STATISTICAL ANALYSIS PLAN
Tom EF Abbott, Gareth Ackland

Normative values of the COMPASS-31 questionnaire in adults: a systematic review and meta-analysis (PROSPERO: CRD420251175235)

Contents

1. Document status and purpose 2
2. Review objectives 2
3. Analysis populations and data sources 2
3.1 Individual participant-level reference population 2
3.2 Aggregate-data population 2
4. Outcomes 2
4.1 Primary outcome 2
4.2 Secondary outcomes 2
5. General statistical principles 2
6. IPD analyses 3
6.1 Overall total-score reference distribution 3
6.2 Bootstrap uncertainty 3
6.3 Domain distributions 3
6.4 Age analyses 3
6.5 Sex analyses 3
6.6 Geographical/dataset variation and influence 3
6.7 Health-status/comorbidity sensitivity analyses 3
6.8 Correlations among domains 3
7. Aggregate-data synthesis 4
7.1 Studies reporting means 4
7.2 Studies reporting medians 4
7.3 Domain aggregate analyses 4
8. Comparison of IPD and aggregate evidence 4
9. Risk of bias and certainty considerations 4
10. Analysis hierarchy and deviations 5
11. Software and reproducibility 5
12. Planned reporting 5
13. Version history 5
14. Registration 6

1. Document status and purpose

This statistical analysis plan (SAP**)** provides the detailed operational specification of the analyses described in the prospectively registered PROSPERO review (CRD420251175235). The registered protocol prespecified derivation of COMPASS-31 reference values using both published aggregate data and individual participant-level data obtained from eligible studies, including targeted acquisition of individual-level data from larger cohorts. The Individual Participant Data (IPD) and aggregate-data components therefore form part of the prospectively registered analytical strategy. This SAP provides the detailed statistical specification of that strategy, including estimation of empirical reference distributions, assessment of demographic and geographical variation, sensitivity and robustness analyses, aggregate-data synthesis, and comparison of aggregate and IPD-derived estimates.

2. Review objectives

The primary objective is to establish empirical reference values for the COMPASS-31 total score in adults meeting the registered reference-population eligibility criteria. Secondary objectives are to describe reference distributions for the six COMPASS-31 domains; assess variation by age, sex, and geographical dataset/region; evaluate robustness to alternative definitions of the reference population; synthesize eligible published aggregate data; and assess concordance between IPD-derived and aggregate-data estimates.

3. Analysis populations and data sources

3.1 Individual participant-level reference population

IPD supplied by investigators or available through eligible study datasets with ≥150 participants will be harmonized. The primary IPD reference population will comprise adults aged 18 years or older represented in eligible healthy/control/reference datasets and with sufficient COMPASS-31 data to derive the relevant total or domain score. The primary pooled analysis will retain participants according to the harmonized reference-population definition used for the study.

3.2 Aggregate-data population

Eligible published cohorts not including the IPD cohort will contribute aggregate COMPASS-31 data for adults described as healthy, control, or without the diseases excluded by the registered eligibility criteria. Extracted outcomes may include total and domain means with dispersion measures and/or medians with quartiles/ranges, as reported.

4. Outcomes

4.1 Primary outcome

The primary outcome is the distribution of COMPASS-31 total score in the pooled IPD reference population. The principal upper reference limit will be the 95th percentile (P95). The median (P50), P90, and P97.5 will also be reported to characterize the upper tail without treating any percentile as a diagnostic threshold.

4.2 Secondary outcomes

Secondary outcomes in the pooled IPD reference population are distributions of the six weighted COMPASS-31 domains (orthostatic intolerance, vasomotor, secretomotor, gastrointestinal, bladder, and pupillomotor), age- and sex-stratified total-score reference values and geographical/dataset variation. For aggregate-data, estimates of total and domain scores will be reported.

