PEER-REVIEWED · OPEN ACCESS
Sriwijaya Journal of Forensic and Medicolegal
Vol. 3, No. 2 (2025), pp. 57–72 · e-ISSN 2987-1530 · Article ID: 249
journal homepage: phlox.or.id/index.php/sjfm

Original Research Article

Comprehension as the Gatekeeper of Valid Surgical Consent: A Cross-Sectional Medicolegal Analysis of 320 High-Risk Indonesian Surgical Patients

Mustafa Mahmud[1]*, Cinthya Callathea[2], Yi-Fen Huang[3], Delia Tamim[4]

[1]*Department of Thoracic Surgery, CMHC Research Center, Palembang, Indonesia*
[2]*Department of Obstetrics and Gynecology, CMHC Research Center, Palembang, Indonesia*
[3]*Department of Biomolecular and Translational Medicine, Thimphu Medical Institute, Thimphu, Bhutan*
[4]*Department of Administrative Law, Enigma Institute, Palembang, Indonesia*

ARTICLE INFO

Article history
Received: September 25, 2025
Revised: October 20, 2025
Accepted: November 19, 2025
Available online: November 28, 2025
Published: November 30, 2025

Keywords
Forensic medicolegal assessment
High-risk surgery
Informed consent comprehension
Medical negligence
Patient autonomy

Corresponding author
Mustafa Mahmud
E-mail:
mustafa.mahmud@sriwijayasurgery.com
ORCID: 0009-0003-7186-3284

DOI
https://doi.org/10.59345/sjfm.v3i2.249

A B S T R A C T

Background. Documented consent without demonstrable comprehension is medicolegally insufficient, yet comprehension-based informed consent (IC) failure in collectivistic surgical populations is largely unquantified.

Methods. This study was a cross-sectional assessed IC efficacy among 320 consecutive adults undergoing high-risk elective surgery (ASA physical status III–IV) at Private Hospital X, Palembang, Indonesia, in 2024, reported per STROBE. Six instruments standardised to 0–100 were administered, including the IC Comprehension Score (ICCS) and the five-domain composite IC Efficacy Score (ICE); adequacy was ICE ≥ 70. Exact intervals and permutation tests, effect sizes, areas under the receiver operating characteristic curve recovered from Mann–Whitney U statistics, E-values, fragility indices, standardisation and a partial-identification audit were added.

Results. Only 19 of 320 patients (5.9%, 95% CI 3.6–9.1) achieved adequate IC efficacy, below both published comparators reporting an explicit adequacy proportion (exact binomial P < 0.001 versus 14.0% and versus 35.8%); standardisation to lower-education distributions reduced this to 2.74–3.79%, so 5.9% is an upper bound. Adequacy rose monotonically across educational strata from 0.0% to 13.3% (exact Cochran–Armitage z = 3.66, P = 0.0003; tertiary versus secondary education or below, odds ratio 8.27, 95% CI 2.36–29.00; E-value 14.29; fragility index 8). ICCS was the only independent predictor in an exploratory multivariable model with 2.71 events per variable (odds ratio 13.75 per SD, 95% CI 3.32–56.92), an association partly constitutive because comprehension is one ICE domain. Neither family influence nor collectivism predicted adequacy.

Conclusion. Comprehension, not cultural collectivism, gates valid surgical consent.

How to cite. Mahmud M, Callathea C, Huang YF, et al. Comprehension as the gatekeeper of valid surgical consent: a cross-sectional medicolegal analysis of 320 high-risk Indonesian surgical patients. Sriwijaya J Forensic Medicoleg. 2025;3(2):57–72. doi:10.59345/sjfm.v3i2.249

1. Introduction

Informed consent (IC) is simultaneously the ethical cornerstone of the surgeon–patient relationship and a discrete medicolegal artefact whose adequacy is adjudicated after the fact. Forensic and legal-medicine doctrine in an increasing number of jurisdictions holds that a signed form unaccompanied by evidence of patient understanding does not discharge the practitioner's duty, and appellate decisions reviewed in the forensic literature turn repeatedly on whether comprehension was verified rather than on whether a signature was obtained.[1,2] In Indonesia, Law No. 17 of 2023 on Health preserves the statutory requirement of comprehensive pre-surgical disclosure, while the Civil Code and the Ministry of Health regulation on medical consent supply the framework within which that requirement is tested; Indonesian legal scholarship has repeatedly identified the absence of comprehension verification as the pivotal vulnerability in hospital consent practice.[3–5] Forensic medicine therefore has a direct stake in an empirical question usually treated as a matter of clinical quality: how often, and for whom, does the consent process actually produce comprehension?

The international evidence is consistently discouraging. A systematic review of empirical studies of consent comprehension found understanding of the basic components to be limited across specialties and settings.[6] Among vascular surgery patients in a high-income setting, only 14% answered every comprehension item correctly, with frailty the sole independent predictor of being informed.[7] Direct observation of 90 audio-recorded preoperative conversations before major cardiothoracic, vascular, oncological and neurosurgical procedures showed that surgeons routinely described the illness, the operation and its complications but were least likely of all to assess whether the patient had understood.[8] Cross-sectional audits in Nepal, Ethiopia, India, Italy and Tanzania reach the same conclusion from very different health systems: patients sign, and frequently report satisfaction, while remaining unable to identify the alternatives, the benefits or the operator.[9–16] Documentation audits and trainee surveys show the same deficit from the clinician side, with consent frequently delegated to staff who have received no formal training in the task.[17–20]

The response has been largely technological and editorial rather than structural. Randomised trials of digital, multimedia and video-assisted consent improve information gain, recall and satisfaction across bariatric, urological, orthopaedic and mixed surgical populations,[21–26] a teach-back protocol improves both consent quality and surgeon trust at negligible time cost,[27] and scoping reviews map the rapid digitalisation of the consent process.[28] In parallel, repeated analyses show that the consent documents themselves are written far above the reading level of the populations they address, with measurable consequences for comprehension and even for refusal rates.[29–32] Health literacy is the common thread: it is associated with surgical outcomes, with participation in shared decision-making, and with multimorbidity at population level.[33–36]

These deficits are amplified in low- and middle-income and in collectivistic settings, where health-literacy gradients are steeper and decision-making is family-centred rather than individual-centred.[37,38] A parallel bioethical literature argues that the individualism–collectivism binary is too coarse to describe patient autonomy in East and Southeast Asian societies, and that relational autonomy better characterises how consent is actually exercised there.[39–43] What that literature lacks is a quantitative test. We therefore treat the proposition as a hypothesis rather than a premise: if collectivistic orientation genuinely threatens consent validity, it should predict consent failure independently of comprehension, and if it does not, the threat has been misattributed. To our knowledge no study has estimated the prevalence of comprehension-based consent failure in a high-risk Indonesian surgical population while measuring family influence and cultural collectivism as competing determinants.

Private Hospital X is a Type B private hospital in Palembang, South Sumatra, accredited by the Indonesian National Accreditation Commission (KARS) and performing more than 1,200 major elective procedures annually across cardiac, thoracic, abdominal and orthopaedic specialties. It operates within the constraints that characterise Indonesian tertiary surgical care generally: a high-volume elective list, a mixed national-insurance and private payer environment, and a workforce for which structured consent-communication training is not a formal component of surgical residency.[44] Its patient population spans the full national educational range, from no formal schooling to postgraduate qualification, which makes it an informative setting in which to test whether a standard consent protocol functions across that range.

This study therefore aimed to (1) estimate the prevalence of adequate comprehension-based IC efficacy among high-risk elective surgical patients at this centre, with exact interval precision, direct standardisation and explicit international benchmarking; (2) quantify the contribution of patient comprehension, health literacy, surgeon communication quality, family influence, cultural collectivism and educational attainment to that outcome, reporting an effect size and a 95% confidence interval for every estimate; and (3) subject the resulting multivariable model, which is exploratory at the available event count, to a formal identification audit, so that the medicolegal inferences drawn from it rest only on quantities the data can support.

2. Methods

2.1. Study design and reporting

A cross-sectional analytical study was conducted and is reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement for cross-sectional studies.[46] The study was conceived as a prevalence-and-determinants investigation with a pre-specified binary primary outcome; no interventional component was involved.

2.2. Setting and period

The study was performed at the Department of Surgery, Private Hospital X, Palembang, South Sumatra, Indonesia, from January 2024 to December 2024. Private Hospital X is a KARS-accredited Type B private hospital providing comprehensive tertiary surgical services and performing more than 1,200 major elective procedures annually across cardiac, thoracic, abdominal and orthopaedic specialties.

2.3. Participants

The study population comprised all adult patients (aged ≥ 18 years) scheduled for high-risk elective surgery, operationally defined as a procedure carrying American Society of Anesthesiologists (ASA) physical status III or IV classification together with a perioperative mortality risk exceeding 1% by the Revised Cardiac Risk Index or by institutional protocol. Eligible categories were (a) cardiac procedures (coronary artery bypass grafting, valve replacement or repair, and combined procedures); (b) thoracic procedures (lobectomy, pneumonectomy, thoracotomy with lung resection, pleurectomy); (c) major abdominal procedures (hepatectomy, pancreatectomy, colectomy with primary anastomosis, Whipple procedure); and (d) orthopaedic procedures (total hip arthroplasty, total knee arthroplasty, major spinal instrumentation).

Patients were excluded if they had documented cognitive impairment (Mini-Mental State Examination score < 24), were unable to communicate in Bahasa Indonesia or a local dialect, were undergoing emergency (non-elective) surgery, or declined written study participation consent. Consecutive sampling was used until the enrolment target was reached. The complete eligibility and analytical pathway, including the three analytical layers and the explicit refusal set, is summarised in Figure 1.

The screening log recording the number of patients assessed for eligibility and excluded under each criterion was not retained, and those counts cannot be reported. Two exclusion criteria have a predictable directional effect on the primary estimate: excluding patients with a Mini-Mental State Examination score below 24, and excluding those unable to communicate in Bahasa Indonesia or a local dialect, both remove groups at elevated risk of comprehension failure and therefore bias the observed adequacy proportion upward.

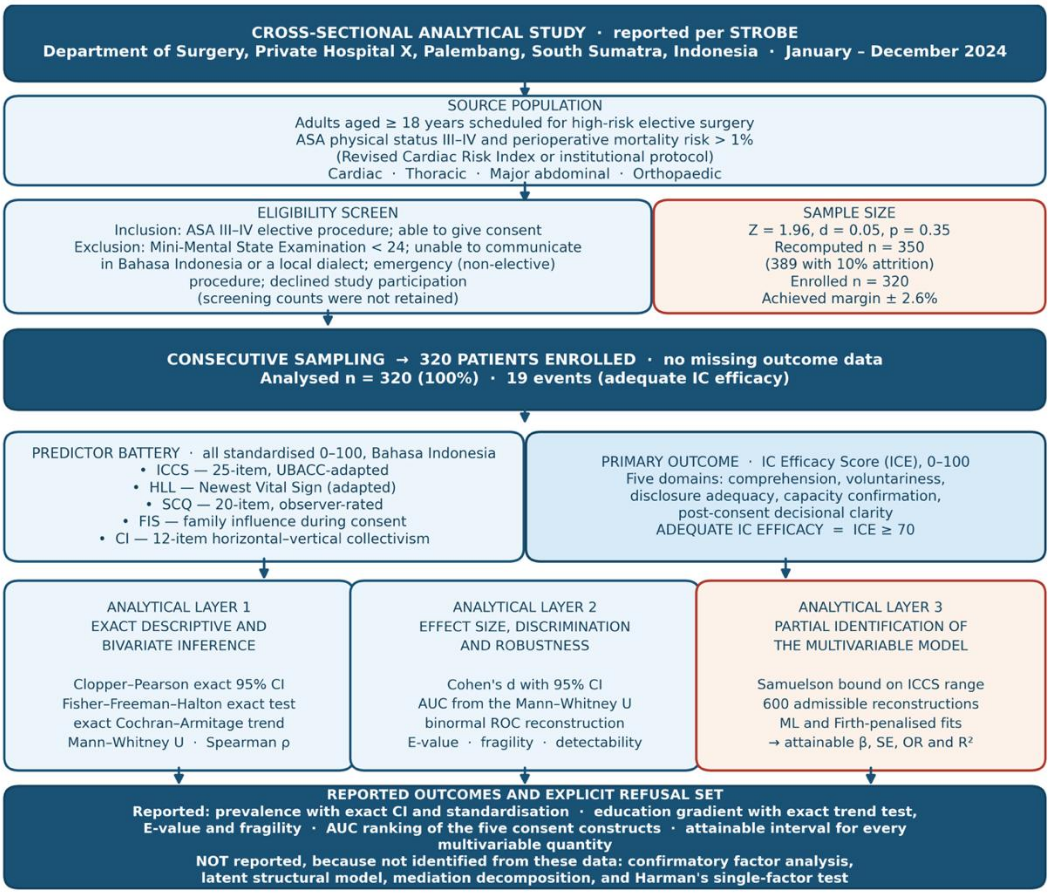

Figure 1. Study design, eligibility pathway and three-layer analytical strategy. The study enrolled 320 consecutive adults undergoing ASA physical status III–IV elective surgery at Private Hospital X, Palembang, between January and December 2024. Layer 1 applies exact rather than asymptotic inference throughout; Layer 2 adds a standardised effect size and a discrimination statistic to every comparison; Layer 3 audits which reported quantities the aggregate data can identify. The analyses explicitly refused as unidentified are listed in the final panel.

