Statistical Inference and Natural Language Analysis of Consent Document Ethics in Clinical Trials

Francis Osei[1*]

[1*]AI-Driven Clinical Innovation & Regulatory Science, Bayezian Limited, 11/13 Weston Street, London, SE1 3ER, United Kingdom.

Corresponding author(s). E-mail(s): francis@bayezian.com;

Abstract

This study evaluates the ethical and linguistic quality of informed consent forms (ICFs) from interventional clinical trials, using natural language processing and logistic regression to analyse their association with trial outcomes. Drawing from a curated corpus of English-language ICFs, we tested whether semantic coherence, ethical clause consistency, and structural variation correlated with trial completion or termination. Results showed that completed trials had significantly higher semantic and ethical consistency, while structural variation (edit distance) was not predictive. Moreover, a higher clause count was inversely associated with trial success, suggesting that excess complexity may reduce comprehension. These findings raise normative concerns about current consent practices and point to the need for clearer, ethically structured communication. The study recommends that ethics committees and sponsors adopt semantic and ethical benchmarks in ICF evaluation and tailor materials for clarity and accessibility. It also highlights the value and limits of computational methods in assessing the moral adequacy of consent.

Keywords: Informed consent, Clinical trial ethics, Natural language processing, Ethical clause richness, Statistical inference

1 Introduction

Informed consent forms (ICFs) play a critical ethical role in upholding participant autonomy and ensuring that enrolment in clinical trials is voluntary, comprehensible, and transparent Ssali et al. (2017). However, despite the foundational importance of informed consent in biomedical research, longstanding concerns persist regarding its efficacy, particularly in contexts characterised by linguistic complexity, regulatory standardisation, and limited participant engagement Grant (2021); Kadam (2017). Research has repeatedly shown that ICFs often exceed recommended readability thresholds, employ specialised legal or technical language, and prioritise institutional liability over participant comprehension, thus undermining the moral purpose of consent and compromising trust Manti and Licari (2018); Rebers et al. (2016).

This ethical shortfall is not geographically constrained. Across both low- and high-resource research settings, individuals may offer formal consent without adequate understanding of study protocols, associated risks, or potential benefits Bhupathi and Ravi (2017); Ssali et al. (2017). Such deficiencies expose a critical tension between regulatory adherence and meaningful communication, as consent documents continue to be designed more for procedural validation than for participant empowerment. Although international frameworks such as the Declaration of Helsinki and the Belmont Report affirm that informed understanding is central to ethical research, the implementation of this principle remains inconsistent and, in many cases, unexamined in practice Rebers et al. (2016).

While existing scholarship has extensively critiqued the readability and ethical content of ICFs through qualitative analysis, a persistent gap remains. No large-scale, empirical investigation has systematically evaluated whether specific linguistic and structural features of consent documentation correlate with tangible research outcomes, such as clinical trial termination. Prior studies have highlighted issues related to literacy, cultural divergence, and ethical oversight Bhupathi and Ravi (2017); Fons-Martinez et al. (2022); Ssali et al. (2017), but few have offered a quantitative model that connects textual features of consent forms to trial success or failure.

This study employs natural language processing (NLP) and statistical inference to analyse informed consent forms (ICFs) from interventional trials classified as either completed or terminated. It investigates whether specific linguistic and ethical characteristics of consent documents are associated with trial outcomes, using semantic similarity (sentence embeddings), ethical clause richness (as a proxy for disclosure quality), and structural variation (edit distance) as key variables. Logistic regression models assess their predictive value.

Three hypotheses structure the inquiry: Hypothesis 1 (H1) posits that completed trials exhibit greater semantic coherence across ICFs; Hypothesis 2 (H2) anticipates higher ethical clause consistency; and Hypothesis 3 (H3) tests whether structural variation independently correlates with trial status.

By advancing from descriptive critique to empirical assessment, this study offers a computational approach to evaluating the ethical integrity of consent documentation. It underscores the role of linguistic clarity and ethical transparency in fostering participant-centred research and enhancing trial validity.