5. General statistical principles

Analyses will emphasize estimation and uncertainty rather than null-hypothesis testing. Continuous COMPASS-31 distributions will be summarized using n, mean, standard deviation, median, interquartile range, and selected empirical percentiles as appropriate. Because COMPASS-31 total and domain distributions may be skewed,

bounded, discrete, and/or zero-inflated, reference limits will be derived empirically rather than by assuming normality. Missing observations will not be imputed for reference-percentile estimation; analyses will use participants with the score required for the analysis. Statistical code and derived outputs will be retained for reproducibility.

6. IPD analyses

6.1 Overall total-score reference distribution

Empirical P25, P50, P75, P90, P95, and P97.5 will be estimated from the pooled reference data using the standard sample quantile implementation used in the analysis code (R quantile type 7). The empirical cumulative distribution function will be used for percentile-position interpretation.

6.2 Bootstrap uncertainty

Uncertainty around reference percentiles, with particular emphasis on P95, will be estimated using 5,000 nonparametric bootstrap replicates sampled with replacement at the participant level within the relevant analysis population. Two-sided 95% bootstrap confidence intervals will be obtained from the 2.5th and 97.5th percentiles of the bootstrap distribution (percentile method).

6.3 Domain distributions

For each COMPASS-31 domain, n, mean, standard deviation, median, selected percentiles, observed range, and prevalence of zero scores will be described. Because domain scores are discrete and may have substantial floor effects and ties, empirical cumulative proportions will be emphasized and conventional continuous-percentile interpretation will be avoided where inappropriate.

6.4 Age analyses

Total-score reference values will be summarized across prespecified adult age bands used in the IPD analysis (18 to ≥80 years). P95 estimates will be accompanied by bootstrap 95% confidence intervals. Age will also be examined as a continuous variable in descriptive/regression analyses where appropriate to assess whether separate age-specific reference limits are supported.

6.5 Sex analyses

Total-score reference values will be summarized separately for participants recorded as male and female, with bootstrap uncertainty for P95. Participants recorded outside these categories will remain in analyses not stratified by sex but will not be assigned to a binary sex category for sex-stratified analyses.

6.6 Geographical/dataset variation and influence

Reference percentiles will be described by contributing dataset/site where sample size permits. Robustness of the pooled reference distribution will be evaluated using leave-one-dataset-out analyses, recalculating P50, P90, P95, and P97.5 after exclusion of each contributing dataset in turn.

6.7 Health-status/comorbidity sensitivity analyses

The overall COMPASS-31 distribution and P95 will be recalculated among participants for whom comorbidity data are available. The influence of individual recorded comorbidities will additionally be explored descriptively by reporting the number of participants with each condition and the total-score median and P95 among participants with and without the condition. Participants with no reported comorbidity will also be compared with those with recorded cardiometabolic and/or neurological disease. These analyses are exploratory and will be used to assess the potential influence of health status on the reference distribution; they will not be used to derive disease-specific reference limits.

6.8 Correlations among domains

Associations among the six domain scores will be described using Spearman rank correlations. These analyses are descriptive and are not used to construct the primary reference limit.

7. Aggregate-data synthesis

The aggregate-data synthesis will be undertaken after completion of eligibility confirmation and risk-of-bias/quality assessment to independently evaluate the findings derived from the IPD analysis. Concordance between aggregate-data and IPD-derived estimates will be assessed for the COMPASS-31 total score and individual domains where sufficient data are available.

7.1 Studies reporting means

For cohorts reporting a mean COMPASS-31 score with an estimable standard deviation and sample size, study-specific means will be synthesized using random-effects meta-analysis. Between-study heterogeneity will be quantified using tau-squared and I-squared with corresponding heterogeneity statistics. The primary random-effects estimator and confidence-interval method will be those implemented in the locked analysis code; alternative reasonable estimators may be presented as sensitivity analyses if materially informative. Studies with unusable or zero dispersion values will be audited and excluded from inverse-variance analyses when mathematically inappropriate, with sensitivity analyses reported where relevant.