2.4. Sample size

Sample size was originally derived from the standard formula for a cross-sectional prevalence study (Z = 1.96, assumed IC adequacy prevalence 35%, absolute margin of error 5%), and a target of 320 participants was set. In the present re-analysis this calculation was recomputed: the stated inputs yield n = 350, or 389 after a 10% attrition allowance, so the realised sample of 320 is smaller than the design specification implies. The discrepancy is, however, conservative in the direction that matters. Because the observed prevalence (5.9%) is far below the assumed 35%, the binomial variance is correspondingly smaller and the achieved absolute margin of error for the primary prevalence estimate is ± 2.6%, comfortably within the intended 5%. The sample is therefore adequate for the primary aim, and its limitations bear on the multivariable aim only, where 19 events across seven model parameters give 2.71 events per variable.

2.5. Variables and measurement instruments

Six composite scores, each standardised to a 0–100 metric and each administered in Bahasa Indonesia by trained data collectors, constituted the study variables. The IC Comprehension Score (ICCS) was derived from a 25-item instrument adapted from the University of California San Diego Brief Assessment of Capacity to Consent, and assessed understanding of diagnosis, procedure, treatment rationale, operative risks, benefits and alternatives.[47] The Health Literacy Level (HLL) was assessed with the Newest Vital Sign instrument adapted to the Indonesian health system, measuring reading comprehension and health numeracy.[48] Surgeon Communication Quality (SCQ) was a 20-item observer-rated scale completed by a trained research assistant during the consent consultation, covering clarity, completeness, pacing, lay-language use, responsiveness to questions and non-verbal quality. The Family Influence Score (FIS) quantified the degree to which family members participated in, redirected or substituted for the patient's autonomous decision-making during the consent process. The Collectivism Index (CI) used the 12-item horizontal–vertical collectivism subscale adapted for the Indonesian sociocultural context.[49]

The IC Efficacy Score (ICE) was the primary composite outcome, integrating five consent domains: patient comprehension, voluntariness, disclosure adequacy, capacity confirmation and post-consent decisional clarity. Adequate IC efficacy was defined a priori as ICE ≥ 70/100, consistent with international benchmark criteria for the medicolegal adequacy of surgical consent.

Three reporting limitations of the instrument battery must be stated. First, no internal-consistency coefficient is reported for any of the six scales, because only composite scores were retained and the item-level responses required to compute Cronbach's α or McDonald's ω are unavailable to this analysis. This is a data-availability limitation rather than a property of the instruments, and its consequence is that unquantified measurement error attenuates every association reported here, including the null associations, where attenuation could in principle produce a spurious null. Second, the SCQ and FIS scales were constructed for this study and have no published validation; SCQ was rated by a single observer present during the consultation, so neither inter-rater reliability nor the possibility of an observation effect on surgeon behaviour can be assessed. Third, and most consequentially, the weighting of the five domains within the ICE composite was not documented at the time of construction and cannot be recovered; the implications for the ICCS–ICE association are set out in the Discussion.

2.6. Statistical analysis

All analyses were performed in Python 3.11.2 with NumPy 1.26 and SciPy 1.11, using custom implementations of the algorithms described below; all Monte-Carlo procedures used a fixed random seed (20260812) and the analysis script is available from the corresponding author. Two-sided α was 0.05, and exact P values are reported to three decimal places unless smaller than 0.001.

Descriptive statistics are frequencies with percentages for categorical variables and means with standard deviations for continuous variables; normality was assessed by the Kolmogorov–Smirnov test. Every proportion, including each stratum-specific adequacy rate, is accompanied by a Clopper–Pearson exact 95% confidence interval. Associations between categorical characteristics and IC efficacy adequacy were tested by the Fisher–Freeman–Halton exact test, implemented by Monte-Carlo permutation with 2×10^5 replicates, because expected cell counts below five occurred in four of the five contingency tables and invalidate the asymptotic Pearson χ^2 statistic; the Pearson χ^2 and Cramér's V are reported alongside for comparability. For the two ordered characteristics (education, monthly income) an exact Cochran–Armitage test for trend with equally spaced scores was computed by the same permutation scheme.

Continuous variables were compared between adequacy groups by the Mann–Whitney U test with a normal approximation. Two effect sizes accompany every such comparison. Cohen's d with a 95% confidence interval was computed from the group means and standard deviations using the pooled standard deviation. The area under the receiver operating characteristic curve (AUC) was recovered from the Mann–Whitney U statistic through the identity

$$AUC = 1 - U / (n_1 n_0)$$

where $n_1 = 19$ is the number of patients with adequate consent and $n_0 = 301$ the number without; the reported U counts pairs in which the inadequate-group value exceeds the adequate-group value, so AUC > 0.5 denotes higher scores among patients with adequate consent. A Hanley–McNeil standard error with a logit-transformed 95% confidence interval was used, and the corresponding curves were reconstructed under a binormal model with unit slope.[50,51] Spearman rank correlations between each continuous predictor and the continuous ICE score are reported with Fisher-z 95% confidence intervals.

Multivariable binary logistic regression was fitted by iteratively reweighted least squares, iterated to a maximum absolute coefficient change below 10^{-10}, with adequate IC efficacy as the dependent variable and ICCS, HLL, SCQ, FIS, CI, sex and education level as predictors, all continuous predictors z-standardised. Odds ratios with 95% Wald confidence intervals and the Nagelkerke pseudo-R^2 are reported. Because the outcome is rare and the strongest predictor produces near-separation, a Firth-penalised fit was computed in parallel as a bias-reduced sensitivity analysis.[52,53] The Hosmer–Lemeshow goodness-of-fit statistic is not reported, because with 19 events its decile groups contain too few events to be informative, and its reliability in exactly this situation has been questioned directly.[45] Variance inflation factors could not be computed, because that requires the predictor covariance matrix, which is unavailable from composite summaries; possible collinearity among ICCS, HLL and SCQ is therefore acknowledged but not quantified.

2.7. Partial-identification audit of the multivariable model

The analysis was conducted on the study's aggregate margins rather than on a record-level file. Rather than assume that this permits every reported quantity to be reproduced, an explicit identification audit was performed. First, Samuelson's inequality was applied to the adequate group's ICCS distribution: with $n_1 = 19$, a mean of 98.3 and a standard deviation of 2.1, every adequate patient's ICCS is bounded below by

$$\bar{x} - s\sqrt{(n_1 - 1)} = 98.3 - 2.1\sqrt{18} = 89.39$$

and the inadequate group's total sum of squared deviations caps at 39 the number of the 301 inadequate patients who could reach that value. Second, the full set of ICCS configurations consistent with the published margins was characterised: the two group means and standard deviations, the [0, 100] instrument range, and — critically — the reported Mann–Whitney U of 233, which fixes the number of discordant pairs and therefore the exact degree of overlap between the groups. Six hundred admissible configurations were generated under these constraints, and maximum-likelihood and Firth-penalised logistic regressions were fitted to each, yielding attainable intervals for β, its standard error, the odds ratio and the Nagelkerke pseudo-R^2.

Falling inside an attainable interval is a necessary but not a sufficient condition: it means a reported value is not refuted, not that it is confirmed. Falling outside the interval is decisive, because such a value cannot have arisen from the data as described. For the pseudo-R^2 a second and stronger argument is available that does not depend on reconstruction at all: because the univariable comprehension model is nested within the full model, the full model's Nagelkerke R^2 cannot be smaller than the univariable value, so any univariable value exceeding the reported full-model figure refutes it directly.

Three classes of analysis were considered and rejected as unidentified. Confirmatory factor analysis, latent structural modelling and Harman's single-factor test require item-level responses, which are not recoverable from composite scores. Mediation decomposition requires the inter-predictor correlation matrix, which is unobserved; positive-definiteness alone bounds the ICCS–HLL correlation only to [−0.085, 0.961] and the SCQ–FIS correlation only to [−0.951, 0.829], intervals far too wide to support any indirect-effect estimate. The general point that a latent correlation may be only partially identified when the underlying distributional assumptions cannot be verified is established in the psychometric literature.[54] These analyses are therefore not reported, and the mediational language used in earlier treatments of this dataset is replaced by explicitly conditional statements.

2.8. Robustness, standardisation and medicolegal sensitivity analyses

For the principal categorical finding, the education gradient, a series of robustness metrics was computed. The E-value quantifies the minimum strength of association, on the risk-ratio scale, that an unmeasured confounder would need with both education and consent adequacy in order to explain away the observed association.[55] The fragility index and the fragility quotient quantify how many outcome events would have to change to render the association non-significant.[56,57] For each non-significant categorical predictor, the observed power and the minimum detectable adequacy proportion at 80% power were computed, so that null results are characterised as informative or uninformative rather than simply reported as absent.

Two further analyses address external validity and measurement. The prevalence estimate was directly standardised to three alternative educational distributions using the stratum-specific adequacy rates, so that the sensitivity of the headline figure to the sample's educational skew could be quantified; the weighting distributions are stated explicitly with the results and are illustrative rather than census-derived. Common method variance was assessed qualitatively rather than by Harman's single-factor test, which requires item-level data: five of the six instruments are patient-reported and were administered in a single sitting, whereas SCQ is observer-rated and is therefore not subject to the same inflation, which permits an internal comparison of their associations with the outcome. Finally, the observed prevalence was compared by exact binomial test against the two published studies that report an explicit adequacy proportion for a comparable surgical population.[7,10]

2.9. Ethical approval

Ethical approval was granted by the Institutional Ethics Committee of the CMHC Research Center, under Reference No. 0024/2024. Written informed consent for research participation was obtained from all participants prior to enrolment. All data were anonymised using participant ID codes, and personal identifiers were stored separately under restricted access in accordance with Indonesian Law No. 27 of 2022 on Personal Data Protection. The study was conducted in adherence with the Declaration of Helsinki (2013 revision) and with Good Clinical Practice guidelines.

3. Results

3.1. Characteristics of the study participants

Three hundred and twenty eligible patients were enrolled with no missing outcome data, and all 320 were analysed. Their sociodemographic and clinical characteristics, together with the six composite instrument scores, are detailed in Table 1. Mean age was 45.9 ± 12.0 years, with 124 patients (38.8%) in the 46–60 year band and 39 (12.2%) older than 60 years; 174 participants (54.4%) were male and 146 (45.6%) female. The educational distribution shown in Table 1 spans the full national range: 22 participants (6.9%) had no formal education, 59 (18.4%) primary schooling, 105 (32.8%) secondary schooling, 104 (32.5%) a diploma or bachelor's degree and 30 (9.4%) a postgraduate qualification. Major abdominal procedures accounted for the largest share of the operative case-mix (116 patients, 36.2%), followed by cardiac (79, 24.7%), orthopaedic (73, 22.8%) and thoracic (52, 16.2%) procedures; general anaesthesia was the most frequent modality (186 patients, 58.1%).