2 Method

2.1 Study Design and Data Source

This study employed a retrospective cross-sectional design. Data were drawn from the U.S. National Library of Medicine (2025), a publicly accessible registry managed by the United States National Library of Medicine. Trials were selected based on predefined eligibility criteria, and full-text informed consent forms (ICFs) were obtained directly from the registry or through linked external repositories. All documents were reviewed to confirm their completeness and relevance to the study objectives.

2.2 Trial Selection Criteria

Trials were included if they satisfied the following conditions: industry sponsorship; designation as Phase II, III, or IV; recruitment status listed as either "Completed" or "Terminated"; and availability of a complete ICF. Phase I studies, as well as trials that were suspended, withdrawn, or lacking a full-text consent form, were excluded. This focus on later-phase, industry-led studies ensured a consistent regulatory context and enhanced the relevance of the analysis to standardised clinical research environments.

2.3 Text Processing

ICFs were extracted from PDF format and subjected to text cleaning and normalisation. Semantic embeddings were generated using the Sentence-BERT model (all-MiniLM-L6-v2). Three measures of textual similarity were applied: cosine similarity (semantic closeness), Jaccard similarity (binary clause overlap), and edit distance (token-level structural variation).

2.4 Ethical Clause Scoring

The presence of ethical clauses was assessed using predefined regular expression patterns covering domains such as study purpose, risks, confidentiality, voluntary participation, and withdrawal rights. Clause detection was carried out through automated pattern matching, supplemented by manual verification to ensure the accurate identification of core ethical disclosures.

2.5 Statistical Analysis

Descriptive statistics were computed for document length, similarity metrics, and ethical clause counts. Mann-Whitney U tests were used to compare linguistic features between completed and terminated trials. Logistic regression was employed to model the likelihood of trial termination, with model evaluation based on the Akaike Information Criterion (AIC), McFadden's R^2, and the Hosmer–Lemeshow test for goodness-of-fit.

2.6 Software and Tools

Data extraction, text processing, and document similarity calculations were performed using Python (version 3.12.1). Statistical analyses were conducted in R (version 4.3.2). All scripts, software environment details, and reproducibility materials are available at the public GitHub repository: https://github.com/Franosei/Consent-Document-Ethics-in-Clinical-Trials.

3 Results

3.1 Sample Characteristics

The final dataset included 85 clinical trials, of which 34 were categorised as terminated and 51 as completed. These trials spanned four broad therapeutic areas: oncology, infectious diseases, cardiology, and other therapeutic categories. Oncology accounted for the largest proportion of trials, making up just over half of the sample. Trials were excluded if their informed consent forms (ICFs) were incomplete, defective, or provided only as scanned images that could not be processed for text extraction.

Table 1: Distribution of Clinical Trials by Trial Status and Therapeutic Area

Trial Status	Oncology	Infectious	Cardiology	Other Therapeutics	Total
Terminated	16	3	2	13	34
Completed	27	9	8	7	51

3.2 Linguistic Characteristics and Similarity Analysis

This section compares the language and structure of informed consent forms (ICFs) from completed and terminated clinical trials. Three aspects were examined: semantic similarity, presence of ethical clauses, and document readability.

3.2.1 Semantic Similarity

ICFs from completed trials showed a higher average semantic similarity ($M = 0.607$) than those from terminated trials ($M = 0.539$), indicating that the language used in the former was more consistent across documents. Visual comparisons, including boxplots, suggested that completed trials tended to produce more uniform documentation. A Mann-Whitney U test confirmed that this difference was statistically significant ($p = 0.0000173$).

