STROBE Statement — Checklist of Items Included in the Report

Designing Out Failure: A Counterfactual Simulation Framework for Reducing Structural Fragility in Clinical Trials

Study design: Retrospective counterfactual simulation using observational registry data (ClinicalTrials.gov / AACT database)

Section / Topic	Item No.	Recommendation	Location in Manuscript
Title and Abstract	**1a**	Indicate the study's design with a commonly used term in the title or the abstract	Title states 'Counterfactual Simulation Framework'; Abstract Methods states 'retrospective counterfactual simulation study'
	1b	Provide in the abstract an informative and balanced summary of what was done and what was found	Structured abstract (Background, Methods, Results, Conclusions) — Page 1
Background / Rationale	**2**	Explain the scientific background and rationale for the investigation being reported	Introduction, Section 1: 'The Clinical Trial Termination Crisis' and 'From Prediction to Prevention' — Pages 2–3
Objectives	**3**	State specific objectives, including any prespecified hypotheses	Introduction, Section 'Study Objectives' — 4 prespecified objectives stated — Page 3
Study Design	**4**	Present key elements of study design early in the paper	Methods, 'Study Design and Data Source': retrospective counterfactual simulation, AACT database, 2016–2021 — Page 4
Setting	**5**	Describe the setting, locations, and relevant dates, including periods of data collection	Methods, 'Study Design and Data Source': ClinicalTrials.gov/AACT, January 2016 – December 2021, accessed January 2025 — Page 4
Participants	**6a**	Give the eligibility criteria, and the sources and methods of selection of participants	Methods, 'Study Design and Data Source': 96,788 interventional trials, status Completed or Terminated, sufficient protocol detail for network reconstruction — Page 4

	6b	Not applicable (no matching); cohort is full registry population meeting inclusion criteria	N/A — full eligible cohort included
Variables	**7**	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers	Methods, 'Network Estimation and SFI Computation' and 'Structural Redesign Scenarios': SFI components (C, H, λ_amp), eligibility count, outcomes, sponsor class — Pages 4–5
Data Sources / Measurement	**8**	For each variable of interest, give sources of data and details of methods of assessment	Methods: AACT database [Ref 14]; SFI computation formula specified; graphical LASSO parameters stated (λ_reg=0.05); weights w1=0.42, w2=0.31, w3=0.27 — Pages 4–5
Bias	**9**	Describe any efforts to address potential sources of bias	Discussion, 'Strengths and Limitations': residual confounding from unmeasured factors acknowledged; prespecified scenarios; bootstrap CIs; sensitivity analyses — Page 8
Study Size	**10**	Explain how the study size was arrived at	Methods: full AACT cohort 2016–2021 used; no sample size calculation required as this is a population-level analysis of all eligible registry trials — Page 4
Quantitative Variables	**11**	Explain how quantitative variables were handled in the analyses	Methods: SFI log-transformed and min-max scaled; quintile thresholds defined (Q3=0.74); scenario thresholds prespecified (eligibility ≤16, C≤Q3) — Page 5
Statistical Methods (a)	**12a**	Describe all statistical methods, including those used to control for confounding	Methods, 'Statistical Analysis': Python 3.10, logistic model, BCa bootstrap (10,000 replicates); validated prediction model [Ref 9] used for counterfactual probabilities — Page 5
Statistical Methods (b)	**12b**	Describe any methods used to examine subgroups and interactions	Methods, 'Sensitivity Analyses': phase-stratified PF; industry vs non-industry stratification — Page 5
Statistical Methods (c)	**12c**	Explain how missing data were addressed	Methods: inclusion criterion 'sufficient protocol detail for

			network reconstruction' excludes incomplete records; no imputation required — Page 4
Statistical Methods (d)	**12d**	Not applicable — no follow-up period; registry status at end of period is the outcome	N/A — registry-based, no loss to follow-up
Statistical Methods (e)	**12e**	Describe any sensitivity analyses	Methods, 'Sensitivity Analyses': conservative redesign (cap ≤20), aggressive optimisation (cap ≤12), phase-stratified PF, industry vs non-industry — Page 5
Participants	**13a**	Report numbers of individuals at each stage of study	Results, 'Baseline Cohort Characteristics': 96,788 total; 81,011 completed (83.7%); 15,777 terminated (16.3%) — Page 6, Table 1
	13b	Give reasons for non-participation at each stage	Methods: trials excluded if insufficient protocol detail for network reconstruction; no further exclusion stages — Page 4
	13c	Consider use of a flow diagram	Figure 1: Counterfactual Redesign Simulation Workflow (4-step pipeline diagram) — Page 10
Descriptive Data	**14a**	Give characteristics of study participants and information on exposures	Results, Table 1: Baseline characteristics (phase, sponsor type, eligibility, endpoints, SFI, load concentration, network entropy) for completed vs terminated — Page 6
	14b	Indicate number of participants with missing data for each variable of interest	Inclusion criterion requires complete protocol detail; missing data excluded at eligibility stage — Page 4
	14c	Not applicable — no follow-up period in registry-based simulation	N/A
Outcome Data	**15**	Report numbers of outcome events or summary measures	Results: 15,777 terminated trials (16.3% termination rate); counterfactual outcomes reported for all 4 scenarios — Page 6, Table 2

Main Results	**16a**	Give unadjusted and confounder-adjusted estimates with precision	Results, Table 2: predicted termination rates with 95% BCa CIs for all 4 scenarios; PF with 95% CIs; absolute and relative reductions — Pages 6–7
	16b	Report category boundaries when continuous variables were categorised	Methods and Results: SFI quintile boundaries defined; Q3=0.74 acceleration threshold; eligibility inflection at 16 criteria — Pages 5–6
	16c	If relevant, translate relative risk into absolute risk for a meaningful time period	Results, Table 3: absolute preventable terminations (4,260; 95% CI 3,865–4,655) and economic savings ($8.61B) over 2016–2021 — Page 7
Other Analyses	**17**	Report other analyses done — subgroups, interactions, sensitivity analyses	Results, 'Sensitivity Analyses': conservative/aggressive redesign scenarios; phase-stratified PF; industry vs non-industry PF — Page 7
Key Results	**18**	Summarise key results with reference to study objectives	Discussion, 'Principal Findings': PF=0.270, 4,260 terminations, $8.61B savings; three mechanistic insights — Page 7
Limitations	**19**	Discuss limitations, including sources of bias or imprecision	Discussion, 'Strengths and Limitations': 5 limitations stated including observational design, residual confounding, generalisability, implementation feasibility, equity implications — Page 8
Interpretation	**20**	Give a cautious overall interpretation considering objectives, limitations, and other evidence	Discussion, 'Comparison with Existing Literature' and 'Policy Implications': findings contextualised against prior registry analyses; policy implications stated with caveats — Pages 7–8
Generalisability	**21**	Discuss the generalisability (external validity) of the study results	Discussion, 'Strengths and Limitations': generalisability beyond 2016–2021 awaits external validation; equity implications for rare disease trials noted — Page 8

Funding	**22**	Give the source of funding and the role of funders	Declarations: 'This research received no funding' — Page 9

Note: This study is a retrospective counterfactual simulation using publicly available registry data (AACT/ClinicalTrials.gov). Some STROBE items designed for classical participant-level cohort studies are not fully applicable (marked N/A) due to the population-level simulation design. All applicable items are addressed. STROBE = Strengthening the Reporting of Observational Studies in Epidemiology. Reference: von Elm E, et al. Lancet. 2007;370(9596):1453–1457.