Table 1. Sociodemographic and clinical characteristics of the study participants (n = 320).

Characteristic	Value	Additional notes
Sociodemographic profile		
Age (years), mean ± SD	45.9 ± 12.0	Kolmogorov–Smirnov P = 0.613 (normal)
Age band, n (%)	< 30: 42 (13.1); 30–45: 115 (35.9); 46–60: 124 (38.8); > 60: 39 (12.2)	
Sex, n (%)	Male 174 (54.4); female 146 (45.6)	
Educational attainment, n (%)	No formal 22 (6.9); primary 59 (18.4); secondary 105 (32.8); diploma or bachelor 104 (32.5); postgraduate 30 (9.4)	Ordered variable; trend tested in Table 2
Monthly income (IDR), n (%)	< 2 m 70 (21.9); 2–5 m 105 (32.8); 5–10 m 106 (33.1); > 10 m 39 (12.2)	Ordered variable; trend tested in Table 2
Clinical and operative profile		
Surgery type, n (%)	Cardiac 79 (24.7); thoracic 52 (16.2); major abdominal 116 (36.2); orthopaedic 73 (22.8)	All ASA physical status III–IV elective procedures: valve replacement or repair and CABG; lobectomy, pneumonectomy and thoracotomy; hepatectomy, pancreatectomy, Whipple and colectomy; THA, TKA and major spinal instrumentation
Anaesthesia type, n (%)	General 186 (58.1); regional 99 (30.9); combined 35 (10.9)	
Composite instrument scores (0–100 scale), mean ± SD		
Health Literacy Level (HLL)	52.9 ± 21.7	KS P = 0.261 (normal)
Surgeon Communication Quality (SCQ)	53.7 ± 19.4	KS P = 0.980 (normal)
Family Influence Score (FIS)	57.3 ± 18.8	KS P = 0.971 (normal)
Collectivism Index (CI)	55.1 ± 17.9	KS P = 0.893 (normal)
IC Comprehension Score (ICCS)	49.4 ± 20.0	KS P = 0.970 (normal)
IC Efficacy Score (ICE)	40.5 ± 19.8	KS P = 0.985 (normal)
Primary outcome		
Adequate IC efficacy (ICE ≥ 70)	19 (5.9%)	Clopper–Pearson exact 95% CI 3.6–9.1
Inadequate IC efficacy (ICE < 70)	301 (94.1%)	Clopper–Pearson exact 95% CI 90.9–96.4

Notes: ASA = American Society of Anesthesiologists; CABG = coronary artery bypass grafting; IC = informed consent; IDR = Indonesian rupiah; KS = Kolmogorov–Smirnov normality test; m = million; SD = standard deviation; THA = total hip arthroplasty; TKA = total knee arthroplasty. Setting: Department of Surgery, Private Hospital X, Palembang, South Sumatra, Indonesia, January–December 2024.

As also detailed in Table 1, mean composite scores were moderate across all six instruments, ranging from 49.4 ± 20.0 for ICCS to 57.3 ± 18.8 for FIS, and all continuous variables were approximately normally distributed by Kolmogorov–Smirnov testing (all P ≥ 0.261). Mean IC Efficacy Score was 40.5 ± 19.8. Only 19 of 320 participants achieved adequate IC efficacy, a prevalence of 5.9% with a Clopper–Pearson exact 95% confidence interval of 3.6% to 9.1%; the remaining 301 patients (94.1%, 95% CI 90.9–96.4) did not. This distribution is displayed in Figure 2A.

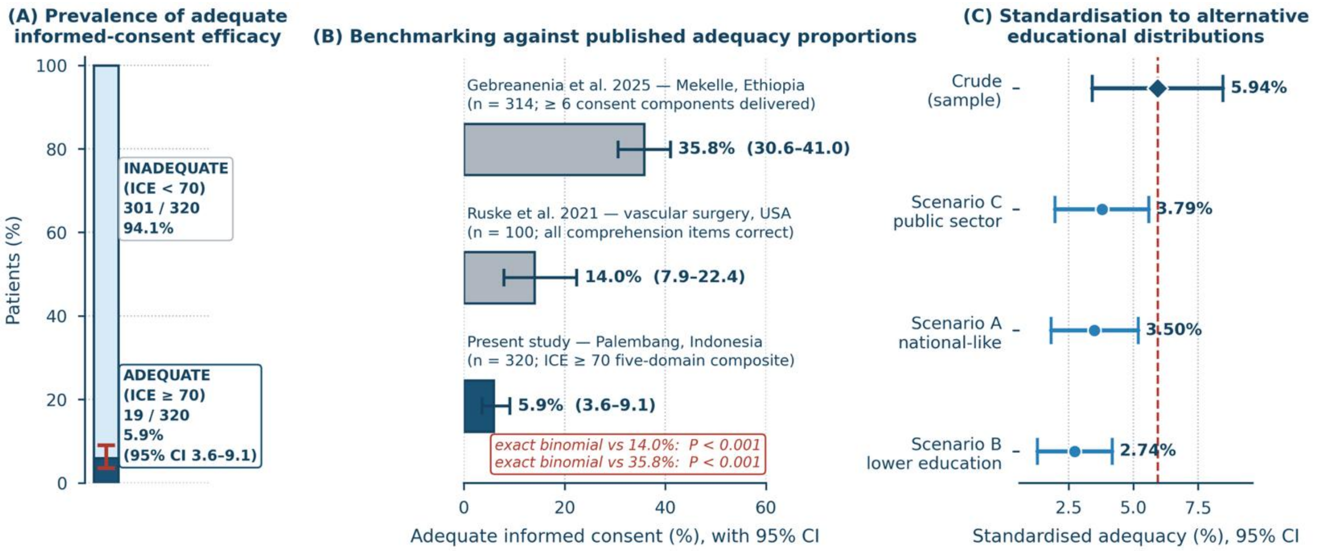

Figure 2. Prevalence of adequate informed-consent efficacy, international benchmarking and standardisation. (A) Nineteen of 320 patients (5.9%, Clopper–Pearson exact 95% CI 3.6–9.1) reached the ICE ≥ 70 threshold. (B) The observed proportion against the two published studies reporting an explicit adequacy proportion for a comparable surgical population; error bars are exact 95% confidence intervals and the comparison is indicative rather than like-for-like because the present outcome is a five-domain composite. (C) Direct standardisation of the stratum-specific adequacy rates to three alternative educational distributions; the dashed line marks the crude estimate.

Figure 2B sets that prevalence against the two published studies identified that report an explicit adequacy proportion for a comparable surgical population. It is lower than the 14.0% (95% CI 7.9–22.4) of vascular surgery patients who answered every comprehension item correctly in a high-income setting,[7] and lower than the 35.8% (95% CI 30.6–41.0) of Ethiopian surgical patients who received at least six of the recommended consent components;[10] the exact binomial comparison against each is significant at P < 0.001. Because the ICE threshold is a five-domain composite whereas both comparators measure a single dimension, this comparison is stated as indicative rather than like-for-like, and the caveat is developed in the Discussion.

Two internal-consistency checks were performed on the source tabulations before any inference. Every row and column margin in Tables 1 to 4 sums correctly to 320 participants and to 19 events, and the overall comprehension mean implied by the two group means and sizes, (19 × 98.3 + 301 × 46.1)/320 = 49.20, agrees with the 49.4 reported in Table 1 to within the rounding of the group means. The contingency tables are therefore internally consistent, and the discrepancies reported below are confined to the test statistics computed from them, not to the cell counts.

3.2. Categorical determinants of informed-consent efficacy

Bivariate associations between categorical characteristics and IC efficacy adequacy, with exact tests and stratum-specific exact confidence intervals, are presented in Table 2. Educational attainment was the only categorical characteristic associated with adequacy. As shown in Table 2 and displayed in Figure 3A, adequacy rose monotonically across the five educational strata: 0.0% (0/22, 95% CI 0.0–15.4) among participants with no formal education, 0.0% (0/59, 95% CI 0.0–6.1) among those with primary schooling, 2.9% (3/105, 95% CI 0.6–8.1) at secondary level, 11.5% (12/104, 95% CI 6.1–19.3) among diploma or bachelor's graduates and 13.3% (4/30, 95% CI 3.8–30.7) among postgraduates. The exact Cochran–Armitage test confirmed a significant monotone trend (z = 3.66, P = 0.0003), the Fisher–Freeman–Halton exact test was likewise significant (P = 0.005), and Cramér's V of 0.221 indicates a moderate association.

No other categorical characteristic was associated with adequacy, as Table 2 shows. Adequacy was 5.2% (9/174, 95% CI 2.4–9.6) in men and 6.8% (10/146, 95% CI 3.3–12.2) in women (exact P = 0.636). Across ascending income strata it was 2.9% (2/70), 5.7% (6/105), 6.6% (7/106) and 10.3% (4/39) (exact P = 0.470; exact Cochran–Armitage z = 1.545, P = 0.138); across cardiac, thoracic, major abdominal and orthopaedic procedures it was 5.1% (4/79), 5.8% (3/52), 6.9% (8/116) and 5.5% (4/73) (exact P = 0.962); and across general, regional and combined anaesthesia it was 5.4% (10/186), 7.1% (7/99) and 5.7% (2/35) (exact P = 0.885). The corresponding exact confidence intervals for every stratum are given in Table 2.

Table 2. Bivariate association between categorical characteristics and adequate informed-consent efficacy (ICE ≥ 70), with exact tests and exact stratum confidence intervals (n = 320).

Characteristic	Adequate IC, n/N (%)	Exact 95% CI (%)	Pearson χ2 (df) †	Exact P ‡	Cramér's V	Cells with expected count < 5
Sex						
Male	9/174 (5.2)	2.4–9.6	0.400 (1)	0.636	0.035	0 of 2
Female	10/146 (6.8)	3.3–12.2				
Educational attainment						
No formal education	0/22 (0.0)	0.0–15.4	15.677 (4)	0.005 §	0.221	3 of 5
Primary school	0/59 (0.0)	0.0–6.1				
Secondary school	3/105 (2.9)	0.6–8.1				
Diploma / bachelor	12/104 (11.5)	6.1–19.3				
Postgraduate	4/30 (13.3)	3.8–30.7				
Monthly income (IDR)						
< 2 million	2/70 (2.9)	0.3–9.9	2.585 (3)	0.470	0.090	2 of 4
2–5 million	6/105 (5.7)	2.1–12.0				
5–10 million	7/106 (6.6)	2.7–13.1				
> 10 million	4/39 (10.3)	2.9–24.2				
Surgery type						
Cardiac	4/79 (5.1)	1.4–12.5	0.329 (3)	0.962	0.032	3 of 4
Thoracic	3/52 (5.8)	1.2–15.9				
Major abdominal	8/116 (6.9)	3.0–13.1				
Orthopaedic	4/73 (5.5)	1.5–13.4				
Anaesthesia type						
General	10/186 (5.4)	2.6–9.7	0.336 (2)	0.885	0.032	1 of 3
Regional	7/99 (7.1)	2.9–14.0				
Combined	2/35 (5.7)	0.7–19.2				
Collapsed educational contrast (added in the present analysis)						
Tertiary (diploma / bachelor / postgraduate)	16/134 (11.9)	7.0–18.7	—	< 0.001 §	—	0 of 2
Secondary school or below	3/186 (1.6)	0.3–4.6				

Notes: † Pearson χ2 recomputed from the cross-tabulation; reported for comparability only, because expected counts below five invalidate the asymptotic statistic in four of the five tables (see Figure 3B). ‡ Fisher–Freeman–Halton exact test by Monte-Carlo permutation, 2×10^5 replicates, random seed 20260812. § Statistically significant. Exact Cochran–Armitage trend test: education z = 3.656, P = 0.00032; monthly income z = 1.545, P = 0.138. Collapsed educational contrast: risk ratio 7.40 (95% CI 2.20–24.90); odds ratio 8.27 (95% CI 2.36–29.00); Fisher exact P = 0.00015. CI = confidence interval; IC = informed consent; IDR = Indonesian rupiah.

