Stem Cell Research and Therapeutics
SCRT, 9(1): 233-240
www.scitcentral.com

ISSN: 2474-4646

Research Article: Open Access

Coding Before Pooling: A Coding Manual and Reporting Checklist for Trials, Registries and Systematic Reviews in Regenerative and Energy-Based Gynecology

Mônica Andréa Probst[1*], Márcio Hiroaki Kume[2], Edson Peracchi[2] and Carmen Austrália Paredes Marcondes Ribas[3]

[*1]*Sugisawa Hospital, Department of Regenerative Medicine, Curitiba, Brazil*

[2]*CeUnina, Department of Biologic Science, Curitiba, Brazil*

[3]*Mackenzie University, Curitiba, Brazil.*

Received October 19, 2026; Revised October 22, 2026; Accepted October 23, 2026

ABSTRACT

Background. Studies of platelet-rich plasma (PRP), vaginal lasers, radiofrequency and other regenerative or energy-based procedures in gynecology are frequently impossible to compare: the treated tissue, the patient phenotype, the product or device dose and the primary outcome are reported inconsistently or not at all. A five-axis classification for this field has been proposed separately as a position paper (the GRACE framework). A classification is only useful, however, if different people apply it in the same way.

Objective. To provide an operational coding manual and a reporting checklist that allow investigators, trial registries, systematic reviewers and editors to apply the five-axis classification reproducibly.

Content. The manual sets out a six-step coding algorithm, ten decision rules for ambiguous cases (mixed phenotypes, multiple compartments, combination therapy, unreported descriptors, mixed comparators and others), an 18-item reporting checklist organised by manuscript section, a structured extraction form for double independent coding, pooling and sensitivity-analysis rules for meta-analysis, three quantitative indicators of codability, and standard wording for patient information at each evidence-maturity level. Three fictitious study descriptions are coded step by step to show how the rules operate.

Conclusions. This document does not grade any intervention; it describes how to code a study and how to report one so that it can be coded. It is intended to be tested in the inter-rater reliability and retrospective recoding phases of the GRACE validation programme and to be revised with the results.

Keywords: regenerative gynecology; reporting guideline; coding manual; systematic review methods; platelet-rich plasma; vaginal laser; inter-rater reliability; evidence synthesis.

Why a coding manual is needed?

Regenerative and energy-based gynecology has a reporting problem before it has an efficacy problem. In our umbrella review of regenerative and energy-based interventions in gynecology, the same structural defects recurred across platelet products, lasers, radiofrequency and reproductive applications: comparators of very different rigour were pooled together, surrogate markers replaced patient-important outcomes, overlapping reviews counted the same trials repeatedly, and interventions were described too vaguely to be reproduced [2]. Our umbrella review restricted to PRP for vulvovaginal disorders illustrated the scale of the issue: across nineteen in-scope reviews, deduplication of the primary studies yielded markedly fewer participants than the

Corresponding author: Mônica Andréa Probst, 80250-190, Iguassu Avenue, 1236, doctorkume@yahoo.com, ORCID 0009-0009-2014-0557.

Citation: Probst MA, Kume MH, Peracchi E & Ribas CAPM. (2026) Coding Before Pooling: A Coding Manual and Reporting Checklist for Trials, Registries and Systematic Reviews in Regenerative and Energy-Based Gynecology. Stem Cell Res Ther, 9(1): 233-240.

Copyright: ©2026 Probst MA, Kume MH, Peracchi E & Ribas CAPM. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI: https://doi.org/10.5281/zenodo.23762067

reviews collectively reported, only a small minority of trials were sham- or placebo-controlled, and few reviews applied a named risk-of-bias tool [3].

The GRACE framework, reported separately, responds to these defects by coding each application along five axes — anatomical compartment (C), clinical-biological phenotype (P), intervention with mandatory descriptors (I), evidence maturity (M1 to M4) and anchor outcome (O) — and by attaching evidence maturity to the compartment–phenotype–intervention triplet rather than to a technology [1]. The framework defines the categories. It does not, by itself, tell a reviewer what to do when a trial enrols women with both atrophic and pain-dominant symptoms, treats vulva and vagina in the same session, combines laser with PRP, or reports platelet counts only for whole blood.

