

Threats to Validity of the PESTO Trial

Threats to Validity of the PESTO Perispinal Etanercept Stroke Trial: A Document-Based Clinical Trial Evaluation

Edward Tobinick¹; Güven Gürbüz²

Edward Tobinick, M.D., Certified by the American Board of Internal Medicine

Director, Institute of Neurological Recovery

Corresponding author: Edward Tobinick - director@nrimed.com

Telephone: (+1) (561) 353-9707 Fax number: (+1) (561) 372-7874

¹Institute of Neurological Recovery (INR PLLC), 1877 S. Federal Highway, Suite 110, Boca Raton, FL 33432, USA

Güven Gürbüz²

Neurologist at RN Clinics

²RN Clinics, Aarhus, Denmark

Threats to Validity of the PESTO Trial

Summary

The PESTO trial's claim to provide Class I evidence that perispinal etanercept does not improve quality of life is precluded by the absence of publicly verifiable trial-wide inclusion/exclusion criteria and by its use of a novel primary responder definition that utilizes a non-accepted scoring pathway, of no established validity or reliability, that was first specified after data collection was completed. PESTO's Statistical Analysis Plan differs from the analysis its operative protocols and registration specify, and no amendment, registry update, or explanation discloses the change. PESTO's published primary quality-of-life measurement is unreliable, and its definitive negative conclusion cannot be relied upon as a statement of what the trial established. Whether the reported estimate would survive the prespecified analysis, the record cannot yet answer. Correction of PESTO's record and conclusions and reclassification of its Class I designation are due now. Further, independent multidisciplinary review, with separate statistical and research-ethics components, is required.

Keywords

CONSORT; etanercept; neuroinflammation; perispinal; PESTO; research integrity; stroke; TNF.

Article highlights

- **The trial utilized a novel, three-route, PESTO-devised responder definition that is not an accepted SF-36 scoring pathway, has no established validity or reliability as a measure of clinically meaningful change, and was first specified after data collection was completed. The primary endpoint analyzed was not the endpoint PESTO prespecified. No disclosed amendment, registry update, or explanation accounts for the difference.**
- **On the published record, PESTO's published primary result is unreliable as a measurement of quality of life; neither its Class I designation nor its definitive negative conclusion can be relied upon as statements of what the trial established.** Correction and reclassification are due now; independent multidisciplinary review is required for the remainder.
- **Through a change in covariate adjustment, the primary-analysis model departed from the model prespecified in PESTO's registration and in every signed protocol.**
- **No disclosed document explains the change in responder definition or in covariate adjustment.** CONSORT 2010 Items 3b and 6b, CONSORT 2025 Item 10, ICH E9 and the reporting requirements Neurology adopted before PESTO was received each require changes of this kind to be reported with reasons.
- **The New Zealand ethics committee prohibited proxy consent at the Christchurch site, but no principal trial document discloses that fact. The Christchurch site's ethics-approved protocol is also not publicly disclosed.**

Threats to Validity of the PESTO Trial

- **Protocol v2.1 required the injection to be performed by an experienced neurologist or stroke physician and to be videotaped. Compliance with these requirements is not confirmed on the published record.**
- **The records that would resolve these issues were not obtainable through the requests described here.** The sponsor declined all seven categories sought, on the contention that they were internal operational records; the trial site stated that governance of the injection recordings was a sponsor function and that it generated, retained and obtained none; and the approving Australian ethics committee replied that operational implementation is a matter for the sponsor.

1. Introduction

Stroke is a leading cause of long-term disability worldwide, and pharmacological options for improving chronic post-stroke impairments remain limited [1-3]. There is no established restorative drug approved for core chronic stroke deficits such as impaired mobility, language, cognition, or chronic post-stroke pain, leaving a large population with substantial unmet need [3].

The Perispinal Etanercept for Stroke Outcomes Study (PESTO) was a three-center Australian and New Zealand (NZ) randomized clinical trial (RCT), funded by a government grant program seeking breakthrough treatments for stroke survivors (SF1, SF2) [4-6]. It was published in *Neurology* as providing Class I evidence that perispinal etanercept (PSE) does not improve quality of life in chronic stroke [5].

This article is not a report of a new clinical trial. It is a document-based evaluation of PESTO's design, conduct, research-ethics governance, analysis, reporting, and evidentiary classification, applying the established framework of internal, construct, statistical-conclusion and external validity together with recognized standards for clinical-trial conduct and reporting [7].

2. Methods of document review

This evaluation is entirely document-based (Table 1). The source material comprised registry entries, published articles and supplements, ethics-committee records, and documents obtained under freedom of information (FOI) from Austin Health, Melbourne (FOI 25/26-1266, 25/26-1561, 25/26-1735, 26-1740, 26-1747), including the site delegation log and the internal-review decision described in §4.2.3, the Australian Government Department of Health, Disability and Ageing (FOI 26-2767), and the Australian Government Department of Industry, Science, Energy and Resources (FOI 300356), together with the Southern Health and Disability Ethics Committee's 2020 and 2021 minutes and its 2020–2024 annual reports. These were evaluated against International Council for Harmonisation (ICH) E3, E6(R2) and E9, the Consolidated Standards of Reporting Trials (CONSORT), Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT), the Template for Intervention Description and Replication (TIDieR), and the approval conditions and protocol commitments in the trial's ethics records [8-15].

Threats to Validity of the PESTO Trial

Protocol searches used the term “perispinal” in the Australian New Zealand Clinical Trials Registry (ANZCTR) and ClinicalTrials.gov; Google searches for “perispinal etanercept” and “stroke trials” identified the Griffith trial and PESTO [5,16]. Searches were last updated on 16 July 2026. “The record reviewed” means the documents so located, the published supplemental record, the registry, and the FOI and ethics materials available to the authors; it is not a claim that no additional record exists. Missing documentation is treated as a limit on verification, not as proof that the protocol was violated, an ethics condition disregarded, or the intervention administered incorrectly.

Factual verification. Confirmation of the protocol-required injector qualifications and videotaping was sought from four quarters, each documented and set out with its outcome in §4.2.3: the sponsor, which declined all seven categories requested (SF13, SF14) [17,18]; the trial site under the Freedom of Information Act 1982 (Vic), which produced the delegation log and an internal-review decision (SF15–SF17) [19–21]; nine PESTO authors, from whom no substantive answer was received; and the Austin Health Human Research Ethics Committee (SF18) [22]. Counsel for the sponsor stated generally that the trial had been conducted in accordance with the approved protocol, but provided no site-specific confirmation or supporting records.

Conflict management and document verification. Both authors disclose relevant professional or financial interests. To reduce interpretive bias, factual assertions were anchored to cited primary documents, documentary observation was separated from interpretation, and categorical claims were avoided where the record could support more than one explanation. The corresponding author rechecked each date and document identifier in Table 1, each quoted protocol provision and each cross-document comparison against the source PDF or registry entry; AI-generated output was not treated as evidence. In February 2026 the corresponding author requested the trial dataset and seven categories of governance records (§5).

3. The trial

3.1 Background

Perispinal etanercept (PSE) is an off-label administration method for the TNF inhibitor etanercept, delivered by perispinal injection followed by positioning intended to promote cephalad transport [23–30]. An earlier double-blind, placebo-controlled RCT of PSE in chronic stroke, conducted at Griffith University by investigators trained in the injection technique, used a prospectively enriched design of the kind U.S. Food and Drug Administration guidance recommends for efficacy trials [16,31,32]. PESTO followed, with an entirely different design [5,32]. Selected protocol-related documentary events are set out in Table 1.