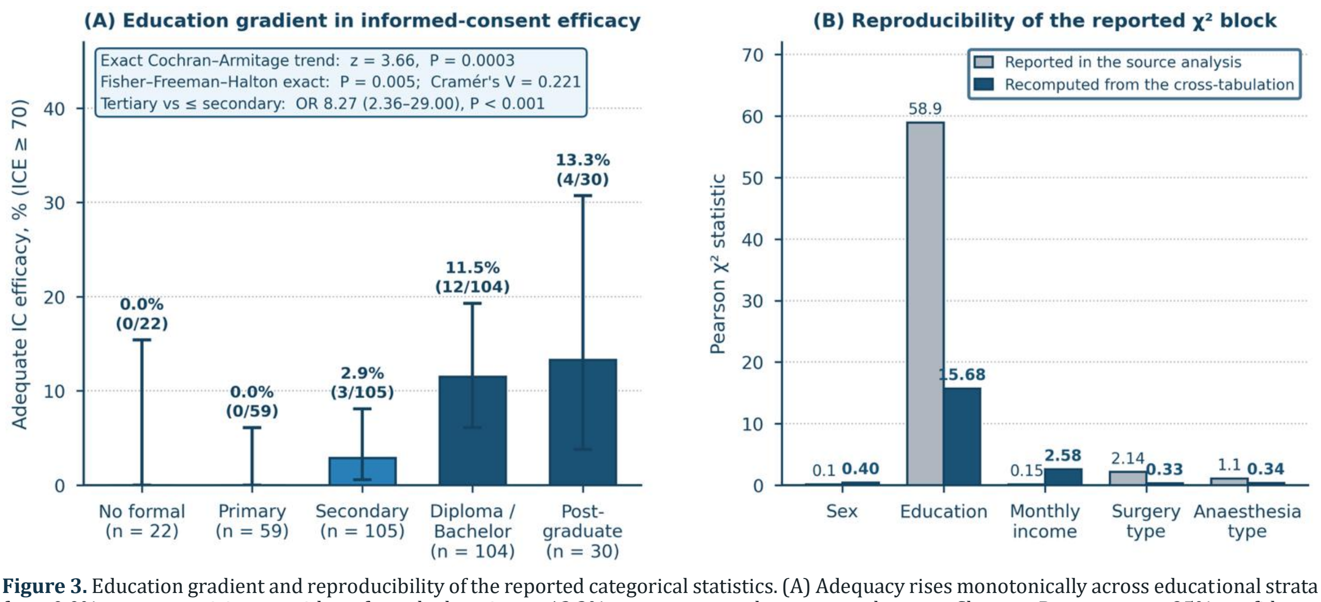

Figure 3. Education gradient and reproducibility of the reported categorical statistics. (A) Adequacy rises monotonically across educational strata from 0.0% among participants with no formal education to 13.3% among postgraduates; error bars are Clopper–Pearson exact 95% confidence intervals and the trend is tested by an exact Cochran–Armitage permutation test. (B) Reported versus recomputed Pearson χ2 statistics for the five categorical comparisons; none of the five reported values reproduces from the published cross-tabulation, which is why the exact tests in Table 2 supersede them.

The right-hand column of Table 2 records that expected cell counts fell below five in four of these five tables, so the exact tests rather than the asymptotic χ^2 statistics are the valid inferential basis. Figure 3B compares the reported and recomputed χ^2 statistics directly: none of the five reported values reproduces from the published cross-tabulation, the largest discrepancy being the education statistic (reported 58.9, recomputed 15.68). Because the underlying cell counts are internally consistent, no substantive conclusion changes.

The educational contrast was quantified further by collapsing the strata into tertiary-educated (diploma, bachelor's or postgraduate) versus secondary-educated or below, shown in the final block of Table 2. Adequacy was 11.9% (16/134, 95% CI 7.0–18.7) versus 1.6% (3/186, 95% CI 0.3–4.6), a risk ratio of 7.40 (95% CI 2.20–24.90) and an odds ratio of 8.27 (95% CI 2.36–29.00; Fisher exact P < 0.001).

3.3. Continuous determinants, effect sizes and discrimination

Continuous comparisons between the adequate and inadequate groups, with standardised effect sizes and discrimination statistics, are presented in Table 3. Three constructs differed significantly. ICCS was 98.3 ± 2.1 in the adequate group versus 46.1 ± 15.8 in the inadequate group (U = 233.0, z = −6.715, P < 0.001), a standardised mean difference of d = 3.400 (95% CI 2.866 to 3.933). HLL was 84.7 ± 12.3 versus 51.1 ± 20.4 (U = 665.0, z = −5.611, P < 0.001; d = 1.678, 95% CI 1.196 to 2.159), and SCQ was 66.4 ± 17.5 versus 53.0 ± 19.3 (U = 1645.5, z = −3.104, P = 0.002; d = 0.698, 95% CI 0.231 to 1.165). As Table 3 also shows, FIS (60.9 ± 15.8 versus 57.1 ± 18.9; P = 0.172; d = 0.203, 95% CI −0.261 to 0.667), CI (58.6 ± 14.2 versus 55.0 ± 18.0; P = 0.490; d = 0.202, 95% CI −0.262 to 0.666) and age (44.1 ± 10.3 versus 45.9 ± 12.1; P = 0.995; d = −0.150, 95% CI −0.614 to 0.314) did not differ.

Table 3. Continuous informed-consent constructs by efficacy adequacy: group comparison, standardised effect size and discrimination (n = 320; 19 adequate, 301 inadequate).

Construct	Adequate IC (n = 19)	Inadequate IC (n = 301)	Mann–Whitney U (z)	P	Cohen's d (95% CI)	AUC (95% CI)	Spearman ρ with ICE (95% CI)
IC Comprehension Score (ICCS)	98.3 ± 2.1	46.1 ± 15.8	233.0 (−6.715)	< 0.001 †	3.400 (2.866 to 3.933)	0.959 (0.826–0.991)	0.773 (0.725–0.814) †
Health Literacy Level (HLL)	84.7 ± 12.3	51.1 ± 20.4	665.0 (−5.611)	< 0.001 †	1.678 (1.196 to 2.159)	0.884 (0.743–0.952)	0.566 (0.487–0.636) †
Surgeon Communication Quality (SCQ)	66.4 ± 17.5	53.0 ± 19.3	1645.5 (−3.104)	0.002 †	0.698 (0.231 to 1.165)	0.712 (0.563–0.826)	0.439 (0.346–0.523) †
Family Influence Score (FIS)	60.9 ± 15.8	57.1 ± 18.9	2325.5 (−1.365)	0.172	0.203 (−0.261 to 0.667)	0.593 (0.451–0.721)	−0.139 (−0.245 to −0.030) †
Collectivism Index (CI)	58.6 ± 14.2	55.0 ± 18.0	2589.5 (−0.690)	0.490	0.202 (−0.262 to 0.666)	0.547 (0.410–0.677)	0.060 (−0.050 to 0.169)
Age (years)	44.1 ± 10.3	45.9 ± 12.1	2857.0 (−0.006)	0.995	−0.150 (−0.614 to 0.314)	0.500 (0.369–0.631)	—

Notes: Values are mean ± standard deviation. † Statistically significant at α = 0.05 (two-sided). AUC = area under the receiver operating characteristic curve, recovered from the Mann–Whitney U statistic as AUC = $1 - U/(n_1 n0)$ with a Hanley–McNeil standard error and a logit-transformed 95% confidence interval; the corresponding curves are shown in Figure 4A. Cohen's d computed from the group means and standard deviations using the pooled standard deviation. Spearman ρ computed against the continuous IC Efficacy Score with Fisher-z confidence intervals. CI = confidence interval; ICE = IC Efficacy Score.

Recovering the area under the receiver operating characteristic curve from each Mann–Whitney U statistic produced the discrimination hierarchy given in the penultimate column of Table 3 and displayed in Figure 4. ICCS discriminated adequate from inadequate consent almost perfectly (AUC 0.959, 95% CI 0.826–0.991), followed by HLL (AUC 0.884, 95% CI 0.743–0.952) and SCQ (AUC 0.712, 95% CI 0.563–0.826). FIS (AUC 0.593, 95% CI 0.451–0.721), CI (AUC 0.547, 95% CI 0.410–0.677) and age (AUC 0.500, 95% CI 0.369–0.631) had confidence intervals that included 0.5 and therefore did not discriminate. As Figure 4B makes visually explicit, the confidence interval for ICCS does not overlap those for FIS, CI or age, so the ordering of the extremes of this hierarchy is itself statistically supported.

Spearman correlations with the continuous ICE score, given in the final column of Table 3, reproduced the same ordering: ICCS ρ = 0.773 (95% CI 0.725–0.814, P < 0.001), HLL ρ = 0.566 (95% CI 0.487–0.636, P < 0.001), SCQ ρ = 0.439 (95% CI 0.346–0.523, P < 0.001), FIS ρ = −0.139 (95% CI −0.245 to −0.030, P = 0.013) and CI ρ = 0.060 (95% CI −0.050 to 0.169, P = 0.285). The family-influence correlation is negative and statistically significant but small, and its confidence interval excludes any association stronger than ρ = −0.245.

3.4. Multivariable model and its attainable intervals

The multivariable logistic regression is presented in Table 4 and displayed as a forest plot in Figure 5A. After simultaneous adjustment, ICCS was the only independent predictor of adequate IC efficacy (β = 2.621, SE = 0.725; odds ratio 13.75 per standard deviation, 95% CI 3.32–56.92, P < 0.001). No other predictor approached significance: HLL (odds ratio 1.04, 95% CI 0.14–7.59, P = 0.968), SCQ (1.42, 95% CI 0.62–3.27, P = 0.404), FIS (1.09, 95% CI 0.57–2.08, P = 0.804), CI (1.15, 95% CI 0.59–2.25, P = 0.682), male sex (0.34, 95% CI 0.08–1.47, P = 0.147) and education entered as an ordinal term (3.09, 95% CI 0.44–21.46, P = 0.255). As Figure 5A makes plain, five of the seven predictors have confidence intervals spanning more than an order of magnitude, and the model should be read as uninformative about those predictors conditional on comprehension rather than as evidence that they have no effect. At 2.71 events per variable the multivariable analysis is exploratory.

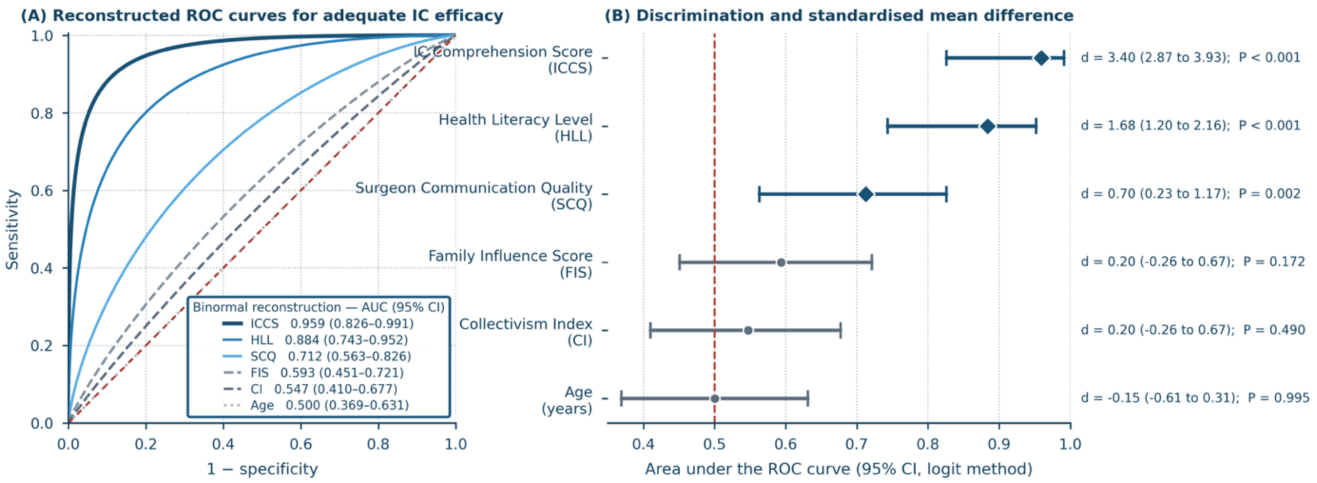

Figure 4. Discrimination of the informed-consent constructs for adequate consent efficacy. (A) Receiver operating characteristic curves reconstructed under a binormal model from the areas recovered from each Mann–Whitney U statistic; the diagonal marks chance performance. (B) Areas under the curve with logit-transformed 95% confidence intervals, annotated with Cohen's d and the exact two-sided P value; diamonds mark statistically significant discrimination and the dashed line marks an area of 0.5.

