HOW THE RUBIC 2026 JURY SCORES THE PROJECTS

Competence-weighted evaluation, uniform handling of conflicts, standardised scores — adopted unanimously by the Jury

The Official Rules set out ten evaluation criteria, grouped in five sections of equal weight, each scored from 0 to 5. What they do not say is who is entitled to judge what. This page explains a decision the Jury has taken on that point, and publishes it in full.

Why we changed something

The RUBIC Jury is deliberately heterogeneous: three robotic urologists, one biomedical robotics engineer, one deep-tech venture builder and one MedTech industry expert. That mix is the point — the projects sit at the intersection of surgery, engineering and entrepreneurship, and no single discipline can assess them alone.

But a simple average across six jurors carries a hidden assumption: that all six are equally able to judge every criterion. Under a plain average, a venture builder's opinion on the clinical impact of a project would count exactly as much as that of a urologist who performs the operation, and a clinician's opinion on market scalability exactly as much as that of the industry expert. Equal weighting is not neutrality. It is a specific assumption, and in a Jury built on complementary expertise it is the wrong one.

What we did

Each member of the Jury has been assigned a coefficient for each of the ten criteria:

- **1** — primary competence: the score counts in full.
- **0.5** — informed but secondary competence: the score counts half.
- **0** — outside the member's declared profile: the criterion is not scored at all.

Within each criterion the coefficients are renormalised so that they sum to one. A member with a coefficient of 1 therefore counts exactly twice a member with 0.5, and a criterion left unscored does not lower anyone's average — it is simply not counted.

Only three levels were used. On small and noisy datasets, coarse weights are as accurate as finely calibrated ones, and a coefficient of 0.85 rather than 0.90 would be impossible to justify to a team that asked why. Three levels can be stated in one sentence and defended in a meeting.

The coefficients, in full

Each member reviewed and confirmed their own row before any project was read, and was free to decline a coefficient of 1 or to request a change. The matrix was then frozen and minuted. It has not been modified since, and it will not be modified while scoring is under way.

Jury member	Declared competence	1.1	1.2	2.1	2.2	3.1	3.2	4.1	4.2	5.1	5.2
Pieter De Backer — ORSI Academy (BE)	Robotic urology, surgical AI	**1**	**1**	**1**	**1**	**1**	0.5	**1**	0.5	**1**	**1**
Karl-Friedrich Kowalewski — DKFZ (DE)	Robotic urology, education, AI	**1**	**1**	0.5	0.5	0.5	0.5	**1**	0.5	**1**	**1**
Cristian Fiori — University of Turin (IT)	Clinical robotic urology	**1**	**1**	0.5	0.5	0	0	**1**	0	**1**	**1**
Leonardo De Mattos — IIT (IT)	Biomedical robotics engineering	**1**	0.5	**1**	**1**	**1**	**1**	0.5	0.5	**1**	**1**
Anilkumar Dave — Darwix (IT)	Deep-tech innovation, venture building	0.5	0	**1**	**1**	0.5	**1**	0	**1**	**1**	**1**
Miguel Fenech — MedTech World (MT)	MedTech industry and market	0.5	0.5	**1**	**1**	0.5	**1**	0	**1**	**1**	**1**
Members scoring each criterion	**6**	**5**	**6**	**6**	**5**	**5**	**4**	**5**	**6**	**6**	

Highlighted: criterion 4.1, clinical and educational impact; criterion 4.2, scalability and diffusion. The two are deliberately mirrored — clinicians carry the first, market and innovation experts carry the second.

1.1	Alignment with the RUBIC theme
1.2	Relevance for robotic urological surgery or its ecosystem
2.1	Degree of innovation
2.2	Originality compared to existing solutions
3.1	Technical feasibility
3.2	Development roadmap clarity
4.1	Potential clinical / educational impact
4.2	Scalability and broader applicability
5.1	Team composition and complementarity
5.2	Use of the RUBIC opportunity

What this changes, stated plainly

Weighting improves the fairness and the defensibility of the process more than it changes the outcome. In simulation on this exact matrix, the weighted result tracks the underlying ranking slightly better than a plain average — the difference is real but modest. We publish this because overstating the effect would be as misleading as hiding the method. No project wins or loses because of a coefficient; the coefficients ensure that each judgement is made by those competent to make it.

Conflicts of interest: one rule for everyone

A member of the Jury who declares a conflict of interest with a project does not score that project. The rule is uniform: it applies to every member and every project, without exception and without individual discretion.

We considered allowing a conflicted member to score anyway, with the score flagged. We tested it and discarded it. Because jurors differ in how strictly they mark, and because a project assessed by five members is not assessed by the same panel as one assessed by six, letting each member choose whether to score would mean that a project gains or loses points according to that individual choice rather than on its merits. In simulation on our own matrix the difference was worth up to several points out of one hundred — enough to move a project in or out of the top five. A uniform rule removes the problem entirely.

The expertise is not wasted. A member who abstains records a written assessment of the project in the scoring workbook. It is read by the Jury and minuted. It carries no numerical weight, but it is on the record and it informs the discussion.

Scores are standardised before they are combined

Each juror's scores are centred on that juror's own average across the eleven projects before the coefficients are applied. Jurors mark differently — some systematically higher, some lower — and without this correction a project would be advantaged or penalised by which members happened to assess it. Centring removes that effect: it does not alter any juror's ranking of the projects, only the level at which they mark. It is standard practice in the peer review of research funding.

Safeguards

- **Every criterion is scored by at least four members.** No criterion depends on a single opinion, and a conflict of interest can never leave one below quorum.
- **Coefficients frozen before scoring.** They were adopted and minuted before any submission was circulated. Adjusting weights after seeing scores would invalidate the exercise.
- **Conflicts of interest declared in writing, project by project.** Every member has signed a declaration on confidentiality, non-use of the submitted ideas and conflict of interest, in which each of the eleven projects is marked individually. Abstentions are recorded and minuted.
- **The President does not vote.** The Chair of the Jury has no voting right and takes part only to break a tie, which is recorded in the minutes.
- **The result is checked without weights.** Before the outcome is announced, the ranking is recomputed with all scores counted equally. Any project whose position changes is discussed by the Jury and the discussion is minuted.
- **Intellectual property remains with the teams.** As stated in Art. 6 of the Official Rules, and reinforced by the non-use obligation signed by every member of the Jury.
- **Submissions from the organising institution are declared.** RUBIC 2026 is organised by the University of Padua, and one of the eleven submissions comes from a team of the University of Padua. It is stated here for transparency. No member of the Jury is affiliated with that team, the President of the Jury does not vote, and the project is assessed under exactly the same rules as every other.

On what basis

The approach is not improvised. Competence-based allocation of evaluators is standard practice in European research and innovation funding, where reviewers are matched to each proposal by expertise profile rather than assigned uniformly. In the literature on combining expert judgements, weighting experts unequally has repeatedly outperformed equal weighting; and studies of grant peer review show that reviewer expertise systematically shifts the scores given — ignoring it does not remove the effect, it only leaves it undeclared. In health technology assessment, explicit participant weights accompanied by a sensitivity analysis are the established protocol. Key references: Cooke (1991) and Colson & Cooke (2018) on the classical model of structured expert judgement; Gallo, Sullivan & Glisson, PLoS ONE 2016, on reviewer expertise in funding decisions; Hummel, Bridges & IJzerman (2014) on group decision-making in benefit–risk assessment; Dawes (1979) and Einhorn & Hogarth (1975) on the robustness of coarse weighting schemes.

Adopted by the Jury and minuted before any submission was circulated. Questions on the evaluation method: info@rubic2026.org