



Assessment rubric

Version 1.0 | 13 October 2026

Reviewers assess what the product is intended to do, how well its principal claims are supported and how it has been designed for its users. Each entry is judged against the published requirements for its category, rather than against a fixed quota of winners.

The six criteria

Intended use and need – 15 points

Reviewers assess the product's intended users, the need it addresses and the conditions in which it is used. Explain its purpose and any limits on appropriate use.

Performance and supporting evidence – 30 points

Provide evidence that addresses the principal claims of the named product and version. Reviewers consider the method, relevance, consistency and limitations of the material. They distinguish what was measured from what is inferred, and direct evidence from evidence concerning another product, ingredient or context.

Usability and accessibility – 20 points

Explain how people learn to use the product, complete important tasks and recognise errors or limitations. Include relevant user research, instructions, fitting or setup requirements and the needs of different users. Show how the findings relate to important user tasks.

Quality, risk and product information – 20 points

Show how product responsibilities, foreseeable risks, instructions, warnings, maintenance and relevant quality information fit together. Reviewers look for material gaps and inconsistencies between the proposed use, evidence and information supplied to users.

Product improvement or innovation – 10 points

Describe a meaningful improvement against a stated starting point or alternative. It may concern function, design, access, reliability or use. A mature product can demonstrate a useful improvement without claiming to be the first of its kind.

Responsible design and lifecycle – 5 points

Explain relevant arrangements for maintenance, useful life, support, disposal, data handling or supply responsibility. Select issues that matter for the product. Environmental or social claims need their own support; a general company statement does not establish a product attribute.

How ratings become a score

Each criterion receives a rating from 0 to 5, with half-points permitted. The rating is divided by five and multiplied by its weight. The six weighted results add to a score out of 100.

Rating	Meaning
0	Information is absent or the criterion cannot be assessed.
1	The material is substantially inadequate or contains a major unresolved inconsistency.



Rating	Meaning
2	Some support is present, with important gaps.
3	The criterion is met with material that can be checked.
4	The criterion is well supported and the product shows a clear strength in this area.
5	Exceptionally strong, relevant and consistent support establishes an outstanding result in this area.

What is required for an award

An award requires at least 80 points overall, a rating of at least 4 for performance and supporting evidence, at least 3 for quality, risk and product information, and at least 2 in every other criterion. An unresolved material authenticity, eligibility or product-risk concern prevents an award even where the numerical total is high enough.

There is one award designation, with no gold, silver or paid upgrade. Scores are not a clinical probability, consumer safety rating or ranking across unlike products.

Differences between categories

Evidence must fit the product. A wearable needs relevant measurement conditions and a reference method; a diagnostic needs clearly defined samples, intended use and performance measures; a supplement must explain whether the evidence applies to its finished formulation. The category examples below show how this applies in practice.

Decisions and feedback

Two reviewers assess an entry independently. A chair checks the reasoning and resolves differences, using a third assessment where required. Applicants receive criterion-level feedback and the reasons for the decision. Purchases, membership and sponsorship do not enter the assessment.

[Read the review process](#) · [Prepare your evidence](#)

Applying the anchors consistently

A rating must cite the entry material and describe its relevance. Record missing information separately from unfavourable evidence. Neither a familiar brand nor an impressive report title supplies a missing finding. Where the evidence is indirect, explain the link and assess its limits.

A rating of 4 requires more than a complete form: the relevant evidence must be sufficiently strong to support the principal claim within its stated scope. A rating of 5 should identify what makes the result exceptional, without relying on company reputation. A half-point reflects a reasoned position between anchors, not false measurement precision.



Category examples for the evidence criterion

Category	Material that may support a strong rating	A gap that must be resolved
Nutrition and supplements	Identified formulation and use level, relevant quality material and a reasoned link between the claims and applicable evidence	Ingredient evidence is presented as a finished-product result without explaining formulation, dose or population differences
Wearables	The entered version assessed against an appropriate reference, with error, conditions and relevant user characteristics reported	A correlation or user satisfaction score is used as the only support for an accuracy claim
Home devices	Model-specific performance material and evidence that important tasks can be completed in the intended home setting	Professional operation in a controlled setting is assumed to prove correct home use
Diagnostics	Defined sample, target, reference and population, with relevant analytical or clinical performance and uncertainty	Performance from one sample type or setting is applied to another without support
Digital health and software	Versioned validation, described data and workflow, relevant results and limitations, with a reviewable demonstration	A changing production model is claimed to match evidence from an unidentified earlier release
Rehabilitation and assistive products	Evidence addressing the user's task, fitting conditions and relevant benefit, with limitations stated	Easier operation is described as improved clinical recovery without evidence for that outcome
Oral and personal health	Evidence for the named formulation or material under relevant contact and use conditions	A general ingredient property is presented as proof of all claimed finished-product benefits
Ingredients and components	Defined specification, relevant test conditions, traceability and downstream limits	A component report is used to imply that all products incorporating it have been assessed

These examples illustrate evidence handling. They do not prescribe one universal study design, guarantee a rating or replace the requirements relevant to the claim and product.

Notes for the other criteria

For intended use, check that the task, user and setting are sufficiently precise to assess. For usability, distinguish observation of representative tasks from preference surveys alone. For quality and risk, look for consistency across labels, instructions, disclosures and reports. For improvement, require a stated baseline. For lifecycle, evaluate the issues relevant to the product rather than adding points for unrelated corporate commitments.

If a criterion appears inapplicable, record the reason and seek guidance before scoring. Do not redistribute its weight or award automatic full marks. The chair must decide how the published criterion applies to that case; a material change to the framework requires a future rule version.



Worked threshold example

An illustrative entry rated 5, 3, 5, 4, 5 and 5 in the six criteria scores 84: $15 + 18 + 20 + 16 + 10 + 5$. It does not qualify because its evidence rating is 3, below the required 4. This is a fictional arithmetic example, not an award decision.

Moderation record

Record both original assessments, points of agreement, differences, requests and responses, any third assessment, and the final reasoned outcome. A third assessment is required if the initial totals differ by more than 15, either is between 75 and 84 inclusive, the reviewers disagree on a mandatory threshold, or a material specialist question remains.

Use a final criterion-by-criterion assessment to explain the decision. A simple average must not conceal an unresolved failure of a required threshold. Record the responsible chair, date and applicable rule version. Retain the original assessments.

Calibration and change control

Before reviewing live entries, reviewers work through common example entry materials and compare their interpretation of the anchors. Differences are discussed and documented. Calibration clarifies the published standard; it does not create an undisclosed award quota.

Version 1.0 uses the weights and thresholds above. A new standard takes effect prospectively and is recorded against each accepted entry.