Table 4. Multivariable binary logistic regression predicting adequate informed-consent efficacy (ICE ≥ 70), with the attainable interval for each quantity from the partial-identification audit (n = 320, 19 events).

Predictor	β (SE)	Odds ratio	95% Wald CI	P	Attainable interval across 600 admissible reconstructions
Constant	−5.784 (0.965)	0.003	0.000–0.020	< 0.001 †	—
IC Comprehension Score (per SD)	2.621 (0.725)	13.75	3.32–56.92	< 0.001 †	β 2.49–3.13; SE 0.65–0.94; OR 12.07–22.84; lower limit 2.81–4.06 — reported values corroborated
Health Literacy Level (per SD)	0.040 (1.014)	1.04	0.14–7.59	0.968	Not separately identified
Surgeon Communication Quality (per SD)	0.354 (0.424)	1.42	0.62–3.27	0.404	Not separately identified
Family Influence Score (per SD)	0.082 (0.331)	1.09	0.57–2.08	0.804	Not separately identified
Collectivism Index (per SD)	0.141 (0.343)	1.15	0.59–2.25	0.682	Not separately identified
Male sex (vs female)	−1.087 (0.750)	0.34	0.08–1.47	0.147	Not separately identified
Education level (per ordinal category)	1.127 (0.989)	3.09	0.44–21.46	0.255	Exact unadjusted contrast: OR 8.27 (2.36–29.00), P < 0.001 (Table 2)
Model summary and sensitivity analysis					
Nagelkerke pseudo-R2	—	0.554 (reported)	—	—	0.661–0.723 — reported value lies below the attainable floor and should not be cited (Figure 5B)
Firth-penalised ICCS estimate (per SD)	2.310 (0.476)	10.08	3.96–25.65	< 0.001 †	OR 8.62–13.40 across all admissible reconstructions
Events per variable	—	2.71	—	—	19 events / 7 parameters; 70 events (n ≈ 1,180) required for 10 events per variable

Notes: † Statistically significant at α = 0.05. All continuous predictors z-standardised prior to entry. The attainable interval is the range of each quantity across 600 datasets simultaneously consistent with the published group means and standard deviations, the 0–100 instrument range, and the reported Mann–Whitney U statistic of 233, which fixes the degree of group overlap. Falling inside an attainable interval means a reported value is not refuted; falling outside it is decisive. β = logistic regression coefficient; CI = confidence interval; OR = odds ratio; SD = standard deviation; SE = standard error.

The identification audit, summarised in the final column of Table 4, corroborated the ICCS estimate and isolated one quantity that the data cannot support. Samuelson's inequality established that every one of the 19 patients with adequate consent had an ICCS of at least 89.39 (Methods 2.8), and that at most 39 of the 301 inadequate patients could reach that value; the reported Mann–Whitney U of 233 pins the actual overlap tighter still, at the equivalent of approximately 12 to 13 inadequate patients scoring within the adequate group's range. Across the 600 admissible reconstructions satisfying all of these constraints, the univariable ICCS coefficient ranged from β = 2.49 to 3.13 (median 2.71) with a standard error of 0.65 to 0.94, corresponding to odds ratios of 12.07 to 22.84 with lower confidence limits of 2.81 to 4.06. The reported multivariable estimate (β = 2.621, SE = 0.725, odds ratio 13.75, lower limit 3.32) falls inside every one of these attainable intervals, so the headline association is reproducible and not an artefact of separation. The Firth-penalised estimate across the same reconstructions was an odds ratio of 10.08 (95% CI 3.96–25.65; range across reconstructions 8.62 to 13.40), approximately 27% smaller than the unpenalised value, indicating modest small-sample upward bias but no change in inferential direction.

The Nagelkerke pseudo-R^2 is the one reported quantity that the data refute, as Figure 5B displays. The argument does not require reconstruction: because the univariable comprehension model is nested within the full model, the full model's pseudo-R^2 cannot be smaller than the univariable value. Across all 600 admissible reconstructions the univariable comprehension model achieved a Nagelkerke R^2 between 0.661 and 0.723, and the range was stable from the first eighty configurations onward. The reported full-model value of 0.554 therefore lies below the attainable floor for the model as specified and should not be cited as an estimate of explained variation. The point does not affect any effect estimate, only the goodness-of-fit summary.

3.5. Robustness of the education gradient and precision of the null findings

The robustness metrics for the education gradient and the precision of each null finding are presented in Table 5. The E-value for the collapsed educational contrast is 14.29, falling to 3.83 at the confidence limit, so an unmeasured confounder would have to be associated with both education and consent adequacy by a risk ratio of at least 3.83 to move the confidence interval across the null. For calibration, the strongest association actually measured in this study, that of health literacy, corresponds to an AUC of 0.884; a confounder of the required strength would therefore have to exceed every variable the study recorded. The fragility index is 8, with a fragility quotient of 0.025, meaning that eight of the sixteen events in the tertiary-educated stratum would have to be reclassified before significance was lost — an unusually high value relative to the event count.

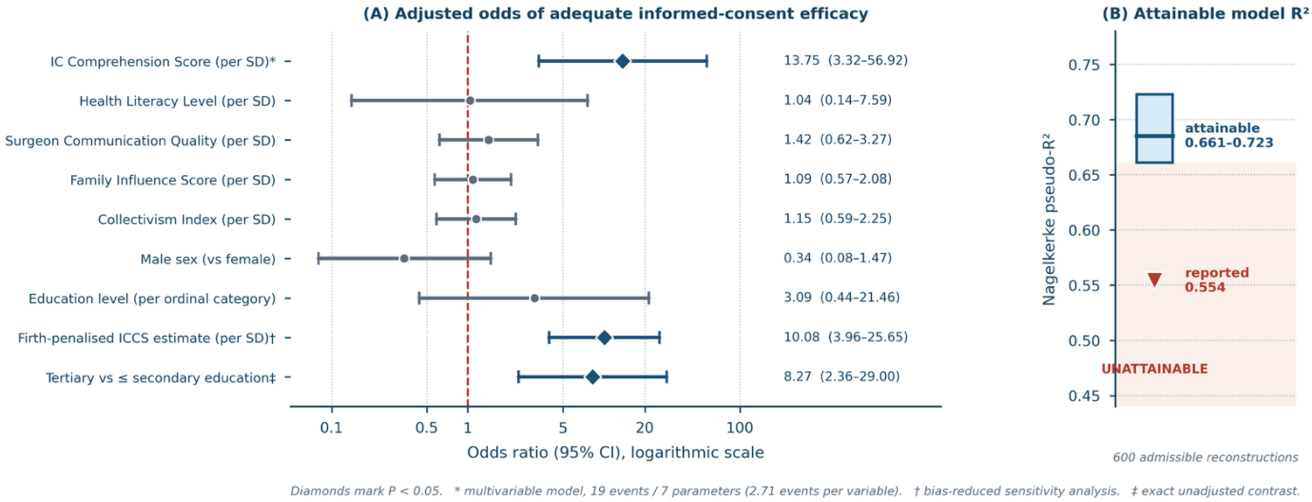

Figure 5. Multivariable model and its identification audit. (A) Odds ratios with 95% confidence intervals on a logarithmic scale; diamonds mark statistically significant estimates, the dashed line marks the null, and the Firth-penalised and exact unadjusted educational contrasts are shown for comparison. (B) The attainable interval for the Nagelkerke pseudo-R2 across all 600 admissible reconstructions, against the reported value of 0.554, which lies in the unattainable region below the floor set by the nested univariable model.

Table 5. Robustness of the education gradient and precision of the non-significant findings.

Quantity	Estimate (95% CI)	Interpretation
Robustness of the education gradient		
Risk ratio, tertiary vs ≤ secondary education	7.40 (2.20–24.90)	Large and precisely signed
Odds ratio, tertiary vs ≤ secondary education	8.27 (2.36–29.00)	Fisher exact P = 0.00015
E-value for the point estimate	14.29	A confounder would need RR ≥ 14.29 with both education and adequacy
E-value at the confidence limit	3.83	RR ≥ 3.83 needed to move the interval across the null
Fragility index (fragility quotient)	8 (0.025)	Eight of the 16 events would have to be reclassified to lose significance
Precision of the non-significant categorical findings		
Female vs male sex	OR 1.35 (0.53–3.41)	Power 0.095; minimum detectable adequacy 14.4% vs 5.2% — uninformative
Regional vs general anaesthesia	OR 1.34 (0.49–3.63)	Power 0.091; minimum detectable adequacy 15.7% vs 5.4% — uninformative
Major abdominal vs other surgery	OR 1.30 (0.51–3.33)	Power 0.085; minimum detectable adequacy 15.0% vs 5.4% — uninformative
Income > 10 million vs ≤ 10 million IDR	OR 2.03 (0.64–6.45)	Power 0.276; minimum detectable adequacy 19.4% vs 5.3% — uninformative
Precision of the two cultural null findings		
Collectivism Index	AUC 0.547 (0.410–0.677); ρ 0.060 (−0.050 to 0.169)	Interval excludes any association of practical magnitude — informative null
Family Influence Score	AUC 0.593 (0.451–0.721); ρ −0.139 (−0.245 to −0.030)	Small negative association; interval excludes ρ < −0.245 — informative null

Notes: AUC = area under the receiver operating characteristic curve; CI = confidence interval; IDR = Indonesian rupiah; OR = odds ratio; RR = risk ratio; ρ = Spearman rank correlation with the continuous IC Efficacy Score. Minimum detectable adequacy is the proportion in the index stratum detectable at 80% power and α = 0.05 against the observed reference proportion.

Because a null result carries medicolegal weight only if the study could have detected the effect in question, Table 5 also reports the observed power and the minimum detectable adequacy proportion for each non-significant categorical predictor. Observed power ranged from 0.085 for the surgery-type contrast to 0.276 for the income contrast, and the minimum adequacy proportion detectable at 80% power ranged from 14.4% to 19.4% against reference rates of 5.2% to 5.4%. This study could therefore only have detected differences of roughly threefold or greater. The null findings for sex, surgery type and anaesthesia type are uninformative rather than negative, and no clinical or medicolegal conclusion of equivalence should be drawn from them.

The same caution does not apply to the two cultural constructs, and Table 5 separates them explicitly for that reason. For both FIS and CI the 95% confidence interval for the AUC excludes values above 0.721, and for the Collectivism Index the Spearman correlation with ICE is bounded above at $\rho = 0.169$. These two nulls are precise enough to interpret.

3.6. Standardisation and the direction of the selection biases

The sample is educationally advantaged relative to the Indonesian adult population: 134 of 320 participants (41.9%) held a diploma or higher, and these are the strata in which consent adequacy is concentrated. Direct standardisation of the stratum-specific adequacy rates to three alternative educational distributions is presented in Table 6 and displayed in Figure 2C. Against a national-like distribution the standardised adequacy rate falls to 3.50% (95% CI 1.82–5.19); against a lower-education skew it falls to 2.74% (95% CI 1.30–4.18); and against a public-sector tertiary-centre profile it is 3.79% (95% CI 1.98–5.60), compared with the crude 5.94% (95% CI 3.41–8.46). These weightings are illustrative rather than census-derived, and the intervals reflect sampling error in the stratum rates only, but the direction is unambiguous across all three scenarios.