Those decisions are where classification systems lose reliability. Generic reporting guidelines — CONSORT 2025 for randomised trials [6], SPIRIT 2025 for protocols [7], TIDieR for intervention description [8] and PRISMA 2020 for systematic reviews [9] — remain necessary, but none of them asks for a platelet-product code, an energy dose or a phenotype. These manual fills that operational gap. It is written for four users: investigators writing protocols and papers, registries and ethics committees reviewing them, systematic reviewers extracting data, and editors and peer reviewers assessing submissions.

Scope and relationship to the framework

The manual applies to any clinical study, registry entry or systematic review of a regenerative, biologic, energy-based, hormonal or standard intervention for a gynecologic, urogynecologic, sexual-health or reproductive indication. Hormonal and standard non-hormonal therapies are included because they are the reference comparators against which regenerative claims must be judged [1,4].

All category definitions are taken from the GRACE framework [1] and are not repeated in full here. For the genitourinary syndrome of menopause (GSM), the phenotype axis maps onto the VAGINA-R classification, reported separately with its validation protocol [5]. Table 1 is a one-page reference card of the codes used in this manual.

Table 1. Reference card of GRACE codes used in this manual (definitions in the framework paper [1]).

Axis	Codes
C — compartment	VU vulva and vestibule; VG vagina; UB urethra and bladder neck; PF pelvic floor; EN endometrium and uterine cavity; OV ovary
P — phenotype	ATR atrophic–hypoestrogenic; ALG algesic–inflammatory; URO genitourinary–urinary; INT compromised tissue integrity; NMS neuromuscular–sexual; MIC microbial alteration; DER dermatologic; STR structural–support; REP-O reproductive–ovarian; REP-E reproductive–endometrial; modifier r (hormonal risk or reserve)
I — intervention	S1–S3 standard non-hormonal; H1–H2 hormonal; B1 platelet-derived; B2 adipose-derived; B3 cell-based or cell-derived; B4 biomaterials; E1 ablative fractional laser; E2 non-ablative laser; E3 radiofrequency; E4 photobiomodulation; E5 other energy; [NR] descriptor not reported
M — maturity	M1 endorsed; M2 signal-positive, not endorsed (M2b lower position); M3 experimental; M4 preclinical
O — anchor outcome	T tissue; F function; S patient-centred symptom; R reproductive (live birth); A adverse events; follow-up suffix, for example -12m

The six-step coding algorithm

Each study, or each study arm when arms differ, is coded in the following fixed order. The order matters: later steps depend on earlier ones, and coding the intervention before the phenotype reproduces the technology-first reasoning the framework is designed to avoid.

Step 1 — Identify the unit of coding. The unit is the comparison, not the publication. A three-arm trial comparing PRP, laser and estrogen generate two codable comparisons against the reference arm. Secondary analyses of the same trial are linked to the parent record and are not coded as independent evidence.

Step 2 — Code the compartment. Record every compartment that received the intervention, in the order of the stated primary target (for example, VG+VU). Code from the methods section, not from the title: a study called “vaginal

rejuvenation" that injects the vestibule and introitus is VU+VG.

Step 3 — Code the phenotype. Record the primary phenotype from the eligibility criteria and baseline assessment, then up to two secondary phenotypes after '>', then the modifier r if any participants have hormone-dependent cancer, aromatase-inhibitor therapy, premature or surgical menopause, or another documented reason that restricts hormonal therapy. If eligibility is defined only by a symptom ("vaginal dryness") without a mechanism, code the most plausible primary phenotype and add the flag [P?] (phenotype inferred).

Step 4 — Code the intervention and its descriptor. Record the Axis I class and the full descriptor in brackets: for platelet products, a PAW [10] or PLRA [11] code for composition and DEPA [12] for dose and purity, plus volume per site, activation, number of sessions and interval; for energy devices, the dose parameters listed in the framework. Where any mandatory element is missing, append [NR] and list the missing elements.