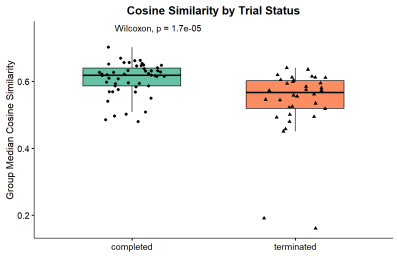
3.2.2 Ethical Clause Coverage

The analysis found that completed trials also included a greater number of ethical clauses in their consent documents. These clauses addressed core areas such as risks,

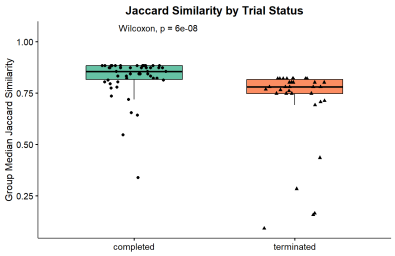
confidentiality, and voluntary participation. The higher frequency of such clauses suggests a stronger emphasis on participant rights and responsibilities in trials that were successfully completed. The difference in clause coverage between the two groups was highly significant ($p = 0.0000000598$).

3.2.3 Readability

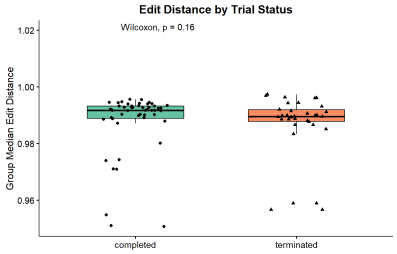
An assessment of readability, using Flesch-Kincaid scores, showed little difference between the two trial groups. In both cases, the language used in the consent forms exceeded the commonly recommended reading level for the general public. This finding highlights an ongoing need to improve the clarity and accessibility of research documentation, regardless of trial outcome.



(a) Cosine Similarity (H1)



(b) Jaccard Similarity (H2)



(c) Edit Distance (H3)

Fig. 1: Group comparisons of linguistic characteristics in informed consent forms (ICFs) by trial outcome.

3.3 Statistical Test Results

To examine differences in document characteristics between completed and terminated trials, statistical tests were conducted for each of the three main hypotheses. The results are summarised in Table 2. As shown, completed trials demonstrated significantly higher semantic similarity and greater consistency in ethical clause coverage.

However, structural differences, measured by edit distance, did not vary significantly between the two groups.

Table 2: Statistical Test Results for Group Differences in Linguistic Characteristics

Hypothesis	Variable	p-value	Interpretation
H1	`group_mean_cosine`	0.0000173	Completed trials have significantly higher semantic similarity
H2	`group_mean_jaccard`	0.0000000598	Completed trials have significantly higher ethical clause overlap
H3	`group_mean_edit_dist`	0.165	No significant difference in structural edit distance

3.4 Regression Results

To further explore the relationship between linguistic features of consent forms and clinical trial outcomes, a logistic regression analysis was conducted. The model incorporated four predictors aligned with the study's hypotheses: semantic coherence (group mean cosine similarity), ethical clause consistency (group mean Jaccard similarity), structural variation (group mean edit distance), and overall document richness (clause count).

The model performed well, explaining approximately 48.4% of the variance in trial status, as indicated by McFadden's pseudo-R^2 value of 0.484. The Akaike Information Criterion (AIC) was 68.55, and the Hosmer–Lemeshow test was non-significant ($\chi^2(8) = 5.16$, $p = 0.741$), suggesting good model calibration.

As presented in Table 3, both semantic similarity and ethical clause consistency were significantly associated with trial completion, supporting Hypotheses 1 and 2. In contrast, edit distance did not show a statistically significant effect, offering no support for Hypothesis 3. Interestingly, a higher total number of clauses, potentially indicative of increased document complexity was associated with a lower likelihood of trial success.

Table 3: Logistic Regression Results Predicting Trial Completion from Linguistic Characteristics

Predictor	Estimate	SE	z-value	p-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept	-71.32	34.49	-2.07	0.0387	1.07×10^{-31}	1.25×10^{-62}	3.34×10^{-2}
H1	25.30	8.15	3.10	0.0019	9.67×10^{10}	6.21×10^{4}	9.68×10^{18}
H2	34.68	8.40	4.13	<0.0001	1.16×10^{15}	7.37×10^{8}	2.61×10^{23}
H3	51.06	32.24	1.58	0.1133	1.50×10^{22}	6.34×10^{-7}	2.47×10^{50}
Clause Count	-1.75	0.45	-3.85	<0.0001	0.17	0.06	0.38