Table 6. International benchmarking, direct standardisation to alternative educational distributions, and verification of the sample-size calculation.

Quantity	Estimate (95% CI)	Comparison or interpretation
Benchmarking against published adequacy proportions		
Present study — Palembang, Indonesia (n = 320)	5.9% (3.6–9.1)	Reference; ICE ≥ 70 five-domain composite
Ruske et al. 2021 — vascular surgery, USA (n = 100)	14.0% (7.9–22.4)	All comprehension items correct; exact binomial P < 0.001
Gebreanenia et al. 2025 — Mekelle, Ethiopia (n = 314)	35.8% (30.6–41.0)	≥ 6 recommended consent components delivered; exact binomial P < 0.001
Direct standardisation of the adequacy rate to alternative educational distributions		
Sample distribution (crude)	5.94% (3.41–8.46)	Observed; 134 of 320 participants (41.9%) held a diploma or higher
Scenario A — national-like (5 / 35 / 40 / 17 / 3%)	3.50% (1.82–5.19)	41% lower than the crude estimate
Scenario B — lower-education skew (8 / 40 / 38 / 12 / 2%)	2.74% (1.30–4.18)	54% lower than the crude estimate
Scenario C — public-sector tertiary centre (6 / 30 / 42 / 19 / 3%)	3.79% (1.98–5.60)	36% lower than the crude estimate
Verification of the sample-size calculation		
Recomputed design requirement (Z = 1.96, p = 0.35, d = 0.05)	n = 350 (389 with 10% attrition)	Realised n = 320 is below the design specification
Achieved margin of error at the observed prevalence	± 2.6%	Adequate for the prevalence aim; the shortfall affects the multivariable aim only
Events per variable in the multivariable model	2.71	19 events / 7 parameters; 70 events (n ≈ 1,180) required for 10 per variable

Notes: Standardisation weights are stated in the order no formal education / primary / secondary / diploma or bachelor / postgraduate; they are illustrative rather than census-derived, and the confidence intervals reflect sampling error in the stratum-specific rates only. The benchmark comparisons are indicative rather than like-for-like, because the present outcome is a five-domain composite whereas both comparators measure a single dimension. CI = confidence interval; ICE = IC Efficacy Score.

That direction is reinforced by the selection structure, and by the sample-size verification recorded in the final block of Table 6. Exclusion of patients with cognitive impairment, exclusion of those unable to communicate in Bahasa Indonesia or a local dialect, and recruitment from a private rather than a public institution all remove or under-represent groups in which comprehension failure is more likely. All three biases, together with the educational skew, point the same way. The observed 5.9% should therefore be read as an upper bound on the adequacy rate at comparable Indonesian institutions rather than as a central estimate.

4. Discussion

This study quantifies comprehension-based informed-consent failure in a high-risk Indonesian surgical population and subjects the resulting model to a formal identification audit. Among 320 consecutive patients undergoing ASA III–IV elective surgery at Private Hospital X, Palembang, only 19 (5.9%, 95% CI 3.6–9.1) achieved adequate consent efficacy, and standardisation to lower-education population distributions reduces that estimate to between 2.74% and 3.79%, so 5.9% is an upper bound. Comprehension discriminated adequacy almost perfectly (AUC 0.959, 95% CI 0.826–0.991) and was the sole independent multivariable predictor (odds ratio 13.75 per SD, 95% CI 3.32–56.92), an estimate the identification audit confirmed to lie within its attainable interval. Educational attainment produced a robust monotone gradient (exact trend z = 3.66, P = 0.0003; E-value 14.29; fragility index 8). Neither family influence nor cultural collectivism predicted consent adequacy.

The prevalence figure is the most arresting result and requires careful interpretation. At 5.9% it is below both published comparators that report an explicit adequacy proportion: 14.0% among vascular surgery patients in a high-income setting,[7] and 35.8% among Ethiopian surgical patients assessed against a six-component consent standard.[10] Two caveats discipline the comparison, and Figure 2B is labelled accordingly. First, ICE is a composite requiring all five of comprehension, voluntariness, disclosure adequacy, capacity confirmation and post-consent decisional clarity, whereas both comparators measure a single dimension; a composite requiring five conditions will necessarily yield a lower proportion than a single-condition measure. Second, the ICE ≥ 70 threshold is stringent. A comprehension-only adequacy rate, which would be directly comparable, cannot be computed here because the domain-level scores underlying the composite were not retained. The finding is therefore best stated as a conservative composite estimate rather than as a claim that Indonesian patients understand less than Ethiopian or American patients on a like-for-like instrument. What survives both caveats is the internal comparison: within a single instrument and a single centre, adequacy is indistinguishable from zero for the least-educated half of the surgical population.

That internal gradient is the study's most defensible finding and the one with the clearest forensic implication. No patient without formal or primary education achieved adequate consent efficacy, against 11.9% of the tertiary-educated. The E-value of 14.29 — 3.83 at the confidence limit — means that an unmeasured confounder would need an association with both education and consent adequacy stronger than any variable measured in this study, and the fragility index of 8 indicates that the finding does not rest on one or two outcome events. The practical reading is uncomfortable: the standard consent protocol at this institution is not merely suboptimal for less-educated patients, it is non-functional for them. This aligns with evidence from Ethiopia, Nepal, India and sub-Saharan Africa more broadly that educational attainment and functional literacy are the dominant correlates of consent comprehension in resource-constrained settings,[9,10,12,37] and with systematic-review evidence that health literacy is associated with surgical outcomes and with participation in shared decision-making.[33,36]

The centrality of comprehension is the study's second principal finding. Comprehension discriminates adequacy with an AUC of 0.959, and Figure 4B shows that its confidence interval does not overlap those of family influence, collectivism or age. In the adjusted model, health literacy, surgeon communication quality and education are all attenuated to non-significance once comprehension is included. The defensible statement is conditional: among patients with equivalent comprehension, health literacy, surgeon communication quality and education carry no further detectable association with consent adequacy. The natural but stronger reading of the same pattern, summarised in Figure 6, is that comprehension sits downstream of those determinants and transmits their effects, and we deliberately stop short of asserting it. Formal mediation decomposition requires the inter-predictor correlation matrix, which the aggregate data do not contain; positive-definiteness alone leaves the ICCS–HLL correlation bounded only to [−0.085, 0.961], and partial identification of a latent correlation under unverifiable distributional assumptions is a recognised problem rather than a technicality.[54] Any indirect-effect estimate computed from this dataset would be an assumption presented as a result. The conditional statement is sufficient for the practical inference and does not require an unidentified causal model to support it.

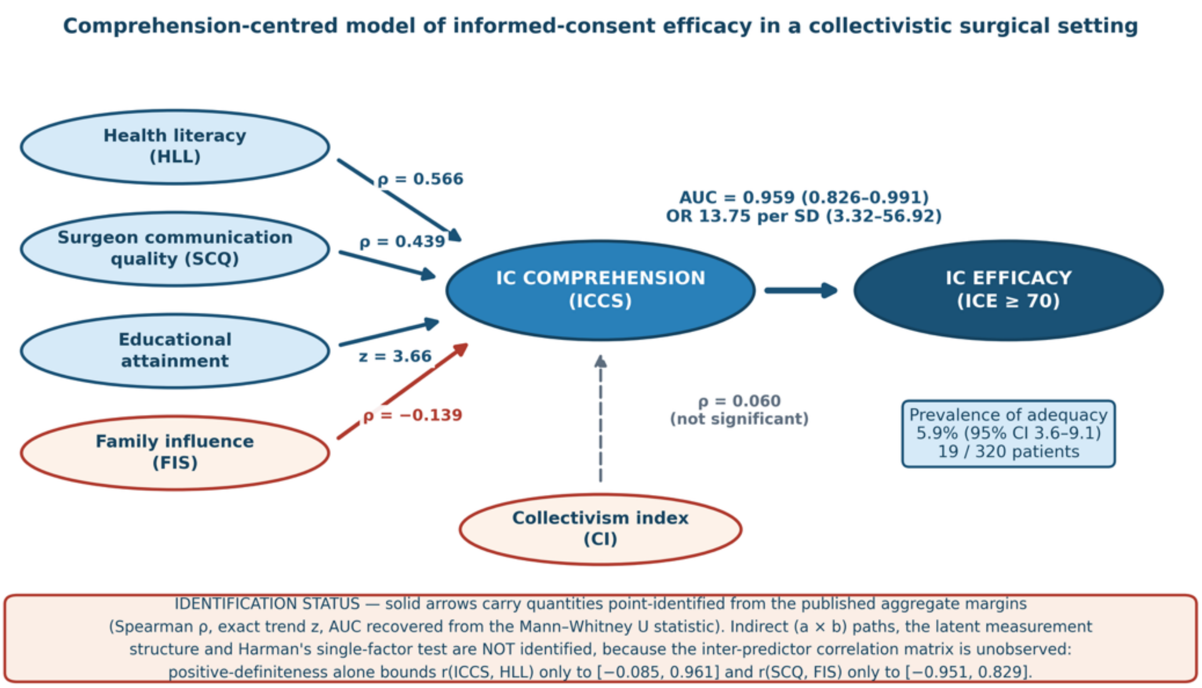

Figure 6. Comprehension-centred conceptual model of informed-consent efficacy, annotated with its identification status. Solid arrows carry quantities point-identified from the published aggregate margins. The indirect (a × b) paths, the latent measurement structure and Harman's single-factor test are not identified from these data because the inter-predictor correlation matrix is unobserved; the bounds shown are those implied by positive-definiteness alone.

The third finding is a negative one with unusual evidential strength. Neither the Family Influence Score nor the Collectivism Index predicted consent adequacy. FIS showed a small negative bivariate correlation with continuous consent efficacy (ρ = −0.139, 95% CI −0.245 to −0.030), whose confidence interval excludes any association stronger than ρ = −0.245, and the Collectivism Index showed no association at all (ρ = 0.060, upper limit 0.169). Unlike the sex, surgery-type and anaesthesia-type nulls — which Table 5 shows to be uninformative, with minimum detectable proportions of 14.4% to 19.4% against reference rates near 5% — these two nulls are precise enough to be interpreted. This is a direct quantitative test of the proposition, prominent in the bioethical literature, that collectivistic cultural orientation is itself a threat to consent validity. The data do not support it. They do support the alternative account, argued on conceptual grounds across several recent analyses, that the individualism–collectivism binary is too coarse an instrument and that relational autonomy better describes how consent operates in these settings.[39–43] If collectivism were the problem, it should have predicted the outcome; it did not. Comprehension did.

This reframing has direct consequences for intervention design. If family presence were a threat to autonomy, the appropriate response would be to exclude relatives from the consent consultation. If, as these data suggest, the operative deficit is comprehension, then family members are better understood as an underused educational resource: structured family consent sessions that educate patient and designated relative together may achieve larger comprehension gains than patient-only approaches, without any autonomy cost — a reading consistent with recent work on the perspectives of patients and their next of kin in lower-middle-income settings.[38] The intervention literature supports the general direction. A teach-back protocol improves both consent quality and surgeon trust at negligible additional consultation time,[27] randomised trials of digital and multimedia consent have improved information gain, recall and satisfaction in bariatric, urological, orthopaedic and mixed surgical populations,[21–26] and the digitalisation of consent is now sufficiently widespread to have been mapped systematically.[28] Given that consent-form readability is itself demonstrably inadequate and measurably affects both comprehension and refusal,[29–32] document simplification is a low-cost first step fully within institutional control.

The medicolegal implications for Indonesian practice are substantial and follow directly from the prevalence estimate. Under Law No. 17 of 2023 on Health, the Indonesian Civil Code and the Ministry of Health regulation governing medical consent, consent is legally valid only where the patient has demonstrably understood the disclosed information; a signature alone does not discharge the duty.[3,4] If 5.9% of high-risk surgical consents at an accredited Type B hospital meet a defensible comprehension standard — and the standardised estimates in Table 6 suggest the true figure at a public-sector centre would be lower still — then the great majority of consents performed there are vulnerable to challenge.

