

A large decorative graphic on the left side of the slide, consisting of a dark grey arrow pointing right, with a white outline and a green arrow pointing right inside it.

# CSR Advisory Council Workgroup: Simplifying Review Criteria for Clinical Trials

**Bruce Reed, PhD**  
Deputy Director  
Center for Scientific Review

**Tonya Palermo, PhD**  
Professor of Anesthesiology,  
Pediatrics, and Psychiatry  
University of Washington

March 29, 2021

# Background

- External feedback and CSR/OER concerns that the review criteria have become too numerous and complex
  - reduced focus on scientific merit
  - increased reviewer burden
- September 2019 CSR Director raises issue with CSRAC

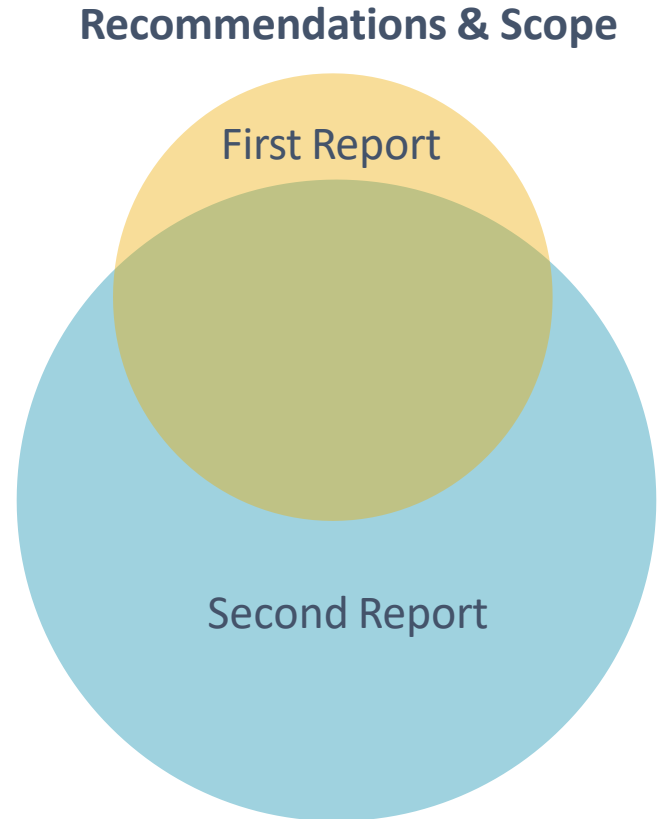
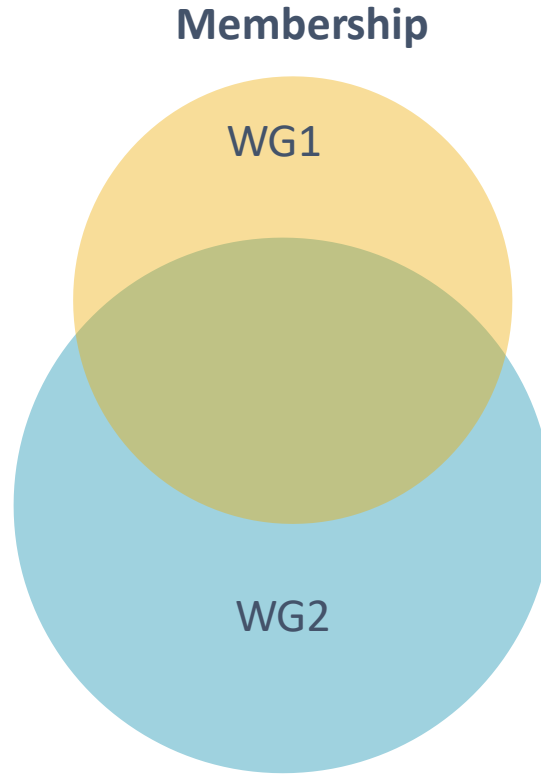
# Process

## WG1

- Simplifying Review Criteria Working Group
- non-clinical trials R01s
- January '20-March'20

## WG2

- Clinical Trials Criteria Working Group
- Clinical trials R01s
- June '20-March'21



# WG2: Clinical Trials Criteria Workgroup Members

## CSR Advisory Council



Alfred George, M.D.