Step 5 — Record the comparator and design. Although the comparator is not one of the five axes, it determines how the study may be pooled and whether the ceiling rules apply. Use one of five comparator labels: SH (sham or placebo, double-blind), AC (active comparator, specify the GRACE intervention code), WL (wait-list or no treatment), BA (before–after, no comparator) and OB (observational comparison).

Step 6 — Code the outcomes and follow-up. Record the anchor outcome required for the primary phenotype [1] and whether the study measured it as primary, secondary or not at all; then the longest follow-up with at least 70% of participants assessed.

The maturity level (Axis M) is not assigned by the person coding an individual study. Maturity is a property of the body of evidence for a triplet and is assigned at the synthesis stage, applying the framework's sham, surrogate, characterisation and redundancy rules [1].

Decision rules for ambiguous cases

Table 2 lists ten situations that, in our experience with the umbrella reviews [2,3], account for most disagreements between coders.

Table 2. Decision rules for ambiguous coding situations.

No.	Situation	Rule
D1	Participants with different primary phenotypes in one arm	Code the phenotype of the majority if at least 70% share it; otherwise code MIXED and exclude from phenotype-specific pooling
D2	Phenotype stated only as a symptom	Infer from eligibility and baseline measures; add [P?]; inferred phenotypes are analysed separately in sensitivity analysis
D3	More than one compartment treated	List all; pool only with studies treating the same set of compartments, or report separately
D4	Combination therapy (for example, E1+B1)	Code the combination as a distinct intervention; it does not inherit the maturity of either component
D5	Background therapy in both arms (for example, vaginal estrogen)	Record as co-intervention after a semicolon (for example, B1[…]; co-H1); the comparison is coded as add-on
D6	Platelet count reported for whole blood only	Code [NR] for composition; do not estimate the product concentration from device manufacturer claims
D7	Energy settings given as a manufacturer "programme" name	Code [NR] unless the numerical parameters are retrievable from the paper, protocol or supplementary material
D8	Sham procedure not credible (for example, device off but audible difference)	Downgrade comparator from SH to AC-unblinded and record the reason
D9	Reproductive study reports clinical pregnancy but not live birth	Code R as not measured; clinical pregnancy is recorded as a secondary outcome only
D10	Follow-up shorter than 12 months	Record actual follow-up (for example, -3m); the study remains codable but cannot support an upgrade of maturity

Two principles underlie these rules. First, uncertainty is recorded rather than resolved by assumption: an inferred phenotype or an unreported descriptor is flagged, not guessed. Second, nothing in a study's title, abstract or marketing language is used for coding if the methods section says otherwise.

The reporting checklist

The checklist (**Table 3**) is designed to be completed by authors at submission and by registries at protocol review. It is complementary to CONSORT 2025, SPIRIT 2025 and TIDieR [6–8], not a substitute. Items are grouped by manuscript section so that they can be checked in reading order. A study that satisfies items 4–12 can be fully coded on all five axes.

Table 3. GRACE reporting checklist (18 items).

Section	Item	Description
Title and abstract	1	Clinical entity named (for example, "atrophic GSM"), not a marketing term ("rejuvenation", "revitalisation")
	2	Full or abbreviated GRACE code stated in the abstract
Introduction	3	Current maturity of the studied triplet declared, with source
Methods — participants	4	Compartment(s) treated
	5	Primary and secondary phenotypes, with the instruments and thresholds used to assign them
	6	Presence and nature of the hormonal-risk modifier r
Methods — intervention	7	Axis I class and full descriptor (product code or energy dose)
	8	Preparation, operator training and setting
	9	Co-interventions and background therapy in each arm
Methods — comparator	10	Comparator type (SH, AC, WL, BA, OB) and sham procedure description
	11	Blinding of participants, operators and outcome assessors
Methods — outcomes	12	Anchor outcome for the primary phenotype, prespecified as primary or co-primary, with MCID
	13	Planned follow-up of at least 12 months, or justification if shorter
	14	Adverse-event definitions and grading system (for example, CTCAE [13])
Results	15	Anchor-outcome results at each time point, with denominators
	16	All adverse events by severity, including procedural pain, bleeding, scarring and new chronic pain
Discussion	17	Statement of whether the results can move the triplet's maturity level, and why
Other	18	Registration number, protocol availability, funding and device or product conflicts of interest

Structured extraction form for double independent coding

For systematic reviews and for the reliability phase of the validation programme, each record should be coded independently by two people using the same form (**Table 4**), with disagreements resolved by discussion and, when needed, a third coder. The form records the source of each coded element (page, table or supplementary file) so that disagreements can be traced.

