



October 13, 2026

Sent electronically via <https://rfi.grants.nih.gov/?s=6a2b1c22c3cfda46e0badd2>

RE: Response to NOT-OD-26-088, *Request for Information (RFI) on Proposed Changes to Reporting Outcomes from NIH Peer Review*

Submitting Organization: University of Colorado Boulder

To Whom It May Concern:

The University of Colorado Boulder appreciates the opportunity to respond to NIH's Request for Information regarding the proposed approach to reporting peer review outcomes. This response reflects input from faculty and research leaders across the biological, behavioral, engineering, and health sciences, including current and former NIH principal investigators, applicants, awardees, study section reviewers, advisory council members, and training grant leaders. The responses reflected strong concern that withholding final overall impact scores and applicable percentiles would diminish the usefulness and transparency of peer review outcomes.

CU Boulder supports NIH's goals of improving the usefulness and transparency of peer review outcomes and recognizes NIH's interest in evaluating applications holistically while avoiding excessive reliance on numerical scores or rigid paylines. Scientific merit, mission relevance, portfolio needs, investigator circumstances, and available resources can all be legitimate considerations. The central concern raised by our research community is that NIH can exercise that discretion without withholding scientific-merit information already produced through expert review. We recommend retaining quantitative reporting alongside written critiques and discussion summaries and making subsequent funding considerations more transparent.

1. Broad categories would not provide sufficient information to applicants.

Under the proposal, "Most Competitive" would cover the top 25% of final overall impact scores and "Competitive" the next 25%. These groupings are not equivalent to funding paylines, and a designation in the top category would not necessarily indicate a strong likelihood of award. Applications at opposite ends of either category may differ substantially in their scientific assessments and realistic resubmission prospects, yet applicants would receive the same label.

Faculty repeatedly emphasized that final overall impact scores and percentiles, where provided, are essential for deciding whether to address a limited set of concerns, conduct substantial additional experiments, redesign a project, or pursue a different funding opportunity. The distinction also matters for institutional advising and for timely decisions about hiring, trainee support, and continuation of research. Category-only reporting would leave investigators with less actionable information at a time when a single application cycle may have significant consequences.

2. Written feedback and quantitative results serve complementary purposes.

CU Boulder values reviewers' written critiques, individual criterion or factor scores, and panel discussion summaries, which NIH proposes to retain. These materials explain an application's strengths and weaknesses, but comments can vary in specificity, conflict with one another, or fail to convey how heavily particular concerns influenced the panel. The final overall impact



score helps applicants interpret the collective judgment that emerged from discussion; it does not replace the scientific reasoning in the written record.

Importantly, the RFI does not propose to provide substantially more detailed critiques to compensate for the loss of reported overall scores. If NIH expects applicants and program staff to rely more heavily on narrative assessments, it should improve the consistency and usefulness of those assessments through guidance, training, and quality assurance without imposing unmanageable additional burdens on reviewers.

3. Preserve the influence of expert peer review while allowing programmatic judgment.

Reviewers would continue to assign numerical overall impact scores, which NIH would use to place applications into categories. The proposed change therefore does not eliminate quantitative peer review; it withholds its final assessment from applicants, institutions, program staff, leadership, and advisory councils. Respondents were concerned that this could make it more difficult to see how expert judgments relate to award decisions and could weaken confidence that scientific merit remains central. Reduced transparency could also heighten concerns about arbitrary or politically influenced funding decisions, regardless of whether such influence occurs, potentially weakening confidence in the integrity of the peer review system.

Programmatic flexibility and quantitative transparency are not mutually exclusive. For example, the National Institute of General Medical Sciences has long considered scientific merit alongside mission, portfolio, early-stage or at-risk investigator status, and other research support without relying on a fixed payline. That approach illustrates how NIH can avoid treating small score differences as dispositive while still using and communicating peer review information. More explicit descriptions of programmatic criteria, accompanied by aggregate data on funding outcomes, would better support the NIH Unified Funding Strategy than withholding scores.

4. Consider the operational and scientific consequences.

Respondents anticipated that broad categories could require program officers to examine and comparatively interpret more full summary statements, prompt additional inquiries from applicants, and make funding recommendations harder to explain. Delays or uncertainty can disrupt research continuity, including commitments to graduate students, postdoctoral scholars, and skilled research staff. Faculty who serve as reviewers also raised concerns about whether withholding the collective scores they help generate would diminish the perceived value of their service and confidence in the process.

These are potential consequences, not established effects of the proposed system. They are nevertheless sufficiently consequential to warrant careful assessment before any agency-wide implementation, particularly across different funding mechanisms, scientific fields, and investigator career stages.

5. CU Boulder recommends that NIH:

- Retain reporting of final overall impact scores for discussed applications and percentiles where applicable to principal investigators, applicant institutions, NIH program staff, Institute/Center/Office leadership, and advisory councils. Categories, if useful, should supplement rather than replace these results.



- Clarify how peer review assessments are weighed alongside mission relevance, portfolio considerations, investigator funding circumstances, and resource constraints, while preserving appropriate programmatic discretion.
- Improve the specificity and consistency of critiques and discussion summaries so that applicants can identify the concerns that most influenced the panel's assessment.
- Publish appropriately aggregated, mechanism- and Institute/Center/Office-level data connecting review results with funding decisions, including the proportion of applications funded within any reported category, while protecting applicant confidentiality.
- Consult applicants, reviewers, advisory councils, and program staff and conduct a limited, independently evaluated pilot before any categorical-only reporting is adopted. Assess applicant understanding, resubmission behavior, reviewer and staff workload, award timing, and consistency across review settings; publish the findings.
- If NIH nevertheless adopts categories, define category assignment and "Not Discussed" precisely, consider narrower groupings, and provide a reliable way for applicants to obtain clarification about their review outcome and general funding considerations.

NIH peer review is a foundational strength of the nation's biomedical research enterprise. The campus feedback was clear that the existing system is not perfect, but that scores, written critiques, panel discussion, and programmatic judgment provide distinct forms of information that are most useful together. NIH should improve the interpretation and use of those results rather than limit access to information its reviewers already produce.

CU Boulder therefore respectfully urges NIH not to replace reporting of final overall impact scores and applicable percentiles with the proposed three-category system. We encourage NIH to retain quantitative outcomes, improve narrative feedback, communicate programmatic considerations clearly, and evaluate any proposed reform before broad implementation.

We appreciate the opportunity to provide feedback. If you have any additional questions regarding our comments, please contact me at massimo.ruzzene@colorado.edu.

Sincerely,

Massimo Ruzzene

Senior Vice Chancellor for Research & Innovation
Dean of the Institutes
Slade Professor of Mechanical Engineering
University of Colorado Boulder
Boulder, Colorado 80309