Table 1. Key PESTO documents and document-related events

Date	SF#	Event
7 Oct 2018	1	MRFF Return to Life, Return to Work Clinical Research Grants Application Guide
18 Mar 2019	2	Stroke Foundation media release: Return to Life, Return to Work grants announced

Threats to Validity of the PESTO Trial

30 May 2019	3	Protocol v1.0 - first signed protocol in the record; does not name New Zealand; permits proxy consent; defines primary response as ≥ 5 -point improvement on the SF-36; primary analysis adjusts for proxy/self completion
8 Sep 2020	4	Southern HDEC directs removal of proxy consent because it is not permitted in New Zealand; New Zealand principal investigator present
7 Oct 2020	5	Date of initial ANZCTR registration; retains the original ≥ 5 -point SF-36 primary response definition and proxy/self primary-analysis covariate through 16 July 2026
11 Dec 2020	6	Southern HDEC approves the revised Christchurch protocol; that ethics-approved protocol was not located in the record reviewed
12 May 2021	7	Stroke Foundation letter reports that the New Zealand site had enrolled 18 participants by April 2021 and that the enrollment target at PESTO's Christchurch trial site was 50
9 Jun 2021	8	Protocol v2.0 widens the stroke window and age limit; does not name Christchurch
31 May 2022	9	Protocol v2.1 —the “most recent” signed protocol; first publicly disclosed protocol to name Christchurch; retains proxy consent, ≥ 5 -point SF-36 response, and proxy/self covariate
1 Jun 2022	10	DSMB Safety Report documents 27 participants enrolled in New Zealand through 28 May 2022
24 Nov 2023	5	Date of completion of data collection listed on PESTO's ANZCTR registration page
18 Dec 2023	11	SAP dated; introduces the PCS/MCS/summed-change responder definition and replaces proxy/self completion with time since stroke
14 Apr 2024	11	SAP signed as final by the Biostatistics Lead and principal investigator
27 Apr 2024	Ref [33]	European Stroke Journal protocol paper published online; includes the three-route outcome definition but retains the proxy/self primary-analysis covariate and trial-wide proxy consent
8 Oct 2024	Ref [34]	Neurology publishes RCT-reporting requirements requiring protocols, SAP, and explanations of substantive changes
27 Mar 2025	Ref [5]	Neurology trial report received
9 Jun 2025	Ref [5]	Neurology trial report accepted
4 Sep 2025	Ref [5]	Neurology trial report published online
28 Sep 2025	12	ANZCTR record updated; the ANZCTR-listed protocol paper is listed as the study protocol; no ethics-approved Christchurch protocol listed on the registry. No mention of proxy prohibition in New Zealand.
11 Feb 2026	15	FOI request to Austin Health for seven categories of site governance records, including the site delegation log, injector qualification and training records, and video-recording governance records
17 Feb 2026	13	Request for sharing of PESTO data by corresponding author to PESTO corresponding author
11 Mar 2026	14	Reply to request for sharing of PESTO data by corresponding author to PESTO corresponding author
16 Mar 2026	16	Austin Health FOI internal-review decision: governance of the injection video recordings was a sponsor function; no such records generated or retained at Austin Health; copies not obtained from the sponsor
16 Mar 2026	17	Site Signature and Delegation of Responsibilities Log, Austin Health site, produced under FOI 25/26-1266 in two productions with differing redactions; records no delegation of the task “Administer IP”
30 Mar 2026	18	Austin Health HREC response declining to address third-party queries regarding trial conduct and identifying operational implementation as a sponsor responsibility

Summary of the key documentary record. Documents and events are listed in chronological order to provide the evidentiary context for subsequent analyses of protocol development, ethics approvals, trial registration, statistical analysis, reporting, and data sharing correspondence.

Threats to Validity of the PESTO Trial

Abbreviations: ANZCTR = Australian New Zealand Clinical Trials Registry; DSMB = Data Safety Monitoring Board; HDEC = Health and Disability Ethics Committee; MRFF = Medical Research Future Fund; SAP = statistical analysis plan; SF = supplementary file.

3.2 Design and analysis as specified

PESTO was a phase 2b, multicenter, double-blind, placebo-controlled RCT. Eligible participants had imaging-confirmed ischemic or hemorrhagic stroke between one and fifteen years before enrollment, a modified Rankin Scale (mRS) score of 2 to 5, and reduced quality of life on the 36-item Short Form Health Survey (SF-36). Participants were randomized 1:1 to etanercept 25 mg or matched saline placebo and again at day 28 for a day-56 secondary analysis. Randomization was stratified by proxy-versus-self completion of the pre-randomization SF-36, baseline mRS category, and time since stroke. Protocols v1.0 and v2.1 and the ANZCTR registration defined the day-28 primary response as improvement of at least 5 points on the SF-36; they did not specify separate qualifying routes through the Physical Component Summary (PCS), Mental Component Summary (MCS), or the sum of their changes [35-37].

Protocol v2.1 (SF9) carries a Statement of Compliance on its cover page: “This study will be conducted in compliance with all stipulation[s] of this protocol, the conditions of the ethics committee approval, the NHMRC National Statement on Ethical Conduct in Human Research (2007) and the Note for Guidance on Good Clinical Practice [GCP] (CPMP/ICH-135/95).”

Protocol v2.1 specified logistic regression with treatment arm as the independent variable and adjustment for baseline SF-36 score and proxy-versus-self SF-36 completion. PESTO’s SAP, first made public with the trial report on 4 September 2025, used a different primary-analysis covariate and a materially different three-route responder definition (SF11) [38].

3.3 Results as reported

Of 147 participants assessed for eligibility, 126 were enrolled and randomized between November 2020 and September 2023, 63 per arm. The trial report states that 126 participants were included in the primary efficacy evaluation; the reported proportions correspond to denominators of 62 per arm, consistent with one participant in each arm having missing primary-outcome data. Median age was 54.5 years; 38% were women; median time since stroke was 3 years; 67% had ischemic stroke overall; 41% had baseline mRS 4–5. Twenty-nine of 126 participants (23%) had their baseline SF-36 completed by a proxy.

On the reported primary complete-case intention-to-treat analysis, the primary outcome was achieved by 33 participants (53%) allocated to etanercept and 36 (58%) allocated to placebo, with adjusted odds ratio 0.82 (95% CI 0.40–1.67) and standardized risk difference –5% (95% CI –22% to 13%).

Five serious adverse events were reported over 56 days, similarly distributed across arms; no participant discontinued the assigned intervention.

4. Threats to validity and unresolved questions of trial conduct and governance

Threats to Validity of the PESTO Trial

A framework for establishing the validity of scientific reports has been established - internal, construct, statistical-conclusion, and external validity [7,39-41]. A second approach belongs to research-ethics governance, and it takes a different form: which approved protocol governed each site, whether site-specific consent conditions were observed, whether substantive changes were authorized and disclosed, and whether the published account gives an accurate picture of how the trial was run.

Two principles frame what follows. The first is that the burden of transparent disclosure rises, rather than falls, when a result is advanced as Class I evidence - the highest level in the American Academy of Neurology (AAN) classification [42]. The second is that the evidentiary limits here are asymmetric. Missing records bear on verifiability; inconsistencies among disclosed documents can be demonstrated directly from the documents themselves. The evidentiary status of each finding, and the inference each will and will not carry, is recorded item by item in Table 3. Statistical reanalysis and ethics-governance review are therefore complementary remedies, not interchangeable ones.

The sections that follow are ordered accordingly. Section 4.1 takes the site-specific ethics-governance, consent, and population questions that the record reviewed does not permit to be resolved. Section 4.2 turns to the primary-outcome, analysis, and intervention-fidelity concerns. Section 4.3 identifies the reporting and research-governance standards most directly relevant to them.

4.1 The New Zealand site: unresolved ethics-governance questions

PESTO's New Zealand site (Christchurch, Canterbury Geriatric Medicine Research Trust) raises questions of a different evidentiary character from those in §4.2. They rest in part on records the authors could not obtain, and they are stated here as verification gaps rather than as established departures (Table 2).

Table 2. New Zealand site: ethics-governance and verification gaps

Item	What the record reviewed shows and does not show
Proxy-consent prohibition	The Southern HDEC prohibited proxy consent at Christchurch. Protocol v2.1, the protocol paper, and the trial report do not disclose the site-specific restriction.
Ethics-approved Christchurch protocol	The Southern HDEC approved a revised Christchurch protocol on 11 December 2020. The approval letter and approved protocol were not located in the record reviewed.
Seventeen-month documentary gap	No publicly disclosed protocol located by the authors names Christchurch between 11 December 2020 and 31 May 2022, although the site was enrolling participants.
Eligibility criteria	The record does not permit verification of whether the eligibility criteria applied at Christchurch were identical to those applied at the Australian sites.
Injector qualifications	No record available to the authors identifies the Christchurch injector as an experienced neurologist or stroke physician.
Videotaping	The record available to the authors does not confirm whether injections were videotaped at Christchurch or at the other sites.
Site-level data	The trial report provides no Christchurch enrollment dates or count, baseline characteristics, or site-level or site-adjusted results.

Summary of the documentary evidence relating to ethics governance and implementation at the New Zealand study site. For each item, the table indicates what the record reviewed supports and what remains unverifiable from the available documentation.

Threats to Validity of the PESTO Trial

Abbreviation: HDEC = Health and Disability Ethics Committee.