Table 4. Minimum fields of the extraction form.

Field	Content
Record identifiers	Citation; registration number; parent trial if secondary analysis
Comparison	Arms compared; unit of coding (Step 1)
C	Compartment code(s); source location
P	Primary and secondary phenotype; modifier r; [P?] flag; instruments
I	Class; descriptor elements present; elements missing; [NR] flag
Comparator	SH, AC, WL, BA or OB; blinding; credibility concerns (D8)
O	Anchor outcome measured (primary, secondary, not measured); follow-up; attrition at final follow-up
Adverse events	Reported (yes/no); grading system
Decision rules applied	D1–D10, with free-text justification
Coder and date	Coder identifier; date; time taken

Agreement should be reported separately for each axis, using Cohen's kappa for nominal axes and percentage agreement for descriptor completeness; values above 0.60 are conventionally considered substantial [14]. Reporting agreement per axis is essential: a system may be reliable for compartment and unreliable for phenotype, and an overall figure would hide that.

RULES FOR EVIDENCE SYNTHESIS

Pooling

Studies may be combined in one meta-analysis only if they share the compartment set, the primary phenotype and the intervention class, and if their descriptors fall within a range prespecified in the review protocol. Comparator categories are never combined: SH, AC, WL and BA studies are analysed in separate strata, and the pooled estimate presented as the main result is the SH stratum where it exists. This stratification makes the sham gradient visible, rather than averaging it away [2].

Mandatory sensitivity analyses

Four sensitivity analyses are recommended whenever the number of studies allows: (i) exclusion of [NR] studies; (ii) exclusion of [P?] studies; (iii) restriction to follow-up of at least 12 months; and (iv) restriction to studies at low risk of bias for blinding of outcome assessment. A result that holds only when [NR] or BA studies are included should be described as not robust.

Overlap between reviews

For overviews and umbrella reviews, overlap between included reviews should be quantified with the corrected covered area [15] and the analysis should be based on the deduplicated set of primary studies. Concordance among reviews that share most of their primary trials is not independent confirmation, and the framework's redundancy rule treats it accordingly [1].

Assigning maturity after synthesis

Maturity is assigned to the triplet from the synthesis, and then the four ceiling rules are applied in order: characterisation (predominantly [NR] evidence caps at M3), sham (benefit not confirmed in SH studies caps at M2), surrogate (reproductive triplets without live-birth benefit cap at M3) and redundancy. The level, the rule that capped it and the evidence that would be needed to move it should be reported together.

Indicators of codability

Applied to a body of literature, the manual yields three simple indicators that describe the state of reporting independently of effect sizes.

Table 5. Codability indicators.

Indicator	Definition	Interpretation
Full-code rate	Proportion of comparisons coded on all five axes without [NR] or [P?]	Share of the evidence that is reproducible and poolable
Axis-specific missingness	For each axis, proportion of comparisons with missing or inferred information	Identifies where reporting fails (typically I and P)
Anchor coverage	Proportion of comparisons that measured the anchor outcome of their primary phenotype as a primary or co-primary outcome	Distance between what is measured and what matters

These indicators can be tracked over time. An increase in the full-code rate in a given indication would show that reporting is improving even before effect estimates change.

Patient information by maturity level

The framework links each maturity level to a consent requirement [1,4]. Inconsistent wording in patient information is itself a source of overstatement, so **Table 6** proposes standard sentences that can be adapted to local language and regulation. They are not legal templates.

Table 6. Suggested standard wording for patient information.

Level	Suggested wording
M1	"This treatment is recommended by professional guidelines for your condition. Its benefits and side effects are well documented."
M2	"Some studies show benefit, but professional societies do not recommend this treatment for routine use. Benefits may last a limited time, and the treatment may need to be repeated. You should try recommended treatments first unless they are unsuitable for you."
M3	"This treatment is experimental for your condition. It should be offered only within a registered research study, and you should not be charged for it as an established therapy."
M4	"There is no evidence from controlled human studies that this treatment works for your condition. It is not offered outside approved early-phase research."