7.2 Studies reporting medians

Studies reporting medians will first be summarized descriptively. Where sample size and quartile information permit, medians will be meta-analyzed using the quantile-estimation approach implemented in the R package metamedian (or its underlying quantile-estimation method), which pools study medians without treating medians as means. Cohorts lacking the information required by that method will remain in the descriptive synthesis. Sensitivity analyses will assess influential or methodologically anomalous cohorts where justified.

7.3 Domain aggregate analyses

Available aggregate data for each of the six domains will be synthesized separately for mean-reporting and median-reporting cohorts using methods consistent with the total-score analyses. Availability and data quality will be audited by domain. Zero or otherwise unusable reported standard deviations will not be treated as valid inverse-variance information; sensitivity analyses excluding such observations will be emphasized.

8. Comparison of IPD and aggregate evidence

The aggregate synthesis serves as an independent assessment of the external concordance of the IPD-derived reference profile. IPD and aggregate estimates will be compared descriptively for total scores and, where data permit, across all six domains. Concordance will be assessed from the direction, magnitude, and pattern of estimates and their uncertainty rather than by a formal equivalence test. The aggregate synthesis will not be used to redefine the empirical IPD P95 unless a separate analysis is explicitly identified as exploratory.

9. Risk of bias and certainty considerations

Risk-of-bias/quality assessment will be completed for eligible aggregate studies before final interpretation of the synthesis. Study quality, adequacy of sample size, participant characteristics, and suitability/consistency of COMPASS-31 outcome measurement will be considered when interpreting pooled estimates and heterogeneity. Any sensitivity analysis based on risk-of-bias judgments will be specified before the final aggregate synthesis where feasible and reported transparently.

10. Analysis hierarchy.

Analysis	Status in SAP v1.0	Interpretation
Overall IPD total-score distribution and P95	IPD analysis already commenced/performed	Primary reference analysis; documented transparently, not claimed as newly prospective
IPD domain distributions; age, sex, region/site	IPD analysis already commenced/performed	Registered secondary/moderator framework; detailed methods documented here
Presence of comorbidity	Additional IPD sensitivity analysis	Robustness to reference-population definition
Leave-one-dataset-out and bootstrap analyses	Additional IPD robustness analyses	Precision/influence assessment
Aggregate mean synthesis	Specified before completion of formal aggregate synthesis	Primary aggregate synthesis where means/SDs are available
Aggregate median synthesis	Specified before completion of formal aggregate synthesis	Complementary synthesis preserving median reporting
IPD–aggregate comparison	Specified before completion of formal aggregate synthesis	External concordance/validation analysis
Individual-comorbidity comparisons	Supplementary/exploratory	Descriptive; not condition-specific reference limits
Online reference calculator	Implementation analysis	Translation of derived empirical distributions; not diagnostic

Any material departure from this SAP during completion of the aggregate synthesis will be dated, justified, and reported as an amendment or deviation. Analyses added after inspection of the relevant results will be identified as post hoc/exploratory.

11. Software and reproducibility

Analyses will be conducted in R 4.5.2. The final manuscript repository will retain the master analysis script, relevant package/version information, derived tables and figures, and code used to generate the public reference calculator. Participant-level datasets will not be made public where data-sharing agreements prohibit redistribution; de-identified derived lookup tables may be released for calculator deployment.

12. Planned reporting

Results will be reported in accordance with applicable systematic-review/meta-analysis and observational/IPD reporting guidance. The primary manuscript will distinguish population reference limits from diagnostic thresholds. The online calculator will report empirical reference position and upper reference limits without diagnosing autonomic dysfunction or classifying an individual as clinically normal or abnormal.

13. Version history

Version	Date	Status	Summary
1.0	1st August 2026	Initial SAP	Operational specification of registered IPD/aggregate framework; documents analyses already undertaken on IPD and prespecifies completion of aggregate synthesis and IPD–aggregate comparison.

14. Registration

PROSPERO registration: CRD420251175235 — “Normative values of the COMPASS-31 questionnaire in adults: a systematic review and meta-analysis.”