On 8 September 2020 the Southern Health and Disability Ethics Committee (HDEC) required removal of proxy (“responsible person”) consent from the protocol governing Christchurch because proxy consent was not permitted in New Zealand; the minutes record that only participants competent to consent would be enrolled there (SF4) [43]. That site-specific restriction is disclosed in none of the principal public trial documents. Protocol v2.1, the first publicly disclosed protocol to name Christchurch, retains proxy consent as a trial-wide provision [35]; the ANZCTR-listed protocol paper presents consent by a medical decision maker as a trial-wide inclusion criterion, without identifying the New Zealand restriction [33,43]; and the trial refers to consent by participants or their legal representatives without identifying the restriction [5,43]. The trial report’s undifferentiated consent language is compatible with no proxy consent having been used in New Zealand. The narrower finding is that no principal document informs readers that one site operated under a binding restriction on a consent pathway presented elsewhere as available trial-wide. That omission matters because proxy-versus-self completion of the SF-36 was one of the trial’s three randomization stratification variables and the covariate every signed protocol prespecified for the primary analysis (§4.2.1). A site-specific restriction defined in relation to that variable bears on the composition of the pooled population.

The Southern HDEC approved a revised protocol naming Christchurch on 11 December 2020, as recorded in its 2021 Annual Report (SF6) [44]. That protocol was not located in the published supplemental record, the registry, or any other record available to the authors. Protocol v2.1, dated 31 May 2022, is the first publicly disclosed protocol to name the site - seventeen months later - and its change history states: “Site Christchurch was not mentioned in prior protocol, now added” [35]. Christchurch had enrolled 18 participants by April 2021 and 27 by 28 May 2022 (SF7, SF10) [45,46]. The public record therefore does not identify the full eligibility and consent provisions governing those enrollments or permit their site-specific contribution to be separated from the pooled result.

The eligibility and consent provisions governing those enrollments therefore cannot be identified from the public record, and the trial report reports no Christchurch enrollment dates, count, baseline characteristics, or site-level results. The AAN scheme requires clearly defined inclusion and exclusion criteria for Class I designation [42]; that predicate cannot be checked here. Protocol v2.0 was obtained only in part under FOI, its substantive pages withheld under the internal-working-documents exemption in section 30(1) of the Freedom of Information Act 1982 (Vic). These questions are curable by disclosure of the 11 December 2020 Christchurch protocol, the approval correspondence, the consent documents and site-specific enrollment information. Statistical reanalysis cannot answer them; they require independent research-ethics review of the complete oversight and site record.

4.2 Outcome, analysis, and intervention-fidelity concerns across validity and conduct

Threats to Validity of the PESTO Trial

Two concerns in this section are affirmative discrepancies among disclosed documents: the primary-analysis model and the primary-outcome definition. The intervention-fidelity concern is different in kind: it is a verification gap. These concerns have distinct dimensions. Participant-level reanalysis can determine whether the outcome and analysis changes affected the treatment estimate; only the complete protocol, ethics, monitoring, and site records can determine whether substantive changes and fidelity safeguards were appropriately authorized, implemented, documented, and reported. Table 3 summarizes the evidentiary status, permissible inference, and independent review required for each principal finding.

Table 3. Principal cross-domain concerns, evidentiary status, and required independent review

Issue	Evidentiary status	What the record supports and does not support	Principal implication of each issue
New Zealand protocol and eligibility	Missing record	The eligibility criteria governing Christchurch enrollment cannot be reconstructed from the public record; this does not prove that different criteria were applied.	Prevents verification of the Class I eligibility predicate and requires independent research-ethics review to identify the protocol and criteria governing Christchurch enrollment.
New Zealand proxy-consent restriction	Affirmative ethics record plus publication omission	The site-specific prohibition of proxy consent is documented in ethics-committee minutes but omitted from the principal trial documents; the record does not establish that prohibited proxy consent was used.	Requires independent review of implementation, consent records, and disclosure of the site-specific ethics condition.
Primary-outcome definition	Affirmative cross-document discrepancy; unexplained departure from prespecification	The SAP and report use a three-route responder definition that differs from that contained in the signed protocols and registry, and no disclosed document explains the change, contrary to the reporting standards applicable to the trial; its numerical effect is unknown.	Requires independent reanalysis of each route, their overlap, and the prospectively specified endpoint, together with review of the amendment and approval history.
Primary-analysis covariate	Affirmative cross-document discrepancy; unexplained departure from prespecification	The reported model substituted time since stroke for proxy/self completion, and no disclosed document explains the change. The result was therefore not produced by the prespecified model. A legitimate pre-unblinding refinement remains possible; material effect is unknown.	Requires independent sensitivity analysis under both models and review of whether the change was appropriately documented and, where required, approved.
Injection-fidelity safeguards	Missing confirmation, together with an affirmative omission in a disclosed site record	Compliance with protocol-required qualifications and videotaping is not independently verified. The site authorization record produced for the Austin Health site does not record delegation of "Administer IP" to any person; this does not establish that the intervention was administered by an unauthorized or unqualified person.	Requires independent review of credentials, delegation records, videos, monitoring, and deviation records to determine whether the protocol-defined safeguards were implemented and documented.

Threats to Validity of the PESTO Trial

Summary of the principal concerns arising from the documentary record, the evidentiary status for each, what the record supports and does not support, and the resulting implications for independent review and interpretation.

4.2.1 The reported primary-analysis model differed from every publicly prespecified model

From Protocol v1.0 (30 May 2019) onward, PESTO prespecified adjustment for whether the SF-36 was completed by the participant or by a proxy [36]. Protocol v2.1 states the model in terms, providing that the primary endpoint “will be tested using a logistic regression model with the treatment arms as an independent variable and baseline value of SF-36, proxy, or self administration of SF-36 as adjustment covariates” [35]. The covariates are named; the specification is not left open. The same covariate appears in Protocol v1.0, Protocol v2.1, the ANZCTR registration, and the ANZCTR-listed protocol paper (Table 4) [33,35-37].

The SAP, dated 18 December 2023 and signed as final on 14 April 2024, replaced proxy-versus-self completion with time since stroke (SF11) [38]. The ANZCTR PESTO page lists 24 November 2023 as the end of data collection. No publicly disclosed protocol amendment, registry update, or explanation reconciles the substitution. The ANZCTR-listed protocol paper, published thirteen days after final SAP signature, continued to present the proxy-versus-self model.

The record does more than fail to show an amendment. The SAP identifies on its cover page the protocol under which it was prepared - Protocol Number 2, Version 2.1, dated 31 May 2022, the same protocol the trial report identifies as most recent - and its approval page records that the signatories approve the plan as final and find it consistent with the requirements of that protocol [38]. Protocol v2.1 carries its own change history recording the amendments that produced versions 2 and 2.1 [35]. On that record the operative protocol when the SAP was signed was v2.1, and v2.1 specifies neither the covariate nor the outcome definition the SAP adopted. Therefore, this is no longer a case in which a document is simply missing from the record. The trial’s own documents indicate what the operative protocol was, and that its primary outcome and primary-analysis covariate are not the ones analyzed.

A legitimate pre-unblinding refinement remains a possible explanation, and it should be given its full weight. Finalizing or refining an SAP before unblinding can comply with ICH E9, and the SAP addresses the point directly: its approval page records that the signatories find the plan compliant with ICH E9 principles and confirm that it was developed in a completely blinded manner, without knowledge of the effect of the intervention being assessed [38]. We take that attestation at face value. It is, however, an attestation by the two signatories rather than an independently verifiable record, it is dated 14 April 2024, and it addresses how the plan was developed rather than whether the specifications it adopted were disclosed. Blinded development and undisclosed departure from a prespecified protocol are not alternatives; the attestation answers the first and is silent on the second. The effect on the treatment estimate cannot be determined here in any event, since that would require participant-level sensitivity analyses under both models.

Threats to Validity of the PESTO Trial

A further feature of the substituted model appears within the SAP itself. The SAP records that randomization was stratified by three variables: whether the pre-randomization SF-36 was completed by the patient or by a proxy, baseline modified Rankin Scale dichotomized as 2-3 versus 4-5, and time since stroke dichotomized as 1-5 years versus more than 5 years [38]. The primary-analysis model it specifies adjusts for baseline SF-36 score and time since stroke in years. Two of the three stratification factors are therefore absent from the primary model, and the one retained enters as a continuous variable rather than in the dichotomized form in which randomization was stratified. Proxy versus self completion is thus omitted twice over: it is both the covariate the protocols prespecified and a factor on which randomization was stratified. The SAP derives its standardized risk difference from FDA guidance on adjusting for covariates in randomized trials [38], guidance directed at the treatment of exactly these variables. We do not suggest the resulting estimate is invalid; we note that the model reported departs from the trial's own randomization structure as well as from its protocols, and that neither departure is explained.