Worked coding exercises

The three study descriptions below are fictitious. They were constructed to exercise the decision rules and do not describe real trials.

Exercise 1. A randomised trial enrols 80 postmenopausal women with vaginal dryness and dyspareunia; eligibility requires vaginal pH above 5 and a vaginal health index below 15. Women receive three sessions of fractional CO2 laser one month apart (power, dwell time and spacing reported) or three sham sessions with the device emitting no energy, with participants and assessors blinded. The primary outcome is most bothersome symptom at 6 months; vaginal maturation index is secondary. Coding: VG-ATR / E1[parameters] / — / S+T-6m; comparator SH. The study is fully codable; because follow-up is under 12 months (D10), it can contribute to the SH stratum but cannot support an upgrade.

Exercise 2. A single-arm study of 40 women with "vaginal laxity and sexual dissatisfaction" injects PRP into the vestibule and anterior vaginal wall; platelet concentration is reported only for whole blood, and the main outcome is the Female Sexual Function Index at 3 months. Coding: VU+VG-STR[P?] / B1[NR: composition, activation] / — / F-3m; comparator BA. The study is not fully codable on three axes (P inferred, I not reported, follow-up short) and would be excluded from the main analysis under the sensitivity rules.

Exercise 3. A randomised trial in women with diminished ovarian reserve compares intraovarian PRP (PAW code and DEPA dose reported) with no injection before IVF; the primary outcome is antral follicle count at 3 months and clinical pregnancy is reported, but live birth is not. Coding: OV-REP-O / B1[descriptor] / — / R not measured (D9); comparator WL. Whatever its result, the trial cannot move the triplet above M3 because of the surrogate rule.

LIMITATIONS

The decision rules and thresholds proposed here — for example, the 70% majority for phenotype assignment and the 70% retention for follow-up — are pragmatic choices that have not yet been tested. The manual inherits the limitations of the framework it operationalises [1], including the judgement involved at the boundary between M2 and M3. Some descriptors require laboratory resources that are not universally available; the manual handles this by flagging

rather than excluding, at the cost of lower codability in under-resourced settings. Coding will initially take longer than conventional data extraction; the extraction form records coding time so that this burden can be measured. Finally, patients were not involved in drafting the standard wording in Table 6, which should be revised with patient representatives during the Delphi phase.

CONCLUSION

A classification is only as good as the agreement between the people who use it. This manual turns the five GRACE axes into a fixed coding sequence, explicit rules for the cases that cause disagreement, a reporting checklist that makes studies codable at the source, and synthesis rules that keep comparators and uncharacterised products apart. Its purpose is narrow: to make studies of regenerative and energy-based gynecology comparable before they are pooled. We invite investigators, reviewers and editors to test it, and we will revise it with the results of the reliability and recoding phases.

DECLARATIONS

Ethics approval and consent to participate. Not applicable. This article proposes methodological guidance and does not involve human participants, human material or human data.

CONSENT FOR PUBLICATION. Not applicable.

AVAILABILITY OF DATA AND MATERIALS. No new data were generated. The extraction form in Table 4 can be obtained from the corresponding author in spreadsheet format.

COMPETING INTERESTS. The authors have no relationship with manufacturers of vaginal energy-based devices, platelet-rich plasma systems, cell-based or exosome products, or menopausal hormone therapy products.

FUNDING. This research received no specific grant from any funding agency.

AUTHORS' CONTRIBUTIONS. MAP: conceptualisation, coding rules and drafting. MAP and MHK: critical revision of the rules, checklist and manuscript. All authors approved the final version.

ACKNOWLEDGEMENTS. None.

USE OF ARTIFICIAL INTELLIGENCE. An AI-based writing assistant was used to support drafting and language editing. The authors verified all statements and references against the primary sources and accept full responsibility for the content.