Yasmin Hurd, Ph.D.



Tonya Palermo, Ph.D.



Michelle Janelins, Ph.D.

## Ad Hoc



Pamela Munster, Ph.D.  
University of California San  
Francisco



Brian Boyd, Ph.D.  
University of Kansas



Brooks King-Casas, Ph.D.  
Virginia Tech



Matthew Carpenter, Ph.D.  
Medical University of  
South Carolina

## NIH Staff



Sally Amero, Ph.D.  
Office of Extramural Research



Bruce Reed, Ph.D.  
Center for Scientific Review

## Charge to CTCWG “WG2”

- Recommend how clinical trial review criteria should be modified to reduce reviewer burden and improve review outcomes.
- Consider the full range of NIH clinical trials; BESH, mechanistic clinical trials, and interventional trials
- Start with the recommendations of WG1



## Additional Questions for WG2



- Concerns about the quality of science and need for improved stewardship of public funds were the basis for the current NIH policies on clinical trials.
  - *How might review of rigor and reproducibility issues be improved?*
- Marked underrepresentation of Blacks, Hispanics and other minorities in science, and persistent racial disparities in NIH funding are troubling.
  - *How might review criteria be modified to help address these issues?*

# WG2 Overview of Work

- Series of meetings from June 2020 - March 2021
- Discussion centered around several themes:
  - Complexity of current criteria
  - Issues of special importance for clinical trials
  - Ensuring rigor and reproducibility
  - Promoting diversity, reducing bias
  - Clarifying the criterion “Innovation”
  - Desirability of general review criteria



# Discussion Themes: Complexity of current criteria

- RPG criteria have become too complex.
  - Additional review considerations distract reviewers from scientific merit
  - The CT criteria definitions add extensive questions to each major criterion.
- The Human Subjects and Clinical Trials Information (HSCTI) forms required for each CT proposed in an application add considerable burden to applicants and reviewers. However, they rarely drive review outcomes.
- **The complexity of current criteria reduce the quality of peer review.**





## Discussion Themes: Issues of special importance for clinical trials

- The most important additional considerations for clinical trials, compared to non-clinical trials, fall under *Approach*.
- Clinical trials considerations should be also be added under *Significance, Innovation, Investigators, and Environment*.
- No consideration is so unique that it requires a separate set of review criteria



# Discussion Themes: Ensuring Rigor and Reproducibility

- Rigor and reproducibility are critical to clinical trials, but not uniquely so. Rigor and reproducibility are important considerations in all research.
- Additional language is needed to direct reviewer attention to rigor and reproducibility in CTs *and* non-CT review
- Rejected a checklist approach
  - Field specificity is challenging; different considerations for phase III, phase I, mechanistic, BESH? Different fields, methods?
  - Difficult to keep current
  - Checklist criteria -> checklist thinking



# Discussion Themes: Make Criteria General

- Criteria definitions should acknowledge the intelligence and adaptability of reviewers and should encourage high-level scientific judgment by reviewers.
- Criteria should be broadly applicable to research project applications across the full range of NIH science
- Use a single set of review criteria for clinical trials and non-clinical trials applications.



# Discussion Themes: Disparities, Diversity and Bias

- The WG was very disturbed by the persistent NIH funding gap for Black and other URM scientists. The lack of progress is frustrating, and urgent action is needed.
- Charge of this group is to evaluate peer review criteria
  - The fundamental purpose and role of peer review at NIH is to evaluate the scientific and technical merit of applications.
  - Modifications to criteria should be focused on improving review to better, more fairly, identify the best science.
- **Bias distorts judgments of merit; conversely, better judgments of merit reduce the impact of bias.**



## Discussion Themes: Diversity and Bias -2-

- Members observed that instances of overt negative bias toward minority PIs are rare. Positive, and potentially biasing statements about the PI's or institution's reputation are common.
- Halo effects tend to favor established senior investigators and elite research institutions; these effects tend to favor non-Hispanic White PIs.
- Regarding proposals to redefine *Innovation* in terms of diversity, or to include evaluation of diversity as part of *Investigators*: While supportive of efforts to improve demographic diversity in science, concerns were expressed:
  - challenges with definition and implementation
  - another “diversity tax”
  - doubts about efficacy