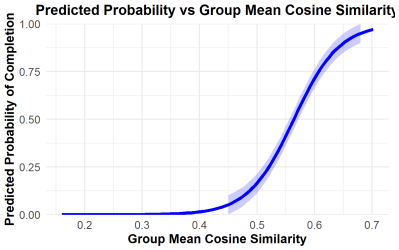
3.5 Predicted Probability of Trial Completion Across Document Similarity Metrics

Logistic regression predictions were used to explore how document characteristics relate to the likelihood of trial completion. Figures were generated for each key predictor corresponding to the study's hypotheses.

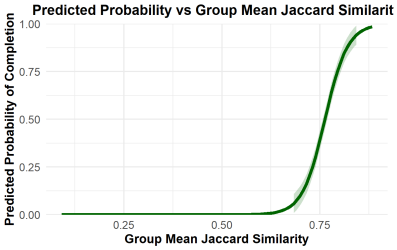
Group Mean Cosine Similarity (H1): The probability of trial completion increased sharply with higher cosine similarity, following a sigmoidal pattern with an inflection point between 0.55 and 0.60. This supports the hypothesis that greater linguistic consistency across ICFs is associated with improved trial success.

Group Mean Jaccard Similarity (H2): A similar trend was observed for Jaccard similarity. The probability of completion rose markedly between values of 0.70 and 0.80, suggesting that consistent inclusion of ethical clauses may contribute to better trial outcomes.

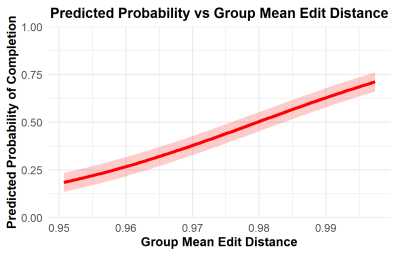
Group Mean Edit Distance (H3): In contrast, edit distance showed a flat, near-linear relationship with trial outcome, indicating limited predictive value. This aligns with the statistical results and provides no strong support for Hypothesis 3.



(a) Cosine Similarity (H1)



(b) Jaccard Similarity (H2)



(c) Edit Distance (H3)

Fig. 2: Predicted Probability of Trial Completion Across Document Similarity Metrics

3.5.1 Odds Ratios and Confidence Intervals for Predicting Trial Completion

Odds ratios and 95% confidence intervals were computed to assess the strength and direction of association between each predictor and trial completion (Table 4). Cosine and Jaccard similarity were strongly associated with trial completion, with confidence intervals well above 1. Edit distance showed no reliable association, with a wide confidence interval spanning both directions. Clause count was inversely associated with trial success.

Table 4: Odds Ratios and 95% Confidence Intervals for Predictors of Clinical Trial Completion

Predictor	**Odds Ratio**	**95% CI Lower**	**95% CI Upper**
Intercept	1.07×10^{-31}	1.25×10^{-62}	3.34×10^{-2}
Group Mean Cosine (H1)	9.67×10^{10}	6.21×10^{4}	9.67×10^{18}
Group Mean Jaccard (H2)	1.16×10^{15}	7.37×10^{8}	2.61×10^{23}
Group Mean Edit Distance (H3)	1.50×10^{22}	6.34×10^{-7}	2.47×10^{50}
Clause Count	0.17	0.06	0.38

4 Discussion

4.1 Principal Findings

This study identified significant associations between the linguistic and ethical characteristics of informed consent forms (ICFs) and clinical trial outcomes. Specifically, consent documents from completed trials exhibited greater semantic coherence and more consistent ethical clause inclusion than those from terminated trials, as evidenced by statistically significant differences in both cosine and Jaccard similarity scores (Table 2; Figure 2). These findings support Hypothesis 1 (H1) and Hypothesis 2 (H2), indicating that linguistic uniformity and ethical transparency contribute meaningfully to the likelihood of trial success.

