



October 13, 2026

Re: NIH RFI on Proposed Changes to Reporting Outcomes from Peer Review

To the National Institutes of Health:

The National Academy of Medicine and the National Academy of Sciences appreciate the opportunity to comment on NIH's proposed changes to the reporting of peer-review outcomes. As part of the National Academies of Sciences, Engineering, and Medicine, NAM and NAS draw on the expertise of their members and other leading scientists, engineers, and health professionals to provide independent, evidence-based advice to government and the public. Our comments are grounded in the National Academies extensive work on federal research allocation, scientific peer review, research integrity, and NIH priority setting, including *Allocating Federal Funds for Science and Technology* (1995), *Scientific Opportunities and Public Needs* (1998), *Spurring Innovation in Food and Agriculture* (2014), *Evaluation of the Congressionally Directed Medical Research Programs Review Process* (2016), and *Fostering Integrity in Research* (2017).

The RFI proposes replacing the reporting of final overall impact scores and percentiles from first-level peer review with three broad categories. Although scores would continue to be calculated, they would no longer be shared with applicants, institutions, NIH program staff, ICO leadership, or NIH Advisory Councils. We agree that an overall impact score is an estimate and should not substitute for expert judgment or function mechanically as a payline. NIH should weigh peer-review comments and panel discussion together with scientific opportunity, public health needs, portfolio balance, workforce considerations, and available funds.

However, we do not support eliminating access to final scores and percentiles. These measures provide valuable information about the collective judgment of peer reviewers that cannot be fully captured by broad categories alone. The final score is the only standardized, panel-wide synthesis of all voting members' judgments after discussion. Percentiles also support comparisons across review groups with different scoring behavior. Replacing these measures with broad categories would remove valuable information without changing the underlying review process and could create new de facto thresholds.

The RFI notes the proposed approach would align NIH's reporting practices more closely with those of the National Science Foundation, Department of Energy, and Department of Agriculture. However, *Allocating Federal Funds for Science and Technology* identified NSF and NIH as leading examples of competitive federal research review, while *Spurring Innovation in Food and Agriculture* found Agriculture and Food Research Initiative review broadly comparable to NIH and NSF. Other major research agencies pair programmatic discretion with safeguards such as access to peer-review information, decision rationales, reconsideration, or external review.

Withholding scores would weaken transparency and accountability, thus eroding trust in the decision process. Applicants would be less able to interpret outcomes and make informed resubmission decisions. Program staff and councils would have less complete peer-review information when comparing applications and constructing portfolios. Broad categories, combined with greater reliance on programmatic considerations under the Unified Funding Strategy, could make decisions appear more discretionary and less clearly grounded in scientific merit, particularly without granular

feedback. Detailed scores also have educational value, helping investigators understand how their work was assessed and identify areas for improvement.

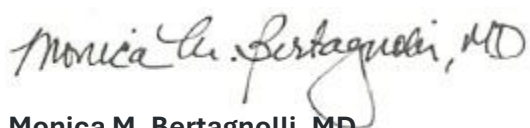
The Academies' 2016 evaluation of the Congressionally Directed Medical Research Programs identified concerns about the transition from peer review to programmatic decision-making and recommended greater transparency, including providing applicants with programmatic rationale alongside peer-review results. Programmatic judgment is legitimate, but it should not obscure independent expert assessment. Elected leaders and agency officials properly set missions, strategies, portfolios, and budgets; qualified peers should remain the principal judges of scientific and technical merit.

We recommend that NIH:

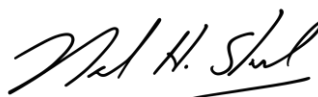
1. Retain final overall impact scores and percentiles for applicants, program staff, ICO leadership, and advisory councils. NIH could add the proposed categories as supplemental and explain that small numerical differences do not represent precise differences in merit.
2. Publish prospectively each ICO's decision factors, provide written rationales when funding decisions depart materially from peer-review assessments, and publish aggregate data sufficient to assess how peer-review outcomes relate to funding decisions.
3. Establish periodic external review of how ICOs apply programmatic judgment, while preserving strong conflict-of-interest rules, audit trails, and clear procedural appeal mechanisms.
4. Pilot and evaluate any change to score reporting before adopting it NIH-wide, with prespecified measures of transparency, consistency, fairness, applicant experience, and the relationship between peer-review outcomes and funding decisions.

NIH can reduce false precision without suppressing useful information. A transparent, contextualized, and evaluated approach would better advance the Unified Funding Strategy's commitment to scientific merit and sustain trust in biomedical research.

Sincerely,



Monica M. Bertagnolli, MD
President, National Academy of Medicine



Neil H. Shubin, PhD
President, National Academy of Sciences