That exposure is not uniform across forums, and the distinction matters to the forensic assessor. In a civil negligence claim the question is whether a reasonable patient would have declined the procedure had the risk been comprehended, so the comprehension deficit must be shown to be material to the decision; retrospective claims series from other jurisdictions show that consent-related allegations are a recurrent and expensive component of surgical litigation.[58–60] In a professional-disciplinary proceeding the question is compliance with the standard of practice, and a documented failure to verify comprehension suffices without any showing of materiality. Where a criminal complaint is lodged, consent adequacy enters through the expert-evidence provisions of the Indonesian Code of Criminal Procedure (KUHAP), and it is the forensic physician who is asked to opine on whether the record supports a finding of comprehension — precisely the exercise for which appellate courts elsewhere have been shown to demand documentary evidence of understanding.[1,2] The exposure is in every case systemic rather than individual, which is the pattern that generates the largest aggregate liability, and it is compounded by the fact that consent documentation errors are themselves an established contributor to wrong-site surgery.[61]

The remedy is administratively simple. A brief structured comprehension check performed and recorded at the end of the consent consultation would simultaneously improve consent quality and generate contemporaneous documentary evidence of comprehension — precisely the evidence whose absence is the current vulnerability. To be evidentially useful, such a record should specify which information items were checked, the method used, who performed the check, whether re-explanation was required and on which items, and the patient's confirmation, with the time and the identity of the clinician. We recommend that verification meeting that minimum specification be incorporated as an explicit element of KARS accreditation standards for surgical services. We further recommend that the Indonesian Association of Forensic Medicine (PDFI) adopt a corresponding standard for the retrospective evaluation of consent adequacy, since its members are the practitioners asked to opine on that question and currently have no agreed benchmark against which to do so.

The Indonesian sociocultural and system context shapes how these recommendations should be implemented, even though collectivism did not predict the outcome. Family participation in medical decision-making is normative rather than exceptional in this setting, hierarchical communication norms discourage patients from asking clarifying questions, and structured consent-communication training is not a formal component of Indonesian surgical residency — a gap consistent with the finding that trainees elsewhere are routinely delegated the consent task without corresponding preparation.[18] The mean SCQ of 53.7 ± 19.4 recorded in Table 1 should be read against those structural constraints — high elective volume and compressed consultation time — rather than as a judgement about individual clinicians. Communication quality nonetheless remains the most directly modifiable system-level determinant of comprehension available within the existing service, and its bivariate association with adequacy was significant (AUC 0.712, 95% CI 0.563–0.826) even though it was attenuated after adjustment. Health-system and patient-safety capacity constraints documented in the Indonesian context make tool-based augmentation — simplified documents, visual risk aids, standardised teach-back scripts — more rapidly deployable than curriculum reform.[5,44]

Two features of the measurement design qualify every association reported here. Five of the six instruments are patient-reported and were administered in a single sitting by the same data collectors, which is the canonical configuration for common method variance; a shared respondent factor would inflate the associations among ICCS, HLL, FIS, CI and ICE. SCQ is the exception, being observer-rated, and it is also the construct with the weakest association with the outcome among the three significant predictors — a pattern consistent with, though not proof of, a common-method contribution to the patient-reported associations. Harman's single-factor test cannot be computed without item-level data and is in any case a weak diagnostic. Working in the opposite direction, the absence of reliability coefficients means that unquantified measurement error attenuates every association, which matters most for the two null findings: an attenuated collectivism association could in principle be a measurement artefact rather than a true absence of effect, and the informative-null claim for the Collectivism Index should be read with that caveat attached.

The study has several strengths. It is, to our knowledge, the first to estimate comprehension-based consent adequacy in a high-risk Indonesian surgical population while measuring family influence and cultural collectivism as competing determinants, and the first in this literature to report an effect size with a 95% confidence interval for every estimate, including exact intervals for all stratum-specific proportions. Consecutive sampling across a full calendar year with no missing outcome data minimises selection bias. Exact rather than asymptotic tests were used wherever expected counts were small, which is the case in four of the five contingency tables and which materially changes the inferential basis. The recovery of discrimination statistics from rank statistics adds a clinically interpretable metric the original analysis did not contain. Most unusually, the multivariable model was subjected to a formal identification audit that both corroborated the principal effect estimate and identified one reported quantity — the Nagelkerke pseudo-R^2 — that the data cannot support.

The limitations are equally important, and one is structural. The exposure and the outcome are not independent constructs: patient comprehension is one of the five domains from which the IC Efficacy Score is composed, so the association between ICCS and ICE is partly constitutive rather than predictive. An AUC of 0.959 and a Cohen's d of 3.400 between a component and a composite containing that component are in part an arithmetic property of the composite's construction rather than evidence of an empirical relationship, and the magnitude of the artefact depends on the weight comprehension carries within the composite. That weighting was not documented and cannot be recovered, so the artefact cannot be bounded; nor can a sensitivity analysis excluding the comprehension domain be performed, because the domain-level scores were not retained. The odds ratio of 13.75 must therefore be read as a measure of internal coherence within the consent construct rather than as a predictive or causal estimate, and the claim that comprehension gates consent validity rests on the conceptual argument and on the education gradient rather than on that coefficient. The remaining findings are unaffected, because none involves a component–composite pairing: the prevalence estimate, the education gradient, the health-literacy association and the two cultural nulls all stand independently of this issue.

Six further limitations apply. The design is cross-sectional and all constructs were measured concurrently, so no temporal ordering and therefore no causal claim is established. The multivariable model rests on 19 events across seven parameters, giving 2.71 events per variable, well below conventional adequacy; the Firth-penalised sensitivity analysis suggests the unpenalised odds ratio is upwardly biased by roughly a quarter, and the wide Wald intervals in Table 4 are the honest summary of precision. A properly powered replication would require approximately 70 events, or about 1,180 patients at the observed rate, on the events-per-variable criterion — a rule-of-thumb minimum rather than a formal sample-size calculation. The study is single-centre and set in a private Type B hospital, so external validity to public-sector and lower-tier Indonesian surgical populations is unestablished, and the standardisation in Table 6 illustrates that sensitivity rather than substituting for multicentre data. Four of the five reported χ^2 statistics could not be reproduced from the published cross-tabulations, as Figure 3B shows, which is why the exact tests reported here supersede them; the cell counts are internally consistent and no conclusion drawn from them is affected. The instruments have limited published psychometric validation in this population and no item-level data were available to evaluate their measurement structure. Finally, the screening log recording exclusions was not retained, so the STROBE participant-flow counts cannot be given.

Future work should proceed in three directions. A prospective multicentre study across public and private Indonesian centres, powered for the multivariable aim at ten events per variable, would establish generalisability and permit the latent-variable analysis the present data cannot support. A pragmatic randomised trial of structured family-inclusive teach-back against standard consent, with comprehension as the primary endpoint, would test the intervention implication directly. And a linked medicolegal cohort, following consent quality forward to complaints and claims, would convert the association reported here into the outcome that ultimately matters for forensic practice.

5. Conclusion

Comprehension-based informed-consent efficacy was profoundly inadequate among patients undergoing high-risk elective surgery at this Indonesian tertiary centre: only 19 of 320 patients (5.9%, 95% CI 3.6–9.1) reached the adequacy threshold, and standardisation to lower-education population distributions places the true figure between 2.74% and 3.79%, so 5.9% is an upper bound. Patient comprehension was the sole independent determinant in an exploratory multivariable model (odds ratio 13.75 per standard deviation, 95% CI 3.32–56.92; AUC 0.959, 95% CI 0.826–0.991), and educational attainment produced a robust monotone gradient (exact trend P = 0.0003; E-value 14.29) with no adequate consent at all below secondary schooling. Neither family influence nor cultural collectivism predicted consent adequacy, and those two null findings are precise enough to interpret. Indonesian surgical consent practice should therefore shift from signature-based procedural compliance to comprehension-verified consent, with a brief recorded comprehension check made an explicit requirement of KARS surgical accreditation standards and of PDFI guidance on the forensic evaluation of consent adequacy.

Declarations

Use of generative artificial intelligence. The authors declare that no generative artificial intelligence (AI) tools were used in the preparation, drafting, analysis, or editing of this manuscript.

Author contributions. MM conceived the study, supervised data collection and drafted the manuscript. CC contributed to study design and to the clinical interpretation of the operative case-mix.

YFH contributed to the analytical strategy and to the verification analyses. DT contributed the medicolegal and regulatory analysis. All authors reviewed and approved the final version of the manuscript.

Funding. This study received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflict of interest. The authors declare that they have no competing interests.

Ethical approval. Ethical approval was granted by the Institutional Ethics Committee of the CMHC Research Center under Reference No. 0024/2024. Written informed consent for research participation was obtained from every participant prior to enrolment. The study was conducted in accordance with the Declaration of Helsinki (2013 revision) and with Good Clinical Practice guidelines.

Consent for publication. Not applicable; no individually identifiable data are reported.

Data availability. The aggregate data supporting the findings, together with the complete analysis script including the reconstruction and standardisation procedures, are available from the corresponding author on reasonable request. Item-level and domain-level data were not retained and are not available.