What can be verified is narrower, and it concerns transparency and whether the reported result withstands the prospectively specified analysis. The model used for the headline result differed from every publicly prespecified model, and a reader has no way of telling whether the estimate is stable to the substitution. Independent reanalysis should report the primary result under the proxy-versus-self model as well as the time-since-stroke model. Whether the change was documented in the protocol record, and where required put to the sponsor and the ethics committees, is a separate question for governance review.

The SAP itself first reached the public in the supplemental file released on the day the trial report was published, 4 September 2025 (esap_1, SF11).

Table 4. Primary-outcome and primary-analysis specifications across PESTO documents

Document	Date	Primary-outcome definition	Primary-analysis covariate	Documentary note
Protocol v1.0	30 May 2019	≥5-point improvement on the SF-36	Baseline SF-36 and proxy/self completion	First signed protocol (SF3)
ANZCTR registration	Registered 7 Oct 2020	≥5-point improvement on the SF-36	Baseline SF-36 and proxy/self completion specified initially	Initial primary outcome definition and primary analysis covariate unchanged through 16 July 2026 (SF5)
Protocol v2.1	31 May 2022	≥5-point improvement on the SF-36	Baseline SF-36 and proxy/self completion	Most recent signed protocol supplied with trial report (SF9)
Statistical analysis plan	Dated 18 Dec 2023; signed 14 Apr 2024	PCS ≥5, MCS ≥5, or sum of PCS and MCS changes ≥5	Baseline SF-36 and time since stroke	No publicly disclosed amendment, registry update, or explanation (SF11)
Registry-listed Protocol paper	Online 27 Apr 2024	PCS ≥5, MCS ≥5, or sum of changes ≥5	Baseline SF-36 and proxy/self completion	Late protocol paper combines revised outcome with earlier covariate
Neurology trial report	Online 4 Sep 2025	Text: PCS ≥5, MCS ≥5, or ≥5 in the sum of the two changes. Results table:	Baseline SF-36 and time since stroke	Uses SAP analysis; the report's text and its results

Threats to Validity of the PESTO Trial

		change in “Mental or Physical or Total SF-36”		table describe the outcome differently
--	--	---	--	--

Summary of the cross-document comparison of the primary-outcome definition and primary-analysis covariate in the PESTO documentary record. Including tracking the evolution of these specifications across trial documents and highlighting any discrepancies relevant to the reported primary analysis.

The trial report is not silent about protocol changes in general. It discloses several eligibility amendments, each attributed to the protocol version that introduced it. What it does not disclose is the change to the primary-analysis covariate, or the materially different primary-outcome definition used for the reported result.

That distinction matters, and its consequence is not confined to transparency. A result generated by a model other than the prespecified one, where no disclosed amendment authorizes the substitution, is not a confirmatory result: the protection prespecification affords against analytic choices made in the presence of data has not been shown to apply to it. What an undisclosed covariate substitution establishes is that the model producing PESTO’s headline result is not the model the trial prespecified, that no reported analysis has ever applied the prespecified model, and that a reader has no means of determining whether the two would agree.

4.2.2 The primary outcome was materially redefined without a publicly disclosed protocol amendment

Protocols v1.0 and v2.1 and the ANZCTR registration all define a responder the same way: a participant improving by at least 5 points on the SF-36 at day 28 [35-37]. Protocol v2.1 defines a responder as “an achievement of at least 5 points improvement on the SF-36 score” [35]. Alternatively, the Statistical Analysis Plan defines a responder as “[t]he change in SF-36 is defined as the sum of changes in the Physical Component Summary Score or the Mental Component Summary or both of the SF-36 between baseline and Day 28” [38]. It is the three-route reading of that sentence - a 5-point improvement in the PCS, a 5-point improvement in the MCS, or a summed change of at least 5 points - that the ANZCTR-listed protocol paper states and the trial report used [5,33].

Two features of this pairing matter, and they are different in kind. The first is that the protocol’s endpoint is under-specified on its face. The SF-36 does not yield a single score, so “the SF-36 score” does not by itself identify a measurand, and a plan resolving the ambiguity was required. The second is when and how the ambiguity was resolved. The SAP bears the date 18 December 2023 and was signed as final on 14 April 2024; the ANZCTR page lists 24 November 2023 as the end of data collection. The specification that fixed what the primary endpoint meant was therefore made after the data existed. That is the circumstance prespecification exists to prevent, and it is not answered by observing that an analysis plan may properly operationalize a protocol endpoint - the objection is not that the SAP made a choice, but that it made this choice at this point and that neither the choice nor its timing is disclosed anywhere a reader would find it.

Threats to Validity of the PESTO Trial

That definition did appear in print before the results, in the European Stroke Journal protocol paper, although the paper was received after data collection had ended. The protocol paper states the three-route definition, but does not identify it as a change from the signed protocols or the registration and gives no reason for it. The definition appears in neither signed protocol and in no version of the prospective registry entry, and nothing in the public record reconciles the difference.

The third route warrants separate attention, because it does not merely add a way of meeting the protocol threshold; it changes what is being measured. The SF-36 has no standard total score[47]. PCS and MCS are separately derived norm-based composites, each formed by weighting all eight domain scores, and they are not two additive halves of a single quantity[48]. Their sum is an investigator-constructed measure rather than a recognized SF-36 score, and no established threshold of clinical importance attaches to it. The 5-point criterion the signed protocols and the registration applied to a single score has been transferred to a different quantity without stated justification[49].

PESTO did not provide any validation or basis to suggest the reliability of its three-route primary outcome construction, and, to our knowledge and after diligent search, no prior randomized trial - in stroke or in any perispinal etanercept study - has used it. Nor is it an accepted SF-36 scoring pathway: the instrument's developers state that the two component summaries measure distinct physical and mental constructs and that combining them into a single overall score is not appropriate [47,51]. The senior developer of the SF-36, John E. Ware, Jr., has confirmed that the three-route construction is not an "accepted scoring pathway" (J.E. Ware, Jr., personal communication, 2026, cited with written permission). This position is reinforced by the component-scoring literature and by published reviews of more than a decade of randomized-trial reporting, in which SF-36 outcomes are reported as PCS and/or MCS or as one or more of the eight SF-36 individual domains that have been validated [47-52]. The validity of a patient-reported outcome endpoint is not inherited from the validation of its constituent scores; the definition actually used in a trial must itself have a demonstrated and interpretable relationship to meaningful within-patient change [47-50].

A consequence follows arithmetically. A participant may meet the summed criterion while deteriorating on one component: a physical-summary gain of 8 points against a mental-summary loss of 3 yields a summed change of +5 and counts as achieving the primary outcome. And the criterion can be satisfied by component changes each substantially smaller than the 5-point improvement either signed protocol required.

The trial report gives none of what a reader would need to weigh this: not the number of participants qualifying by each route, not the overlap among the routes, and not a sensitivity analysis applying the outcome as the signed protocols and the registry specify it [53]. There is accordingly no way to confirm that the headline result corresponds to the prospectively specified endpoint. Quantifying the effect of the three-route rule is a task for independent reanalysis, and establishing whether the redefinition was prospectively authorized is a task for independent review of the amendment and ethics record. What can be said without either is that the primary result PESTO reports is not the result of the analysis PESTO prespecified, that the prespecified analysis has never been reported, and

Threats to Validity of the PESTO Trial

that the data required to perform it are not currently available to anyone outside the trial team (§5).

The summary scores the redefinition relies on have had their validity in stroke specifically questioned. Hobart et al. found that the scaling assumptions required to generate PCS and MCS scores were not satisfied in their ischemic-stroke sample, and concluded that reporting those summary scores was unsupported [49]. That study examined a different stroke population and did not test PESTO's summed-change rule, so it settles nothing here; it does show why a composite responder definition of this kind calls for prospective justification and validation.

4.2.3 Compliance with protocol-defined intervention-fidelity safeguards is not independently verified

Protocol v2.1 treated operator experience, procedure-specific training, and videotaped verification as safeguards for intervention fidelity. It mandated three specific safeguards [35]:

- “Active treatment or placebo ... will be administered ... by an experienced neurologist or stroke physician.”
- “The injection will be videotaped, such that correct administration can be verified independently.”
- “The doctors administering the IP [investigational product] will be trained to correctly administer the injection.”

All three safeguards are absent from the ANZCTR-listed protocol paper [33,35].

