, comprehension, timely correction of deviations and alarm management. If favourable associations are observed, subsequent studies should evaluate whether changes in these process metrics translate into improved patient-centred outcomes.

Author affiliations

- ¹Department of Anaesthesiology, Weill Cornell Medicine, New York, New York, USA
- ²Institute of Anaesthesiology and Perioperative Medicine, University Hospital Zurich, Zurich, Switzerland
- ³University of Zurich, Zurich, Switzerland
- ⁴Seminar for Statistics, ETH Zurich, Zurich, Switzerland
- ⁵Mathematics, ETH Zurich, Zurich, Switzerland
- ⁶Department of Anaesthesiology, University Medical Centre Groningen, Groningen, The Netherlands
- ⁷Technical University of Munich, Munich, Germany

Contributors SYJ: conceptualisation, methodology, writing - original draft, writing - review and editing, supervision, project administration, funding acquisition. TRR: conceptualisation, methodology, study preparation, writing - original draft, writing - review and editing, supervision, project administration. VET: statistical

methodology, software, writing - review and editing. ME: methodology, study preparation, writing - review and editing. MK: statistical methodology, software, writing - review and editing. GD: statistical methodology, software, writing - review and editing. KC: statistical methodology, software, writing - review and editing. LM: statistical methodology, software, writing - review and editing. MT: study preparation, methodology, writing - review and editing. AM: study preparation, methodology, writing - review and editing. CBN: conceptualisation, study preparation, methodology, writing - review and editing. ZAT: conceptualisation, methodology, writing - original draft, writing - review and editing, supervision, project administration, funding acquisition. DWT: conceptualisation, methodology, writing - original draft, writing - review and editing, supervision, project administration, funding acquisition. DWT is the guarantor. Generative AI (ChatGPT model 5.2, OpenAI, San Francisco, CA, USA) was used to enhance the clarity and readability of the manuscript. All AI-assisted revisions were reviewed and approved by the authors to ensure accuracy and completeness.

Funding TRR and DWT have received research funding from Philips through a joint development agreement between Philips Medizin Böblingen GmbH and the University of Zurich for the Visual Patient Avatar (dated 18 December 2018). In addition, TRR and DWT received research grants from Philips Healthcare, Philips North America (Exhibit B: USZ-0007; Exhibit B: USZ-0010). SYJ, VET and ZAT received research funding from Philips Healthcare, Philips North America (SRA_233936/WCM-0001). All funding and remuneration were managed through institutional agreements. This is an investigator-initiated study. Philips had no role in the study design, data collection, data analysis, data interpretation, manuscript preparation or decisions regarding publication.

Competing interests CBN and DWT are named inventors of the Visual Patient Avatar technology. This technology has been licensed by the University of Zurich to Philips. In accordance with University of Zurich regulations, inventors are entitled to an appropriate share of licensing revenues, and CBN and DWT receive royalties related to this technology. TRR, CBN and DWT have received honoraria and travel support for lectures from Philips. TRR and DWT have received research funding from Philips through a joint development agreement between Philips Medizin Böblingen GmbH and the University of Zurich for the Visual Patient Avatar (dated 18 December 2018) and research grants from Philips Healthcare, Philips North America (Exhibit B: USZ-0007; Exhibit B: USZ-0010). SYJ, VET and ZAT have received research funding from Philips Healthcare, Philips North America (SRA_233936/WCM-0001). KC, GD, LM, ME, AM, MT and MK have no competing interests to declare.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <https://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Lukas Meier <https://orcid.org/0000-0001-8508-1222>