# Major Outcomes and Recommendations

- ❖ A new proposed set of review criteria applicable to all R01s, both clinical trial and non-clinical trial applications
  - Starts with the 3 Factor Review Criteria structure proposed by WG1
  - Modifies each factor and the underlying criteria with the aims of
    - improving review of rigor and reproducibility
    - reducing positive bias in review
    - clarifying the criterion innovation
  
- ❖ WG2 recommends dropping the additional HSCTI form, dropping the criterion “Timeline” and shifting most “additional review considerations” out of peer review



# Proposed Structure for Review Criteria



Reorganize the five core review criteria into 3 factors to focus reviewers' attention on the big picture questions that should drive scores:

## 1. *Should it be done?*

- **Factor 1 “Importance of the Science”**
- Consider the criterion *Significance*

## 2. *Can it be done well?*

- **Factor 2 “Feasibility and Rigor”**
- Consider the criteria *Approach* and *Innovation*

## 3. *Will it be done?*

- **Factor 3 “Investigators and environment”**
- Consider the criteria *Investigators* and *Environment*

Factor 1. Importance of the Science:	1-9
Factor 2. Feasibility and Rigor	1-9
Factor 3. Investigators and Environment	1-9
Overall Impact Score	1-9

# Recommendations for improving review of Rigor and Reproducibility

- Factor 1 should include language directing attention to the empirical foundation of the proposal.  
“Evaluate the rationale for undertaking the study. An empirical foundation is important to most studies and critical to clinical trials. Evaluate the rigor of the scientific background.”
- Factor 2 should be named “Feasibility and Rigor”
- In defining *Approach*, add a bullet that highlights important general elements of rigor and reproducibility





# Recommendations to reduce bias

1. Endorsed the proposal of WG1 to have a distinct Factor 3 Investigators and Environment

This structure lends itself to 2-stage, partially blinded review.

2. Modify the criterion definitions for Investigator and Environment to reduce positive bias.

*“Rather than general reputation, consider the specific strengths or weaknesses of the investigators and environment with respect to the specific science proposed.”*

Reducing halo effects that favor the status quo was seen as a point of potential effective change.



# Recommendations regarding the criterion Innovation

- *Innovation* is relevant to both Factors 1 (Importance) and 2 (Feasibility & Rigor). WG1 proposed including Innovation under both factors. WG2 thought listing it as a criterion twice is confusing. Instead:
  - incorporate the concept of innovation into the definition of *Significance* in Factor 1, emphasize the value of “creation”
  - Include the criterion *Innovation* under Factor 2



# Recommendations to reduce burden

- Simplify reviewer responsibility for evaluating the budget
- Relieve reviewers of responsibility for most “additional review considerations”.
  - Biohazards
  - Applications from Foreign Organizations
  - Select Agent Research
  - Resource Sharing Plans
  - Authentication of Key Biological and/or Chemical Resources
- Drop the additional clinical trials criterion “Timeline”
- **Drop the HSCTI forms now required of clinical trial studies**



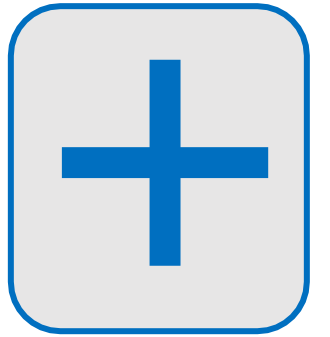
# Advantages: Better Review

- The 3-factor structure directs reviewers to focus on 3 big questions; better focus, better answers
- Respects the intelligence of reviewers and defines for them the judgements NIH asks them to make
- Encourages attention to Significance, Discourages attention to minor weakness
- Should make review fairer by reducing halo effects that contribute to the status quo
- Provides a framework that is easily adapted to 2-stage, partially blinded review models
- Will result in better review of elements critical to rigor and reproducibility
- Provides a consistent framework for review that can be adopted across a wide range of mechanisms and science



## Advantages: Reduced Burden

- Eliminating the HSCTI forms reduces applicant and reviewer burden of CT applications substantially
- Shifting most “additional review considerations” to administrative review reduces reviewer burden
- Frees up reviewer time and attention to focus on scientific merit
- A simplified conceptual framework for criteria should simplify training and reduce the cognitive demands of dealing with multiple sets of detailed criteria.





# Discussion