In contrast, structural variation, as measured by edit distance, did not significantly differ between the two trial groups and showed limited predictive value in logistic regression modelling (Table 3). This result does not support Hypothesis 3 (H3) and suggests that surface-level structural consistency is not, in itself, a reliable indicator of ethical communication or operational viability.

Moreover, a higher overall clause count, often assumed to reflect comprehensiveness, was inversely associated with trial completion, suggesting that excessive document complexity may hinder rather than help participant engagement. This counterintuitive outcome highlights the importance of prioritising clarity and relevance over volume in ethical disclosures (Table 4).

Taken together, the results reinforce the hypothesis that ethically robust and linguistically coherent consent documents not only reflect better communicative intent but also predict stronger operational outcomes in clinical research contexts.

4.2 Ethical Interpretation and Implications

The empirical associations identified in this study raise substantive ethical questions about the adequacy and moral function of current informed consent practices. The finding that higher semantic similarity correlates with trial completion suggests that linguistic coherence is not merely a stylistic concern but a structural enabler of ethical communication. Consent documents that maintain internal consistency are more likely to support participant understanding, thereby fulfilling the ethical obligation to secure genuinely informed consent.

Similarly, the significant predictive power of ethical clause richness reinforces the normative claim that informed consent is not ethically valid unless core disclosures, risks, voluntariness, confidentiality, and withdrawal rights are meaningfully conveyed. The absence or inconsistent presence of such clauses undermines participant autonomy and violates principles articulated in international ethics guidelines such as the Declaration of Helsinki and the Belmont Report.

Conversely, the inverse association between clause count and trial success challenges the bureaucratic assumption that more content equates to better ethical practice. When consent documents become dense compilations of legal and procedural jargon, they risk shifting from instruments of participant empowerment to mechanisms of institutional self-protection. This reflects a broader concern that procedural compliance has displaced ethical intent in many research contexts.

These dynamics are particularly troubling in low-literacy or marginalised populations, where cognitive and linguistic burdens disproportionately affect participants' ability to engage with trial materials. Ethically, this constitutes a form of epistemic injustice, where individuals are excluded from full participatory agency due to inaccessible communication. The results, therefore, call for a re-evaluation of consent form design through the lens of justice in communication: one that respects diversity in comprehension while upholding universal standards of ethical clarity.

4.3 Comparison with Previous Research

The results of this study are consistent with longstanding ethical concerns surrounding informed consent and extend several strands of earlier empirical and normative work. Numerous studies have criticised ICFs for their complexity and poor accessibility. Grant (2021) and Kadam (2017) both emphasised that consent forms are often unreadable for non-specialist audiences, while Manti and Licari (2018) showed that excessive legal language impedes comprehension. The finding that semantic coherence correlates with trial completion reinforces these concerns by quantifying the operational consequences of incoherent consent language.

Clause-level analysis builds on findings from Ssali et al. (2017), who noted that the formal presence of consent does not always equate to informed understanding, especially in low-resource settings. Bhupathi and Ravi (2017) similarly observed that procedural

adherence can obscure ethical communication, a theme also reflected in Rebers et al. (2016), who highlighted the gap between ethical guidelines and real-world implementation.
This study also complements efforts to improve clause content and standardisation, such as those outlined in the i-CONSENT project Fons-Martinez et al. (2022), the FDA guidance Geetter and Siegfried (2023), and the European Medicines Agency recommendations Enpr-EMA Working Group on Ethics (2021). However, our findings caution against over-standardisation, as a higher number of clauses was inversely related to trial success, echoing the concerns raised by ACRP Association of Clinical Research Professionals (ACRP) (2013), Coleman et al. (2021), and Health Research Authority (HRA) (2018) about cognitive overload in consent forms.
The underreporting of ethical elements, identified in studies by Trung et al. (2017) and Koonrungsesomboon et al. (2015), is echoed in our data, where many terminated trials lacked consistent disclosure of core ethical domains. Moreover, this study aligns with Schwarz (2025) and International Conference on Harmonisation (ICH) (2025) in recognising the importance of tailoring consent processes to participant needs, but highlights the need for empirical validation of those adaptations.
By quantitatively linking document quality to trial outcomes, this study bridges normative critique and operational evidence, offering a concrete framework to evaluate consent adequacy in clinical research.