The trial report states that injectors were “trained according to published methods” but does not identify them as experienced neurologists or stroke physicians and does not mention videotaping [5]. It reports the cervical levels and four-minute Trendelenburg position but omits the midline interspinous approach, needle gauge and length, and degree of incline-details relevant to independent replication of the procedure [23-30,32].

Neither the published record reviewed nor the inquiries described in Methods confirmed that the injector-qualification or videotaping requirements were met at any site.

4.2.4 The site authorization record does not record delegation of investigational-product administration.

ICH E6(R2) requires the investigator to maintain a list of appropriately qualified persons to whom significant trial-related duties have been delegated [9]. For the Austin Health site that record is the Site Signature and Delegation of Responsibilities Log, produced under FOI in two separate productions with differing redactions (SF17) [21]. Its printed Study Task Key lists thirty delegable tasks, of which task 21 is “Administer IP.” Every row records a combination of other tasks, among them the management, preparation and dispensing of the investigational product delegated to pharmacy staff. No row on either production records delegation of task 21 to any person, so the injector's identity and qualifications at

Threats to Validity of the PESTO Trial

that site cannot be established from the record that exists to establish them. The Study Task(s) column is legible and identical across both productions, so the observation does not depend on a single image. It is confined to the Austin Health site; the corresponding logs for the other sites were not available to the authors, and material withheld under the third-party personal-affairs exemption may bear on it.

The absence of the remaining records is not for want of asking, and it is now documented from both directions. In February 2026 the corresponding author requested them from the sponsor in terms framed to avoid any disclosure of personnel information: clinician names could be redacted and replaced by substitute numbers, the video-recording request sought only aggregate counts by site of the participants for whom recordings were and were not made, and the Trial Master File request sought the index rather than the file (SF13) [18]. All seven categories were declined, on the ground that such material constitutes internal operational records not ordinarily made available outside the trial's oversight structures (SF14) [17]. The equivalent categories were requested from the site under freedom-of-information law (SF15) [19]. The internal-review decision of 16 March 2026 stated that making the recordings and transmitting them to the sponsor was the site investigator's responsibility, that their storage, verification, review and management was a sponsor function, that no such governance records were generated or retained at Austin Health, and that Austin Health did not obtain copies from the sponsor (SF16) [20]. The Austin Health HREC, asked whether it could confirm the two safeguards at its own site, replied that it does not investigate third-party queries regarding trial conduct and that operational implementation falls to the sponsor (SF18) [22].

The consequence is structural rather than accusatory. The records that would show who administered the investigational product, with what qualifications and training, and whether the videotaping requirement was met are held by the sponsor, which declines to release them; the site states that it does not hold them and has not sought them; and the committee that approved the protocol treats their implementation as a matter for the sponsor. No party reachable by an external inquirer presently holds itself out as able to answer whether the protocol's own safeguards were implemented.

The resulting concern is both construct-validity and conduct-related: intervention fidelity remains unverified rather than disproven. Only a body with standing to see the operational record can determine whether the mandated safeguards were implemented and appropriately documented.

4.3 Reporting and research-governance standards relevant to the documented concerns

The central questions are whether changes to outcomes and analyses were identified with reasons, whether the personnel and safeguards for delivering a technically demanding intervention were described, and whether the record permits verification of conduct under the approved protocols and site-specific ethics conditions. These requirements predated or were in force when PESTO was conducted, submitted, and published.

4.3.1 Reporting changes to outcomes and analyses

Threats to Validity of the PESTO Trial

CONSORT 2010 Item 3b requires important changes to trial methods after commencement to be reported with reasons, and Item 6b requires the same for changes to trial outcomes [11]. CONSORT 2025 Item 10 consolidates this requirement and expressly includes outcomes or analyses that were not prespecified [12,13]. ICH E9 likewise requires the primary variable and, for a composite variable, the combining algorithm to be specified, with substantive changes documented [10]. Neurology's 2024 RCT-reporting editorial additionally required submission of the signed original and most recent protocol and SAP, together with explanations of substantive changes [34]. PESTO's report was received and published after Neurology announced those requirements [5,34]. Nevertheless, the trial report, registry, protocol paper, and disclosed SAP do not explain the differences in either the primary-outcome definition or the primary-analysis covariate.

The requirement is explicit, it applied to PESTO, and it was not met. Two substantive changes to the prespecified primary outcome and primary analysis are reported without the explanation each of these standards requires, in a trial received and published after its own journal adopted that requirement in terms. This is not an inference about what may have occurred; it is a property of the published record. Nor is it cured by the possibility that an authorizing amendment exists, since the standards require the change and its reasons to be reported, not merely to have been approved somewhere. The consequence for the reported result follows from the same fact. A reader cannot determine from the public record whether the reported primary result corresponds to the prospectively specified outcome and analysis; and because it demonstrably corresponds to neither as disclosed, the result cannot be treated as a confirmatory test of the hypothesis the trial prespecified.

4.3.2 Reporting who delivered the intervention

TIDieR requires trial reports to identify who provided each intervention and to describe relevant expertise and specific training [15]. CONSORT 2025 Item 12b similarly calls for eligibility criteria for intervention providers and the settings in which the intervention was delivered [12,13].

Protocol v2.1 required an experienced neurologist or stroke physician, procedure-specific training, and videotaped verification. The trial report does not identify the protocol-required injector qualifications or report whether videotaping occurred (§4.2.3).

4.3.3 Ethics-approved protocols, amendments, and site conditions

ICH E6(R2), the Declaration of Helsinki applicable during the trial, and the Australian Code for the Responsible Conduct of Research require clinical research to proceed under an approved protocol and ethics oversight, with substantive changes and deviations appropriately documented and reviewed [9,54,55]. Protocol v2.1 itself states that PESTO would comply with all protocol stipulations and the conditions of ethics approval.

The record reviewed does not identify the complete ethics-approved Christchurch protocol, reconcile the New Zealand proxy-consent condition with the principal trial documents, or disclose whether the materially different primary-outcome and analysis specifications were incorporated into a protocol amendment and, where required, reviewed by the relevant oversight bodies. It also does not confirm implementation of the

Threats to Validity of the PESTO Trial

protocol-defined injector-qualification and videotaping safeguards. These are questions of trial conduct and research governance that statistical reanalysis cannot answer and that require independent access to the complete ethics, protocol, monitoring, and site records.

5. Expert opinion

PESTO was randomized, double-blind and placebo-controlled, and it retained primary-outcome data for 124 of 126 randomized participants. Its reported primary estimate was neutral, and serious adverse events were similarly distributed between the arms. These are design strengths, and we do not dispute them [5].

Set against them, the record contains affirmative discrepancies in the primary-outcome definition and the primary-analysis model, together with unresolved verification gaps concerning the Christchurch protocol, the site-specific proxy-consent restriction, and the intervention-fidelity safeguards. Four questions follow. Which approved protocol governed participants at each site, and were the site-specific ethics conditions implemented? Were the changes to the primary outcome and the analysis, made after data collection ended, authorized and disclosed? Was the intervention delivered under the mandated safeguards? And does the numerical conclusion withstand the specifications prospectively disclosed? Table 3 records, for each finding, what the record supports and what it does not. Two require no further inquiry to state: the primary outcome and the primary-analysis model used to generate the published result are not those the trial prespecified, and no disclosed document explains either change.

The Class I designation depends on two predicates the AAN scheme requires: clearly defined inclusion and exclusion criteria, and clearly defined primary outcomes [42]. Neither holds on the record reviewed. The Christchurch protocol is unavailable, so the eligibility criteria governing the pooled population cannot be reconstructed; and the outcome definition analyzed is demonstrably not the one the trial defined in advance. A designation that presupposes clearly defined criteria and outcomes cannot stand on a record that supplies neither. Each deficiency alone precludes the designation. PESTO's own Statistical Analysis Plan describes the trial as phase 2b [38], a stage of development at which a definitive negative conclusion would ordinarily not be drawn.

The stakes attached to the primary endpoint were fixed in advance. In May 2021, mid-recruitment, the funder's chief executive advised the Commonwealth that "if the trial does not achieve the primary endpoint, research should focus on other ways to improve outcomes after disabling stroke" (SF7). The primary endpoint to which that consequence was attached was the endpoint then prespecified: "an achievement of at least 5 points improvement on the SF-36 score," as the signed protocols and the registration defined it. The three-route construct that ultimately adjudicated it did not yet exist.