References

1. Şeker Z, Özesen TA, Kaya K, et al. Evaluation of Supreme Court Decisions Regarding Informed Consent in Lawsuits Opened with Alleged Medical Malpractice: Cross-Sectional Research. Turkiye Klin J Forensic Med. 2023;20(1):11-23. doi:10.5336/forensic.2022-93835
2. Vural T, Erbaş M. Evaluating the Opinions of the Supreme Court and the Council of State on “Informed Consent (Assent)” in Compensation Cases Filed with Medical Malpractice Claim. Bull Leg Med. 2024;29(1):20-28. doi:10.17986/blm.1665
3. Sutarno S, Maryati M. Information of medical malpractice and risks in the informed consent process before surgery in Indonesia. Yustisia J Huk. 2021;10(2):269. doi:10.20961/yustisia.v10i2.52132
4. Lazuardi I, Marwiyah S. Analysis of informed consent as the legal protection of physician relationships and patients in malpractice cases. POLRI J Law Policy. 2023;2(4):327-338. doi:10.55047/polri.v2i4.774
5. Budi RC, Wibowo A. A Well-understood Surgical Informed Consent : A Scoping Review. Unnes J Public Health. 2023;12(1):21-32. doi:10.15294/ujph.v12i1.57731
6. Pietrzykowski T, Smilowska K. The reality of informed consent: empirical studies on patient comprehension—systematic review. Trials. 2021;22(1):57. doi:10.1186/s13063-020-04969-w
7. Ruske J, Sharma G, Makie K, et al. Patient comprehension necessary for informed consent for vascular procedures is poor and related to frailty. J Vasc Surg. 2021;73(4):1422-1428. doi:10.1016/j.jvs.2020.06.131
8. Long KL, Ingraham AM, Wendt EM, et al. Informed Consent and Informed Decision-Making in High-Risk Surgery: A Quantitative Analysis. J Am Coll Surg. 2021;233(3):337-345. doi:10.1016/j.jamcollsurg.2021.05.029
9. Basukala S, Shrestha O, Thapa N, et al. How informed is informed consent?—Evaluating the quality of informed consent among surgical patients in a tertiary care hospital in Nepal. PLoS One. 2023;18(7):e0288074. doi:10.1371/journal.pone.0288074
10. Gebreanenia FA, Kahsay HB, Belay DS, et al. Surgical informed consent practice and associated factors among adult postoperative patients in public hospitals of Mekelle, Tigray, Ethiopia 2023/2024. Cross sectional. PLoS One. 2025;20(2):e0310920. doi:10.1371/journal.pone.0310920
11. Ayele TT, Negash TT, Oumer KE, et al. Patients’ satisfaction and associated factors towards preoperative informed consent process: A cross-sectional study. Ann Med Surg. 2022;79:104104. doi:10.1016/j.amsu.2022.104104
12. Kebede BF, Tesfa TB, Hiwot AY, et al. Knowledge of surgical informed consent and associated factors among patients undergone obstetric and gynecologic surgery at Jimma Medical Center, Jimma, Ethiopia, 2020: an institutional based cross-sectional study. Perioper Med. 2023;12(1):18. doi:10.1186/s13741-023-00295-2
13. Valente EP, Mariani I, Covi B, et al. Quality of Informed Consent Practices around the Time of Childbirth: A Cross-Sectional Study in Italy. Int J Environ Res Public Health. 2022;19(12):7166. doi:10.3390/ijerph19127166
14. Molla MT, Bizuneh YB, Nigatu YA, et al. Evaluate the practice of preoperative informed consent for elective surgical patients at the university hospital, Ethiopia, in 2022. Int J Surg Open. 2022;47:100546. doi:10.1016/j.ijso.2022.100546
15. Mondal S, Bhattacharya S, Jana PK, et al. A Comparative Study on Comprehension of Informed Consent Before Emergency and Elective Surgical Operative Procedures. Acta Med Int. 2022;9(2):153-158. doi:10.4103/amit.amit_90_22
16. Michael S, Kyejo W, Ismail A, et al. Patients’ perceptions and understanding of preoperative informed consent: A qualitative thematic analysis from Tanzania. SAGE Open Med. 2025;13:20503121251331122. doi:10.1177/20503121251331122
17. Williams E, Selvam R, Hopman W, et al. Are patients truly informed? A retrospective chart review of the documentation of informed consent in laparoscopic cholecystectomy. Can J Surg. 2021;64(4):E435-E441. doi:10.1503/cjs.004120
18. Atkin J, Incoll IW, Owen J, et al. Informed consent: perceptions and practice of orthopaedic trainees. ANZ J Surg. 2022;92(4):819-824. doi:10.1111/ans.17517
19. Kumru S, Yiğit P, Demirtaş M, et al. Informed consent in surgical practice with patients’ experiences: A cross-sectional study. Patient Exp J. 2023;10(3):42-48. doi:10.35680/2372-0247.1768
20. Bär E, Bock T, Bock B, et al. Limitations of informed consent in routine clinical practice - An observational survey study to characterize patient populations with limited understanding of the routine informed consent process in an anaesthesia clinic. BMC Anesthesiol. 2025;25(1):263. doi:10.1186/s12871-025-03275-9
21. Zevin B, Almakky M, Mancini U, et al. Digital approach to informed consent in bariatric surgery: a randomized controlled trial. Surg Endosc. 2021;36(1):809-816. doi:10.1007/s00464-020-08277-x
22. Haack M, Fischer ND, Frey L, et al. Digital informed consent for urological surgery - randomized controlled study comparing multimedia-supported vs. traditional paper-based informed consent concerning satisfaction, anxiety, information gain and time efficiency. Prostate Cancer Prostatic Dis. 2023;27(4):715-719. doi:10.1038/s41391-023-00737-4
23. Abujarad F, Peduzzi P, Mun S, et al. Comparing a Multimedia Digital Informed Consent Tool With Traditional Paper-Based Methods: Randomized Controlled Trial. JMIR Form Res. 2021;5(10):e20458. doi:10.2196/20458
24. Aslam AM, Kennedy J, Seghol H, et al. A randomized controlled trial of patient recall after detailed written consent versus standard verbal consent in adults with routine orthopaedic trauma. Bone Jt Open. 2023;4(2):104-109. doi:10.1302/2633-1462.42.bjo-2022-0148.r1
25. Osterloh J, Müller-Seubert W, Cai A, et al. Is there an impact of a video-based patient informed consent in elective hand surgery?. Arch Orthop Trauma Surg. 2024;144(6):2899-2903. doi:10.1007/s00402-024-05291-9
26. Liu Z, Yuan L, Rong W, et al. The effects of written versus video aids on informed consent across different types of surgeries: A randomized controlled study. Curr Probl Surg. 2025;67:101747. doi:10.1016/j.cpsurg.2025.101747
27. Seely KD, Higgs JA, Butts L, et al. The “teach-back” method improves surgical informed consent and shared decision-making: a proof of concept study. Patient Saf Surg. 2022;16(1):12. doi:10.1186/s13037-022-00342-9
28. Goldschmitt M, Gleim P, Mandelartz S, et al. Digitalizing informed consent in healthcare: a scoping review. BMC Health Serv Res. 2025;25(1):776. doi:10.1186/s12913-025-12964-7
29. Bothun LS, Feeder SE, Poland GA. Readability of Participant Informed Consent Forms and Informational Documents. Mayo Clin Proc. 2021;96(8):2095-2101. doi:10.1016/j.mayocp.2021.05.025
30. García-Álvarez JM, García-Sánchez A. Readability of Informed Consent Forms for Medical and Surgical Clinical Procedures: A Systematic Review. Clin Pract. 2025;15(2):26. doi:10.3390/clinpract15020026

31. García-Álvarez JM, García-Sánchez A. Readability and Comprehension of Anesthesia Informed Consent Forms in a Spanish County Hospital: An Observational Study. Nurs Rep. 2024;14(2):1341-1354. doi:10.3390/nursrep14020101
32. Kamath YV, Shetty YC, Lanjewar IC, et al. Readability of informed consent documents and its impact on consent refusal rate. Perspect Clin Res. 2024;16(1):38-43. doi:10.4103/picr.picr_322_23
33. Trutner ZD, Furlough K, Martinez A, et al. Is Health Literacy Associated With Surgical Outcomes? A Systematic Review. J Surg Res. 2023;291:720-733. doi:10.1016/j.jss.2023.06.044
34. Pedersen SE, Aaby A, Friis K, et al. Multimorbidity and health literacy: A population-based survey among 28,627 Danish adults. Scand J Public Health. 2021;51(2):165-172. doi:10.1177/14034948211045921
35. Shao Y, Hu H, Liang Y, et al. Health literacy interventions among patients with chronic diseases: A meta-analysis of randomized controlled trials. Patient Educ Couns. 2023;114:107829. doi:10.1016/j.pec.2023.107829
36. Toapanta N, Salas-Gama K, Pantoja PE, et al. The role of low health literacy in shared treatment decision-making in patients with kidney failure. Clin Kidney J. 2023;16(Supplement_1):i4-i11. doi:10.1093/ckj/sfad061
37. Pittalis C, Sackey C, Okeny P, et al. Surgical informed consent practices and influencing factors in sub-Saharan Africa: a scoping review of the literature. BMJ Qual Saf. 2023;33(10):653-662. doi:10.1136/bmjqs-2023-016823
38. Nsaful J, Attah D, Sefogah PE, et al. The informed consent process for surgical procedures: perspectives of patients and next-of-kins in a lower middle-income country. BMC Med Ethics. 2026;27(1):14. doi:10.1186/s12910-026-01407-5
39. Lim MYH. Patient autonomy in an East-Asian cultural milieu: a critique of the individualism-collectivism model. J Med Ethics. 2022;50(9):640-642. doi:10.1136/medethics-2022-108123
40. Akhtar MU, Bhatti ME, Fredericks S. What factors influence patient autonomy in healthcare decision-making? A systematic review of studies from the Global South. Nurs Ethics. 2024;32(3):875-891. doi:10.1177/09697330241272794
41. Sylvestre P, Gaucher NO, Perez T, et al. Rethinking informed consent in the age of behavioural sciences and relational autonomy. Ethics Med Public Health. 2021;19:100708. doi:10.1016/j.jemep.2021.100708
42. Lewis J. Re-envisioning autonomy: From consent and cognitive capacity to embodied, relational, and authentic selfhood. Clin Ethics. 2025;20(1):14777509251324841. doi:10.1177/14777509251324841
43. Zhai G. Improving Chinese patients' autonomy in medical decision-making through policy frameworks. Front Public Health. 2025;13:1553721. doi:10.3389/fpubh.2025.1553721
44. Wahyuningsih CS, Asih EB, Masitoh R, et al. Patient Safety Culture in Hospitals based on Agency For Health Care Research And Quality (AHRQ). J Health Policy Manage. 2024;09(02):197-210. doi:10.26911/thejhpm.2024.09.02.06
45. Yang A, Yang K. A Study on Reproducibility and the Reliability of the Hosmer-Lemeshow Test in Published Research. N Engl J Stat Data Sci. 2025:1-14. doi:10.51387/25-nejsds81
46. Ghaferi AA, Schwartz TA, Pawlik TM. STROBE Reporting Guidelines for Observational Studies. JAMA Surg. 2021;156(6):577-578. doi:10.1001/jamasurg.2021.0528
47. Jeste DV, Palmer BW, Appelbaum PS, et al. A New Brief Instrument for Assessing Decisional Capacity for Clinical Research. Arch Gen Psychiatry. 2007;64(8):966. doi:10.1001/archpsyc.64.8.966
48. Weiss BD. Quick Assessment of Literacy in Primary Care: The Newest Vital Sign. Ann Fam Med. 2005;3(6):514-522. doi:10.1370/afm.405
49. Triandis HC, Gelfand MJ. Converging measurement of horizontal and vertical individualism and collectivism. J Pers Soc Psychol. 1998;74(1):118-128. doi:10.1037/0022-3514.74.1.118
50. Zhou QM, Zhe L, Brooke RJ, et al. A relationship between the incremental values of area under the ROC curve and of area under the precision-recall curve. Diagn Progn Res. 2021;5(1):13. doi:10.1186/s41512-021-00102-w
51. Pang J, Ju W, Welch M, et al. Nonparametric bootstrap methods for interval estimation of the area under the ROC curve with correlated diagnostic test data: application to whole-virus ELISA testing in swine. Front Vet Sci. 2023;10:1274786. doi:10.3389/fvets.2023.1274786
52. Stolte M, Herbrandt S, Ligges U. A comprehensive review of bias reduction methods for logistic regression. Stat Surv. 2024;18(none):1-51. doi:10.1214/24-ss148
53. Guan Y, Fu GH. A Double-Penalized Estimator to Combat Separation and Multicollinearity in Logistic Regression. Mathematics. 2022;10(20):3824. doi:10.3390/math10203824
54. Moss J, Grønneberg S. Partial Identification of Latent Correlations with Ordinal Data. Psychometrika. 2023;88(1):241-266. doi:10.1007/s11336-022-09898-y
55. Gaster T, Eggertsen CM, Støvring H, et al. Quantifying the impact of unmeasured confounding in observational studies with the E value. BMJ Med. 2023;2(1):e000366. doi:10.1136/bmjmed-2022-000366
56. Placer-Galán C, Arocena EB. Fragility Index, Reverse Fragility Index, Fragility Quotients, and Susceptibility Index. Novel tools for assessing of randomized clinical trials. Cir Esp. 2024;102(1):41-43. doi:10.1016/j.ciresp.2023.05.012
57. Grimes DR. The ellipse of insignificance, a refined fragility index for ascertaining robustness of results in dichotomous outcome trials. eLife. 2022;11:e79573. doi:10.7554/elife.79573
58. Magowan D, Evans M, Burton L, et al. Medicolegal claims in general surgery: a 10-year retrospective review of claims against the NHS in England. Ann R Coll Surg Engl. 2023;105(7):664-671. doi:10.1308/rcsann.2022.0121
59. Hansrani V, Ghatwary Tantawy T, Al-Khaffaf H. Clinical Negligence Claims Against Vascular Surgery in the United Kingdom: An Observational Study. Ann Vasc Surg. 2021;70:549-554. doi:10.1016/j.avsg.2020.09.020
60. Øhrn FD, Årøen A, Aae TF. Medical negligence compensation claims in knee meniscal surgery in Norway: a cross-sectional study. Patient Saf Surg. 2025;19(1):9. doi:10.1186/s13037-025-00427-1
61. Taylor MA, Yonash RA. Risk Factors for Wrong-Site Surgery: A Study of 1,166 Reports of Informed Consent and Schedule Errors. Patient Saf. 2024;6(1). doi:10.33940/001c.117084