REFERENCES

1. Kume MH, Probst MA, Ribas CAPM. GRACE: a five-axis classification of evidence for regenerative and energy-based gynecology — compartment, phenotype, intervention, maturity and anchor outcome. Position paper. Manuscript submitted to J Stem Cell Res; 2026.
2. Probst MA, Kume MH, Peracchi E, Ribas CAPM. Regenerative and energy-based interventions in gynecology: an umbrella review of systematic reviews on platelet-rich plasma, vaginal laser, radiofrequency, and reproductive tissue regeneration. J Stem Cell Res. 2027;8(1). Available from: https://www.genesispub.org/jscr
3. Probst MA, Kume MH, Peracchi E, Ribas CAPM. Platelet-rich plasma for vulvovaginal disorders: an umbrella review of nineteen systematic, scoping and narrative reviews published since 2020. J Stem Cell Res. 2027;8(1). Available from: https://www.genesispub.org/jscr
4. Probst MA, Kume MH, Peracchi E, Ribas CAPM. Conventional hormonal therapy versus emerging regenerative interventions for genitourinary syndrome of menopause: a technical comparative matrix and evidence-tiered feasibility analysis of 2024–2026 systematic reviews and meta-analyses. J Clin Pract Med Case Rep. 2026;3(1):1-21. doi:10.52793/JCPMCR.2026.3(1)-38. Available from: https://www.genesispub.org/jcpmcr/conventional-hormonal-therapy-versus-emerging-regenerative-interventions-for-genitourinary-syndrome-of-menopause-a-technical-comparative-matrix-and-evidence-tiered-feasibility-analysis-of-2024-2026-systematic-reviews-and-meta-analyses
5. Kume MH, Probst MA, Peracchi E, Ribas CAPM. VAGINA-R: a phenotype classification for the genitourinary syndrome of menopause and the REGEN-VAG validation protocol. Manuscript in preparation; 2026.
6. Hopewell S, Chan AW, Collins GS, Hróbjartsson A, Moher D, Schulz KF, et al. CONSORT 2025 statement: updated guideline for reporting randomized trials. Nat Med. 2025;31(6):1776-83. doi:10.1038/s41591-025-03635-5. Available from: https://pubmed.ncbi.nlm.nih.gov/40229553/
7. Chan AW, Boutron I, Hopewell S, Moher D, Schulz KF, Collins GS, et al. SPIRIT 2025 statement: updated guideline for protocols of randomized trials. Nat Med. 2025;31(6):1784-92. doi:10.1038/s41591-025-03668-w. Available from: https://pubmed.ncbi.nlm.nih.gov/40295741/
8. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of

interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ. 2014;348:g1687. Available from: https://pubmed.ncbi.nlm.nih.gov/24609605/

9. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. Available from: https://pubmed.ncbi.nlm.nih.gov/33782057/
10. DeLong JM, Russell RP, Mazzocca AD. Platelet-rich plasma: the PAW classification system. Arthroscopy. 2012;28(7):998-1009. doi:10.1016/j.arthro.2012.04.148. Available from: https://pubmed.ncbi.nlm.nih.gov/22738751/
11. Mautner K, Malanga GA, Smith J, et al. A call for a standard classification system for future biologic research: the rationale for new PRP nomenclature. PM R. 2015;7(4 Suppl):S53-9. doi:10.1016/j.pmrj.2015.02.005. Available from: https://doi.org/10.1016/j.pmrj.2015.02.005
12. Magalon J, Chateau AL, Bertrand B, et al. DEPA classification: a proposal for standardising PRP use and a retrospective application of available devices. BMJ Open Sport Exerc Med. 2016;2(1):e000060. doi:10.1136/bmjsem-2015-000060. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC5117023/
13. National Cancer Institute. Common Terminology Criteria for Adverse Events (CTCAE). Bethesda, MD: NCI. Available from: https://ctep.cancer.gov/protocoldevelopment/electronic_applications/ctc.htm
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74. Available from: https://pubmed.ncbi.nlm.nih.gov/843571/
15. Pieper D, Antoine SL, Mathes T, Neugebauer EA, Eikermann M. Systematic review finds overlapping reviews were not mentioned in every other overview. J Clin Epidemiol. 2014;67(4):368-75. Available from: https://pubmed.ncbi.nlm.nih.gov/24581293/