We do not ask the reader to disregard our conflict of interest. The corresponding author originated the perispinal etanercept treatment method and has a direct and continuing interest in it; the interests of both authors are declared in full following the conclusion. We ask readers to set aside the favorable observational - but not randomized - PSE chronic stroke results that have been published; considering them is unnecessary to an objective analysis of PESTO's conduct and reporting.

Threats to Validity of the PESTO Trial

What is required is an independent multidisciplinary review rather than reanalysis by the original trial team alone. The panel should combine independent statisticians, clinical-trial methodologists, research ethicists, stroke specialists and expertise in the protocol-defined intervention, and no member should have had any role in PESTO or any financial or professional interest in the treatment. The corresponding author sought access to the trial data himself; given the interests declared above, we have concluded that a reanalysis will carry weight only if conducted by people with no stake in its outcome, and we therefore no longer seek that access, nor any role in the review or in nominating its members. The panel should receive the participant-level data and code, all protocol versions and amendments, ethics submissions and approvals, consent materials, site-level enrollment records, and injector qualification, training, videotaping, monitoring and deviation records.

An established mechanism already exists, and we have used it. PESTO's registry entry states that the trial will consider sharing de-identified participant data on stated conditions, and expressly lists studies testing whether findings can be repeated or confirmed among the permitted purposes (SF12) [37]; the trial report's data-availability statement is to the same effect [5]. The corresponding author requested the dataset and seven categories of governance documentation (SF13) [18], and the reply of 11 March 2026, from the principal investigator on behalf of the PESTO Steering Committee, is filed as SF14 [17].

The reply is open to considering access to the dataset subject to a data-sharing agreement, Steering Committee approval, a detailed analysis plan with power analysis, institutional affiliation and ethics approval, and public pre-registration before access. We do not dispute those conditions; a plan fixed and registered before access is precisely what is proposed here. Two features are nonetheless material. Any resulting manuscript would have to be given to the PESTO investigators for review before submission. And all seven categories of governance documentation were declined as internal operational records [17].

That answer marks the limits of what any external party can establish. The records that would show which protocol governed enrollment at Christchurch, who administered the injections and with what training, and whether the videotaping requirement was met are the same records that fall outside the data-sharing arrangement. We do not suggest this position is irregular; it may reflect ordinary practice. It does mean that the questions set out in §§4.1 and 4.2.3 cannot be resolved by a data request, and that answering them needs a body with standing to see the operational record. That is why the research-ethics component is not an optional adjunct to the statistical one.

It is worth naming those bodies, because the question is otherwise left in the air. Four have standing an external inquirer does not. The sponsor holds the operational record, and its own research-integrity function may examine it. The approving institution has a formal process for concerns about research conducted under its approval, and a concern has been lodged through it; the Austin Health HREC has stated that operational implementation is a matter for the sponsor rather than for the committee (SF18) [22]. The Australian Code for the Responsible Conduct of Research provides for review of an institution's handling of such concerns where its own process does not resolve them [55]. And PESTO was publicly funded, which gives the funding body a conventional interest in whether the trial was

Threats to Validity of the PESTO Trial

conducted under the protocol it approved. We ask none of them to reach a particular conclusion, and advance no allegation for any of them to adjudicate. The point is narrower and, we think, unavoidable. Whether a publicly funded trial was conducted under the protocol approved for it is a question the record now shows cannot be answered by request from outside. The mechanisms capable of answering it exist and are ordinary ones. They will answer it only if one of them asks.

The issue is not merely that the public record fails to supply an amendment history. PESTO's own trial report states that the first and "most recent" signed protocols were supplied with the publication. The most recent protocol so supplied is version 2.1, dated 31 May 2022, and it expressly prespecifies the ≥ 5 -point SF-36 endpoint and adjustment for baseline SF-36 and proxy-versus-self administration. On the trial investigators' own published account, therefore, no later signed protocol superseded version 2.1 before the SAP adopted the materially different three-route responder definition and substituted time since stroke as a covariate. Those specifications departed from the most recent signed protocol identified by the trial report; no disclosed protocol amendment authorizes either change.

It may be said that a trial's result cannot be called unreliable merely because records are absent, and we agree. Nor do we claim the reported estimate is arithmetically wrong; whether it would change under the prespecified endpoint and covariates is what the sensitivity analyses specified below will determine, and we say nothing about their outcome in advance. Our claim is a measurement claim. A result offered as Class I evidence derives its authority from two sources: specification before the data were seen, and measurement by an instrument definition of established validity and reliability. The result PESTO reported has neither. The definition that produced it appears in no operative protocol or registry entry; it is a novel PESTO-devised construct; it is not an accepted SF-36 scoring pathway; and no evidence of its validity, reliability, responsiveness, or meaningful-change threshold has been published (§§4.2.1–4.2.2). Validity in this context is demonstrated before use, not presumed until disproven. On the published record, PESTO's published primary result is therefore unreliable as a measurement of quality-of-life response, and its definitive negative conclusion cannot be relied upon as an answer to the question the trial posed. Under the present terms of data access, no one outside the trial team can establish it.

The sensitivity analyses specified below remain necessary, and they carry real evidentiary weight, because the endpoint they apply predates the data. But they will be performed with full knowledge of the published results, and their reach runs in one direction. If the prespecified analysis contradicts the published conclusion, the conclusion falls. If it agrees, the agreement might support a corrected report, if the additional issues extant are satisfactorily resolved; but it cannot retrospectively convert an analysis that departed from the most recent signed protocol into the prospectively specified confirmatory analysis.

The statistical component should fix its analysis plan before receiving unblinded allocation codes. It should report how many participants qualified through each of the three routes and their overlap; apply the endpoint specified in the signed protocols and the registry so far as contemporaneous records allow; repeat the primary analysis under both covariate

Threats to Validity of the PESTO Trial

models and under the trial's three randomization stratification factors; and report whether responder classification, effect estimates or the substantive inference change across those analyses.

The research-ethics component has a different task. It should determine which approved protocol governed each site during each period of enrollment; whether enrollment at Christchurch complied with the proxy-consent prohibition; whether the outcome and analysis changes required, and received, protocol amendment, sponsor authorization and ethics approval; whether the intervention-fidelity safeguards were implemented and any deviations recorded; and whether the publication accurately represents the protocols, consent conditions, amendments and procedures under which the trial was conducted.

Three determinations can be made now, from a small number of documents. Whether an approved amendment exists for the changes to the primary-outcome definition and the primary-analysis covariate, and if so when and by whom it was approved: that requires the amendment file and the corresponding ethics submissions. Which approved protocol governed enrollment at Christchurch in each period: that requires the Southern HDEC approval file. And whether the operational record supports the trial report's account of how the intervention was delivered - that injectors were "trained according to published methods" - together with the protocol's qualification and videotaping requirements: that requires the delegation logs for all three sites, the training records and the video-recording governance records. The first of these is a single page per site, and the Austin Health page has already been produced; what it does not record is set out at §4.2.3. Each is a documentary question with a determinate answer, settleable well before the statistical work is complete.

What follows should turn on those answers. If the operative protocol was amended and the amendment approved, the specification changes are a failure of disclosure, and disclosure corrects them. If no such amendment exists, the analysis reported departed from the most recent signed protocol - the protocol the trial report identifies as operative and supplies with the paper - and that departure is established on the face of the record [35,38]. The intervention question resolves the same way: if the operational record substantiates the published account, that issue closes; if it does not, the publication has misdescribed the conduct of the trial. In either case the first response is correction of the published record and referral to the bodies identified above, which have authority to determine how the departure occurred and how serious it was. What is not open, on the present record, is to leave the reported conclusion standing unqualified while the documents that would support it remain unproduced.

The panel's methods, documentary findings, analytic code, statistical results and ethics-governance conclusions should be made public, subject to protecting participant confidentiality. Two remedies do not depend on the outcome of that review, and should not wait for it. The published record should be corrected to disclose the changes made to the primary-outcome definition and the primary-analysis model and the reasons for them, as the applicable reporting standards require. And the Class I designation, which presupposes clearly defined eligibility criteria and clearly defined outcomes that the record does not

Threats to Validity of the PESTO Trial

supply, should be reclassified by those responsible for it. Both follow from facts already established on the published record and require no further inquiry.

Whether any remedy beyond correction and reclassification is required is a different question, and it is for independent editorial review informed by the panel's findings. Retraction is warranted where a review establishes that the publication's principal conclusion cannot be relied upon, or that the trial was conducted outside the terms of its ethics approval [56]. That review may be convened by any of the bodies identified above; it need not await action by the publishing journal. We do not ask for that determination to be made in advance of the review that should make it.