David W Tscholl <https://orcid.org/0000-0002-3677-3705>

REFERENCES

- 1 Gelb AW, Morris WW, Johnson W, et al. World Health Organization-World Federation of Societies of Anaesthesiologists (WHO-WFSA) international standards for a safe practice of anesthesia. *Anesth Analg* 2018;126:2047–55.
- 2 Klein AA, Meek T, Allcock E, et al. Recommendations for standards of monitoring during anaesthesia and recovery 2021: guideline from the Association of Anaesthetists. *Anaesthesia* 2021;76:1212–23.
- 3 American Society of Anesthesiologists. *Standards for basic anesthetic monitoring*. Schaumburg, IL, Available: <https://www.asahq.org/standards-and-practice-parameters/standards-for-basic-anesthetic-monitoring>
- 4 Schulz CM, Endsley MR, Kochs EF, et al. Situation awareness in anesthesia: concept and research. *Anesthesiology* 2013;118:729–42.
- 5 Drews FA, Westenskow DR. The right picture is worth a thousand numbers: data displays in anesthesia. *Hum Factors* 2006;48:59–71.
- 6 Tscholl DW, Handschin L, Rössler J, et al. It's not you, it's the design - common problems with patient monitoring reported by anesthesiologists: a mixed qualitative and quantitative study. *BMC Anesthesiol* 2019;19:87.
- 7 Ford S, Birmingham E, King A, et al. At-a-glance monitoring: covert observations of anesthesiologists in the operating room. *Anesth Analg* 2010;111:653–8.
- 8 Tscholl DW, Rössler J, Said S, et al. Situation awareness-oriented patient monitoring with visual patient technology: a qualitative review of the primary research. *Sensors (Basel)* 2020;20:2112.
- 9 Garot O, Rössler J, Pfarr J, et al. Avatar-based versus conventional vital sign display in a central monitor for monitoring multiple patients: a multicenter computer-based laboratory study. *BMC Med Inform Decis Mak* 2020;20:26.
- 10 Pfarr J, Ganter MT, Spahn DR, et al. Effects of a standardized distraction on caregivers' perceptive performance with avatar-based and conventional patient monitoring: a multicenter comparative study. *J Clin Monit Comput* 2020;34:1369–78.
- 11 Tscholl DW, Handschin L, Neubauer P, et al. Using an animated patient avatar to improve perception of vital sign information by anaesthesia professionals. *Br J Anaesth* 2018;121:662–71.
- 12 Roche TR, Said S, Braun J, et al. Avatar-based patient monitoring in critical anaesthesia events: a randomised high-fidelity simulation study. *Br J Anaesth* 2021;126:1046–54.
- 13 Sendelbach S, Funk M. Alarm fatigue: a patient safety concern. *AACN Adv Crit Care* 2013;24:378–86.
- 14 Lewandowska K, Weisbrodt M, Cieloszyk A, et al. Impact of alarm fatigue on the work of nurses in an intensive care environment—a systematic review. *Int J Environ Res Public Health* 2020;17:8409.
- 15 Koomen E, Webster CS, Konrad D, et al. Reducing medical device alarms by an order of magnitude: a human factors approach. *Anaesth Intensive Care* 2021;49:52–61.
- 16 Wesselink EM, Kappen TH, Torn HM, et al. Intraoperative hypotension and the risk of postoperative adverse outcomes: a systematic review. *Br J Anaesth* 2018;121:706–21.
- 17 Kouz K, Thiele R, Michard F, et al. Haemodynamic monitoring during noncardiac surgery: past, present, and future. *J Clin Monit Comput* 2024;38:565–80.
- 18 Rauch S, Miller C, Bräuer A, et al. Perioperative hypothermia—a narrative review. *Int J Environ Res Public Health* 2021;18:8749.
- 19 Koninklijke Philips NV. Visual patient avatar—a new way to see your patient's vitals takes shape, 2025. Available: <https://www.usa.philips.com/healthcare/technology/visual-patient-avatar>
- 20 Gasciauskaitė G, Lunkiewicz J, Roche TR, et al. Human-centered visualization technologies for patient monitoring are the future: a narrative review. *Crit Care* 2023;27:254.
- 21 Milovanovic P, Braun J, Hunn CA, et al. Avatar-based versus conventional patient monitoring with distant vision: a computer-based simulation study. *J Clin Monit Comput* 2025;39:1065–75.
- 22 Endsley MR, Jones DG. *Designing for situation awareness: an approach to user-centered design*. 3rd edn. Boca Raton, FL: CRC Press, 2025.
- 23 Tramontana A, Rulli M, Falegnami A, et al. Visual avatar to increase situational awareness in anaesthesia: systematic review of recent evidence. *World J Methodol* 2025;15:100459.
- 24 Andrade E, Quinlan L, Harte R, et al. Novel interface designs for patient monitoring applications in critical care medicine: human factors review. *JMIR Hum Factors* 2020;7:e15052.
- 25 Drews FA, Doig A. Evaluation of a configural vital signs display for intensive care unit nurses. *Hum Factors* 2014;56:569–80.
- 26 Wachter SB, Markewitz B, Rose R, et al. Evaluation of a pulmonary graphical display in the medical intensive care unit: an observational study. *J Biomed Inform* 2005;38:239–43.
- 27 Michels P, Gravenstein D, Westenskow DR. An integrated graphic data display improves detection and identification of critical events during anesthesia. *J Clin Monit* 1997;13:249–59.
- 28 Tscholl DW, Weiss M, Handschin L, et al. User perceptions of avatar-based patient monitoring: a mixed qualitative and quantitative study. *BMC Anesthesiol* 2018;18:188.

- 29 Pfarr J, Ganter MT, Spahn DR, *et al.* Avatar-based patient monitoring with peripheral vision: a multicenter comparative eye-tracking study. *J Med Internet Res* 2019;21:e13041.
- 30 Tscholl DW, Rössler J, Handschin L, *et al.* The mechanisms responsible for improved information transfer in avatar-based patient monitoring: multicenter comparative eye-tracking study. *J Med Internet Res* 2020;22:e15070.
- 31 Wetli DJ, Bergauer L, Nöthiger CB, *et al.* Improving visual-patient-avatar design prior to its clinical release: a mixed qualitative and quantitative study. *Diagnostics (Basel)* 2022;12:555.
- 32 Lunkiewicz J, Fries D, Milovanovic P, *et al.* Pediatric anesthesia providers' perspective on the real-life implementation of the Philips Visual Patient Avatar: a qualitative study. *Children (Basel)* 2023;10:1841.
- 33 Hunn CA, Lunkiewicz J, Noethiger CB, *et al.* Qualitative exploration of anesthesia providers' perceptions regarding Philips Visual Patient Avatar in clinical practice. *Bioengineering (Basel)* 2024;11:323.
- 34 Radaviciute I, Hunn CA, Lunkiewicz J, *et al.* Survey-based qualitative exploration of user perspectives on the philips visual patient avatar in clinical situation management. *Sci Rep* 2024;14:22176.
- 35 Des Jarlais DC, Lyles C, Crepaz N, *et al.* Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Public Health* 2004;94:361–6.
- 36 Benchimol EI, Smeeth L, Guttman A, *et al.* The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med* 2015;12:e1001885.
- 37 van Zadelhoff AC, Poppe JA, Willemsen S, *et al.* Age-dependent changes in arterial blood pressure in neonates during the first week of life: reference values and development of a model. *Br J Anaesth* 2023;130:585–94.
- 38 de Graaff JC, Pasma W, van Buuren S, *et al.* Reference values for noninvasive blood pressure in children during anesthesia: a multicentered retrospective observational cohort study. *Anesthesiology* 2016;125:904–13.
- 39 Jakobsen JC, Gluud C, Wetterslev J, *et al.* When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. *BMC Med Res Methodol* 2017;17:162.