4.4 Methodological and Ethical Limitations

This study has several limitations. First, it relies on publicly available ICFs, which may not represent the broader population of trials. Selection bias is likely, as incomplete or inaccessible documents were excluded. Second, all forms analysed were in English, limiting applicability to multilingual or non-English-speaking contexts where consent practices may differ substantially.
Third, local adaptations, verbal consent processes, and informal communication, key elements of ethical engagement, were not captured in the textual data. Fourth, although statistically informative, the sample size constrains detailed subgroup analysis by therapeutic area or sponsor type. Lastly, the observational design permits association but not causality; findings should be interpreted as indicative rather than determinative.
Beyond methodological constraints, computational models cannot fully account for contextual nuance or participant experience. While semantic and structural metrics offer valuable proxies, they cannot replace direct engagement with the lived realities of informed consent, especially in vulnerable populations.

4.5 Recommendations for Ethics Committees and Sponsors

The findings suggest several actionable steps to improve the ethical and operational quality of informed consent. Ethics committees should move beyond checklist-based review and evaluate ICFs for semantic coherence and ethical clause sufficiency. Clearer, standardised disclosure requirements focusing on purpose, risks, rights, and withdrawal should be mandated to ensure ethical parity across studies.

Sponsors should invest in linguistic review and usability testing of consent documents. Readability alone is insufficient; documents must be internally consistent and ethically legible. Cultural and linguistic tailoring is especially critical in multilingual or low-literacy settings. ICFs should be adapted to reflect participant norms without compromising ethical substance.
Finally, regulatory frameworks should encourage empirical evaluation of consent quality, not just procedural compliance. Embedding semantic and ethical benchmarks in ICF review would strengthen both participant protection and trial viability.

4.6 Future Research Directions

Further research should evaluate consent practices across linguistic and cultural contexts to assess whether the associations observed here hold in non-English ICFs. Comparative studies using translated corpora or multilingual NLP models may help uncover structural barriers to understanding in global trial settings.
Participant-centred approaches, such as feedback surveys, comprehension testing, or ethnographic interviews should complement text-based analysis to better capture the lived experience of consent. These methods would provide essential context that computational tools alone cannot.
In addition, prospective studies that integrate semantic and ethical metrics into ICF drafting and ethics review could test whether document optimisation improves participant comprehension or retention. However, any future work must remain sensitive to the ethical limits of automation in domains requiring human trust, judgment, and care.

5 Conclusion

This study demonstrates that informed consent forms with higher semantic coherence and consistent ethical clause inclusion are more likely to be associated with successful clinical trial completion. Structural variation alone was not predictive, and excessive clause counts may reduce clarity rather than enhance it. These findings reinforce that consent documents are not just regulatory instruments but ethical tools that must communicate clearly, transparently, and accessibly. Evaluating ICFs through both linguistic and ethical lenses offers a practical path toward improving participant engagement and strengthening the integrity of research design.

Declarations

Funding: This study was conducted as part of Bayezian Limited's ongoing work applying artificial intelligence and statistical modelling to improve clinical trial design and outcomes. No external funding was received.

Conflict of interest/Competing interests: Not applicable.

Ethics approval and consent to participate: Not applicable. This study analysed publicly available, anonymised documents and did not involve human subjects

or identifiable personal data.

Consent for publication: Not applicable.

Data availability: The original informed consent forms used in this study were retrieved from publicly accessible clinical trial registries. A processed version of the dataset is available at: https://github.com/Franosei/Consent-Document-Ethics-in-Clinical-Trials

Materials availability: Not applicable.

Code availability: All analysis scripts and code used in this study are publicly available at: https://github.com/Franosei/Consent-Document-Ethics-in-Clinical-Trials

Author contribution: Francis Osei conceived the study, collected and analysed the data, interpreted the results, and wrote the manuscript.