6. Five-year view

The question perispinal etanercept poses remains open. A trial capable of closing it would be adequately powered, with outcomes specified unambiguously before data collection ends, an analysis plan and every protocol amendment publicly accessible, and a documented record of who administered the injections and how fidelity was assured.

7. Conclusion

PESTO's design strengths include its randomization, double-blinding, placebo-control and high primary-outcome follow-up. They do not eliminate the need to reconcile the trial report with the prospectively disclosed protocol, registry and ethics record. The record obtainable does not permit verification of the ethics-approved protocol governing Christchurch, implementation of the New Zealand proxy-consent restriction, or compliance with the three mandated intervention-fidelity safeguards; the site authorization record produced for the Austin Health site records no delegation of investigational-product administration, and the records that would resolve the question are held by the sponsor, which has declined to release them. Independently, the primary-outcome definition and primary-analysis model used for the reported result differ from those specified in the signed protocols and registry without public reconciliation.

These findings establish that the reported primary result was not produced by the outcome definition or the analytic model PESTO prespecified, that no disclosed document explains either change, and that the reporting standards applicable to the trial required both to be explained. They also establish that the questions raised cannot be resolved by a single sensitivity analysis or by statistical reasoning alone. They bear both on whether the reported result withstands the prespecified specifications and on whether the trial was conducted, governed, amended, and reported in a manner consistent with its approved and publicly represented specifications. Disclosure of the complete trial record and an independent multidisciplinary review with separate statistical and research-ethics components are required by the affirmative discrepancies and unresolved verification gaps identified here. Until that disclosure and review occur, PESTO's definitive negative conclusion and its Class I designation cannot be relied upon as statements of what the trial established: the result reported was produced by a novel, unvalidated responder definition that is not an accepted SF-36 scoring pathway, and the analysis the trial prespecified has never been reported. Whether the reported estimate would survive that analysis is a question the record cannot yet answer, and we do not answer it here.

Threats to Validity of the PESTO Trial

Statistical reanalysis should examine the three responder routes and both covariate models; ethics review should determine the governing site protocols at all three sites, the amendment and approval history, compliance with the intervention-fidelity safeguards, and the accuracy of the published account, beginning with who performed the injections and with what qualifications. On the published record, correction of PESTO's record and its conclusions and reclassification of its Class I evidence designation are due now and depend on nothing further. Whether any remedy beyond those is required is for independent editorial review, informed by the panel's findings.

Declaration of interest

Edward Tobinick is the originator of the perispinal etanercept treatment discussed in this article, uses it at his private clinic, holds multiple U.S. and international patents that cover the procedure, and receives royalties from those patents, including from the RN Clinics in Denmark. Güven Gürbüz is a fully certified neurologist and uses perispinal etanercept as an off-label treatment at the RN Clinics. We have acknowledged this as a limitation. While no independent investigator without financial interest participated in the analysis, bias was minimized through transparent reporting, explicit documentation of our interpretive approach, and direct access to the primary sources to allow readers to assess our conclusions independently. The authors have no other competing interests. The authors received no funding for this work. Dr. Tobinick's clinic paid the fees associated with publication of this article. This evaluation rests entirely on publicly available publications, Dr. Tobinick's personal correspondence, and documents obtained under freedom-of-information law. Both authors contributed to the drafting of the manuscript and take full responsibility for its content.

Funding

This paper was not funded.

Acknowledgments

Assistance in the preparation of this manuscript by Genavieve Prieto, Danielle Ucci, Samantha Asseraf, and Kamryn Adams is acknowledged. Declaration of generative AI use: during the preparation of this manuscript the authors used Claude (Opus 5; Anthropic), ChatGPT (GPT-5.5; OpenAI), and Gemini (3.1 Pro; Google) to assist with drafting and editing text, organizing the manuscript's structure, and analyzing and cross-referencing the source documents on which the evaluation relies. The corresponding author directed all such use and verified all factual content and citations against the primary source documents. No generative-AI tool was used to generate, alter, or analyze research data, images, or figures, and no AI tool is listed as an author. The authors take full responsibility for the accuracy, originality, and integrity of the work. Data availability statement: there is no original data in this review.

Table Legends

Table 1: Summary of the key documentary record. Documents and events are listed in chronological order to provide the evidentiary context for subsequent analyses of protocol

Threats to Validity of the PESTO Trial

development, ethics approvals, trial registration, statistical analysis, reporting, and data sharing correspondence.

Table 2: Summary of the documentary evidence relating to ethics governance and implementation at the New Zealand study site. For each item, the table indicates what the record reviewed supports and what remains unverifiable from the available documentation.

Table 3: Summary of the principal concerns arising from the documentary record, the evidentiary status for each, what the record supports and does not support, and the resulting implications for independent review and interpretation.

Table 4: Summary of the cross-document comparison of the primary-outcome definition and primary-analysis covariate in the PESTO documentary record. Including tracking the evolution of these specifications across trial documents and highlighting any discrepancies relevant to the reported primary analysis.

Reference Annotations

Papers of special note have been highlighted as either of interest (*) or of considerable interest (**) to readers.

For Ref [13], Hopewell: * Item 10 requires important changes after trial commencement, including outcomes or analyses that were not prespecified, to be reported with reasons.

For Ref [11], Schulz: * Items 3b and 6b establish that the obligation to report changes to methods and to trial outcomes long predates PESTO.

For Ref [35], Thijs: ** The signed protocol identified by the trial report as the most recent; the reference standard for the outcome definition, the analysis model, and the intervention-fidelity safeguards.

For Ref [37], ANZCTR: * *The registry entry, which retains the superseded outcome and analysis specifications and records the trial's commitment to consider sharing individual participant data for confirmatory reanalysis.*

For Ref [38], Thijs: ** The statistical analysis plan, released only with the trial report, in which both the primary-outcome definition and the primary-analysis model differ from the signed protocols.

For Ref [42], Gross: ** Sets out the evidence-classification criteria, including clearly defined inclusion and exclusion criteria and a clearly defined primary outcome, on which the Class I designation depends.

For Ref [43], Southern HDEC: ** The ethics-committee minutes recording the New Zealand prohibition of proxy consent; the central research-governance document of this evaluation.

For Ref [34], Ziegler: ** Establishes that the transparency requirements at issue were the publishing journal's own, in force while PESTO was under editorial consideration.

References

Threats to Validity of the PESTO Trial

- [1] Liu S, Li Y, Lan X, et al. Global, regional, and national trends in ischaemic stroke burden and risk factors among adults aged 20 + years (1990-2021): a systematic analysis of data from the Global Burden of Disease study 2021 with projections into 2050. *Front Public Health*. 2025;13:1567275.
- [2] Lee TH, Uchiyama S, Kusuma Y, et al. A systematic-search-and-review of registered pharmacological therapies investigated to improve neuro-recovery after a stroke. *Front Neurol*. 2024;15:1346177.
- [3] Okabe N, Wei X, Abumeri F, et al. Parvalbumin interneurons regulate rehabilitation-induced functional recovery after stroke and identify a rehabilitation drug. *Nat Commun*. 2025;16(1):2556.
- [4] Stroke Foundation [Australia] Media Release, "Return to Life, Return to Work research grants announced," 18 March 2019. [Internet]. 18 March 2019
- [5] Thijs VN, Cloud GC, Gilchrist N, et al. Safety and Efficacy of Perispinal Etanercept for Chronic Stroke: A Randomized Clinical Trial. *Neurology*. 2025;105(6):e213981.
- [6] Return to life, return to work Clinical Research Grants Application Guide. Stroke Foundation; 2018.
- [7] Garattini S, Jakobsen JC, Wetterslev J, et al. Evidence-based clinical practice: Overview of threats to the validity of evidence and how to minimise them. *Eur J Intern Med*. 2016;32:13–21.
- [8] FDA. International Council for Harmonisation (ICH) E3 (R1) Structure and Content of Clinical Study Reports: Questions and Answers. 2013.
- [9] FDA. ICH E6(R2) Good Clinical Practice: Integrated Addendum to ICH E6(R1)
- [10] ICH E9. Statistical Principles for Clinical Trials, European Medicines Agency, September 1998.
- [11] Schulz KF, Altman DG, Moher D, et al. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *PLoS Med*. 2010;7(3):e1000251.
- [12] Hopewell S, Chan AW, Collins GS, et al. CONSORT 2025 statement: Updated guideline for reporting randomised trials. *PLoS Med*. 2025;22(4):e1004587.
- [13] Hopewell S, Chan AW, Collins GS, et al. CONSORT 2025 explanation and elaboration: updated guideline for reporting randomised trials. *BMJ*. 2025;389:e081124.
- [14] Hrobjartsson A, Boutron I, Hopewell S, et al. SPIRIT 2025 explanation and elaboration: updated guideline for protocols of randomised trials. *BMJ*. 2025;389:e081660.
- [15] Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ*. 2014;348:g1687.
- [16] Ralph SJ, Weissenberger A, Bonev V, et al. Phase I/II parallel double-blind randomized controlled clinical trial of perispinal etanercept for chronic stroke: improved mobility and pain alleviation. *Expert Opin Investig Drugs*. 2020:1–16.
- [17] Reply to Request for PESTO Data Sharing, 11 March 2026.
- [18] Letter to Florey for PESTO Data Sharing, 17 February 2026.
- [19] Austin Health FOI Request, 11 February 2026.
- [20] Austin Health Internal Review Decision, produced 16 March 2026 in response to an FOI request dated 11 February 2026.
- [21] Site Signature and Delegation of Responsibilities Log prepared 16 March 2026, produced under FOI 25/26-1266 request.

Threats to Validity of the PESTO Trial

- [22] Austin Health HREC Clarification Request Response, produced on 30 March 2026 in response to a request for clarification of HREC complaint process dated 29 March 2026.
- [23] Breschet G. Essai sur les veines du rachis [Theses presentees et soutenues publiqu. devant les juges concours le 28. Avril 1819]. Paris: Faculte de Medecine de Paris; 1819.
- [24] Breschet G. Recherches anatomiques physiologiques et pathologiques sur le syst eme veineux. Paris,: Rouen fr eres; 1829.
- [25] Batson OV. The Function of the Vertebral Veins and Their Role in the Spread of Metastases. *Ann Surg.* 1940;112(1):138–49.
- [26] Batson OV. The vertebral vein system. Caldwell lecture, 1956. *Am J Roentgenol Radium Ther Nucl Med.* 1957;78(2):195–212.
- [27] Tobinick E. The cerebrospinal venous system: anatomy, physiology, and clinical implications. *MedGenMed.* 2006;8(1):53.
- [28] Nathoo N, Caris EC, Wiener JA, et al. History of the vertebral venous plexus and the significant contributions of Breschet and Batson. *Neurosurgery.* 2011;69(5):1007–14; discussion 1014.
- [29] Griessenauer CJ, Raborn J, Foreman P, et al. Venous drainage of the spine and spinal cord: a comprehensive review of its history, embryology, anatomy, physiology, and pathology. *Clin Anat.* 2015;28(1):75–87.
- [30] Carpenter K, Decater T, Iwanaga J, et al. Revisiting the Vertebral Venous Plexus-A Comprehensive Review of the Literature. *World Neurosurg.* 2021;145:381–395.
- [31] FDA. Enrichment Strategies for Clinical Trials to Support Determination of Effectiveness of Human Drugs and Biological Products, Guidance for Industry, FDA Document FDA-2012-D-1145-0040, available online at <https://www.regulations.gov/document/FDA-2012-D-1145-0040>. Silver Spring, Maryland: U.S. Food and Drug Administration, Center for Biologics Evaluation and Research; 2019 March 15.
- [32] Tobinick E, Ucci D, Bermudo K, et al. Perispinal etanercept stroke trial design: PESTO and beyond. *Expert Opin Biol Ther.* 2024;24(10):1095–1108.
- [33] Thijs V, Cloud GC, Gilchrist N, et al. Perispinal Etanercept to improve STroke Outcomes (PESTO): Protocol for a multicenter, international, randomized placebo-controlled trial. *Eur Stroke J.* 2024;9(3):789–795.
- [34] Ziegler A, Eliasziw M, Howard VJ, et al. New Requirements in the Reporting of Randomized Controlled Trials Published in Neurology to Foster Greater Transparency. *Neurology.* 2024;103(7):e209909.
- [35] Thijs V. Protocol v2.1 - A Prospective multicentre, phase 2b randomised controlled double-blind trial, to determine the safety and efficacy of perispinal etanercept on quality of life at 28 days post treatment PESTO, Protocol Number 2, Version: 2.1 Date: 31 May 2022 Author: Thijs, V.N.. Sponsor: Florey Institute of Neuroscience and Mental Health. Published as online supplemental material within the file esap_1, along with the corresponding Neurology® trial report, Volume 105, Issue 6, 4 Sept 2025. [Clinical Trial Protocol].
- [36] Thijs V. PESTO Protocol No.1, version 1.0 (30 May 2019) - A Prospective multicentre, phase 2b randomised controlled double-blind trial, to determine the safety and efficacy of perispinal etanercept on quality of life at 28 days post

Threats to Validity of the PESTO Trial

- treatment Author: Thijs, V.N.. Sponsor: Florey Institute of Neuroscience and Mental Health. Published as online supplemental material within the file esap_1, along with the corresponding Neurology® trial report, Volume 105, Issue 6, 4 Sept 2025.
- [37] ANZCTR: Australian New Zealand Clinical Trials Registry Page for the PESTO perispinal etanercept stroke trial, registered 7 October 2020, last updated 28 Sept 2025 [Internet]. [cited 2026/01/31]. Available from: <https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=378447&isReview=true>.
- [38] Thijs V, Churilov L. ESAP_1, including PESTO Protocol v1.1(2019), 2.1(2022) and the Statistical Analysis Plan (SAP, 2024) [Online Supplementary Material]. Neurology. 2025;105(6).
- [39] Garcia-Perez MA. Statistical conclusion validity: some common threats and simple remedies. Front Psychol. 2012;3:325.
- [40] Rankupalli B, Tandon R. Practicing evidence-based psychiatry: 1. Applying a study's findings: The threats to validity approach. Asian J Psychiatr. 2010;3(1):35–40.
- [41] Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". Lancet. 2005;365(9453):82–93.
- [42] Gross RA, Johnston KC. Levels of evidence: Taking Neurology to the next level. Neurology. 2009;72(1):8–10.
- [43] Southern Health and Disability Ethics Committee, New Zealand, 08 Sept 2020 minutes, 20/STH/139 Perispinal Entanercept [sic] in post stroke patients with moderate to severe disability.
- [44] Southern Health and Disability Ethics Committee, 2021 Annual Report, page 12, 20/STH/139 Approval on 11 December 2020.
- [45] DSMB Report - "PESTO SAFETY REPORT for Data Safety Monitoring Board," prepared 1-June-2022, produced under FOI 26-1740 request, on 1 July 2026.
- [46] Letter 12 May 2021 from the CEO of the Stroke Foundation to the Department of Industry, Science, Energy and Resources, produced under FOI 26-2767 request to the Australian Government Department of Health, Disability and Ageing, Document 25, entitled "RE: MRFF-N-SF A17.04 Accelerated Research Program: 'Return to life, return to work'"; available online by searching: <https://www.health.gov.au/resources/foi-disclosure-log/foi-26-2767-clinical-trial-grant-funding>.
- [47] Lins L, Carvalho FM. SF-36 total score as a single measure of health-related quality of life: Scoping review. SAGE Open Med. 2016;4:2050312116671725.
- [48] Ware JE, Jr. SF-36 health survey update. Spine (Phila Pa 1976). 2000;25(24):3130–9.
- [49] Hobart JC, Williams LS, Moran K, et al. Quality of life measurement after stroke: uses and abuses of the SF-36. Stroke. 2002;33(5):1348–56.
- [50] Taft C, Karlsson J, Sullivan M. Do SF-36 summary component scores accurately summarize subscale scores? Qual Life Res. 2001;10(5):395–404.
- [51] Ware JE, Jr., Maruish M, Editor. User's manual for the SF-36v2 Health Survey. Quality Metric Incorporated. 2011.
- [52] Frendl DM, Ware JE, Jr. Patient-reported functional health and well-being outcomes with drug therapy: a systematic review of randomized trials using the SF-36 health survey. Med Care. 2014;52(5):439–45.

Threats to Validity of the PESTO Trial

- [53] Calvert M, Brundage M, Jacobsen PB, et al. The CONSORT Patient-Reported Outcome (PRO) extension: implications for clinical trials and practice. *Health Qual Life Outcomes*. 2013;11:184.
- [54] WMA. WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Participants. The World Medical Association. 2024.
- [55] Australian Code for the Responsible Conduct of Research.
<https://www.nhmrc.gov.au/about-us/publications/australian-code-responsible-conduct-research-2018>.
- [56] COPE. Committee on Publication Ethics (COPE): *Retraction Guidelines*. August 2025.