References

Association of Clinical Research Professionals (ACRP). 2013. The process of informed consent. ACRP White Paper.

Bhupathi, P.A. and G.R. Ravi. 2017. Comprehensive format of informed consent in research and practice: A tool to uphold the ethical and moral standards. *International Journal of Clinical Pediatric Dentistry 10*(1): 73–81. https://doi.org/10.5005/jp-journals-10005-1411 .

Coleman, E., L. O'Sullivan, R. Crowley, M. Hanbidge, S. Driver, T. Kroll, A. Kelly, A. Nichol, O. McCarthy, P. Sukumar, and P. Doran. 2021. Preparing accessible and understandable clinical research participant information leaflets and consent forms: A set of guidelines from an expert consensus conference. *Research Involvement and Engagement* 7: 31. https://doi.org/10.1186/s40900-021-00265-2 .

Enpr-EMA Working Group on Ethics. 2021. Assent / informed consent guidance for paediatric clinical trials with medicinal products in Europe. European Medicines Agency (EMA), Document No. EMA/671028/2019.

Fons-Martinez, J., C. Ferrer-Albero, and J. Diez-Domingo. 2022. Keys to improving the informed consent process in research: Highlights of the i-CONSENT project. *Health Expectations 25*(3): 1183–1185. https://doi.org/10.1111/hex.13427 .

Geetter, J.S. and S. Siegfried. 2023. Fda issues final guidance on informed consent in clinical investigations. McDermott Will & Emery. Retrieved from https://www.fda.gov/media/88915/download.

Grant, S.C. 2021. Informed consent—we can and should do better. *JAMA Network Open 4*(4): e2110848. https://doi.org/10.1001/jamanetworkopen.2021.10848 .

Health Research Authority (HRA). 2018. Applying a proportionate approach to the process of seeking consent: HRA guidance. Version 1.02. Available from http://www.hra.nhs.uk/resources/before-you-apply/consent-and-participant-information/.

International Conference on Harmonisation (ICH). 2025. Informed consent of trial subjects (ich-gcp). https://ichgcp.net/publications.

Kadam, R.A. 2017. Informed consent process: A step further towards making it meaningful. *Perspectives in Clinical Research 8*(3): 107–112 .

Koonrungsesomboon, N., J. Laothavorn, and J. Karbwang. 2015. Understanding of essential elements required in informed consent form among researchers and institutional review board members. *Tropical Medicine and Health 43*(2): 117–122. https://doi.org/10.2149/tmh.2014-36 .

Manti, S. and A. Licari. 2018. How to obtain informed consent for research. *Breathe 14*(2): 145–152. https://doi.org/10.1183/20734735.001918 .

Rebers, S., N.K. Aaronson, F.E. van Leeuwen, and M.K. Schmidt. 2016. Exceptions to the rule of informed consent for research with an intervention. *BMC Medical Ethics 17*(9). https://doi.org/10.1186/s12910-016-0092-6 .

Schwarz, G. 2025. Informed consent considerations. Slides presented at the ACT-EU Workshop, European Medicines Agency.

Ssali, A., F. Poland, and J. Seeley. 2017. Exploring informed consent in HIV clinical trials: A case study in Uganda. *Heliyon 3*(2): e00196. https://doi.org/10.1016/j.heliyon.2016.e00196 .

Trung, L.Q., M.E. Morra, N.D. Truong, T. Turk, A. Elshafie, A. Foly, D.N.H. Tam, A. Iraqi, T.T.H. Van, A. Elgebaly, T.N. Ngoc, T.L.H. Vu, N.T. Chu, K. Hirayama, J. Karbwang, and N.T. Huy. 2017. A systematic review finds underreporting of ethics approval, informed consent, and incentives in clinical trials. *Journal of Clinical Epidemiology* 92: 1–7. https://doi.org/10.1016/j.jclinepi.2017.08.007 .

U.S. National Library of Medicine. 2025. Clinicaltrials.gov. https://clinicaltrials.gov/.