Evaluation of the NIH Grant Review Process

November 2018

Presentation for the Health Research Alliance

Richard Nakamura, Ph.D., Retired

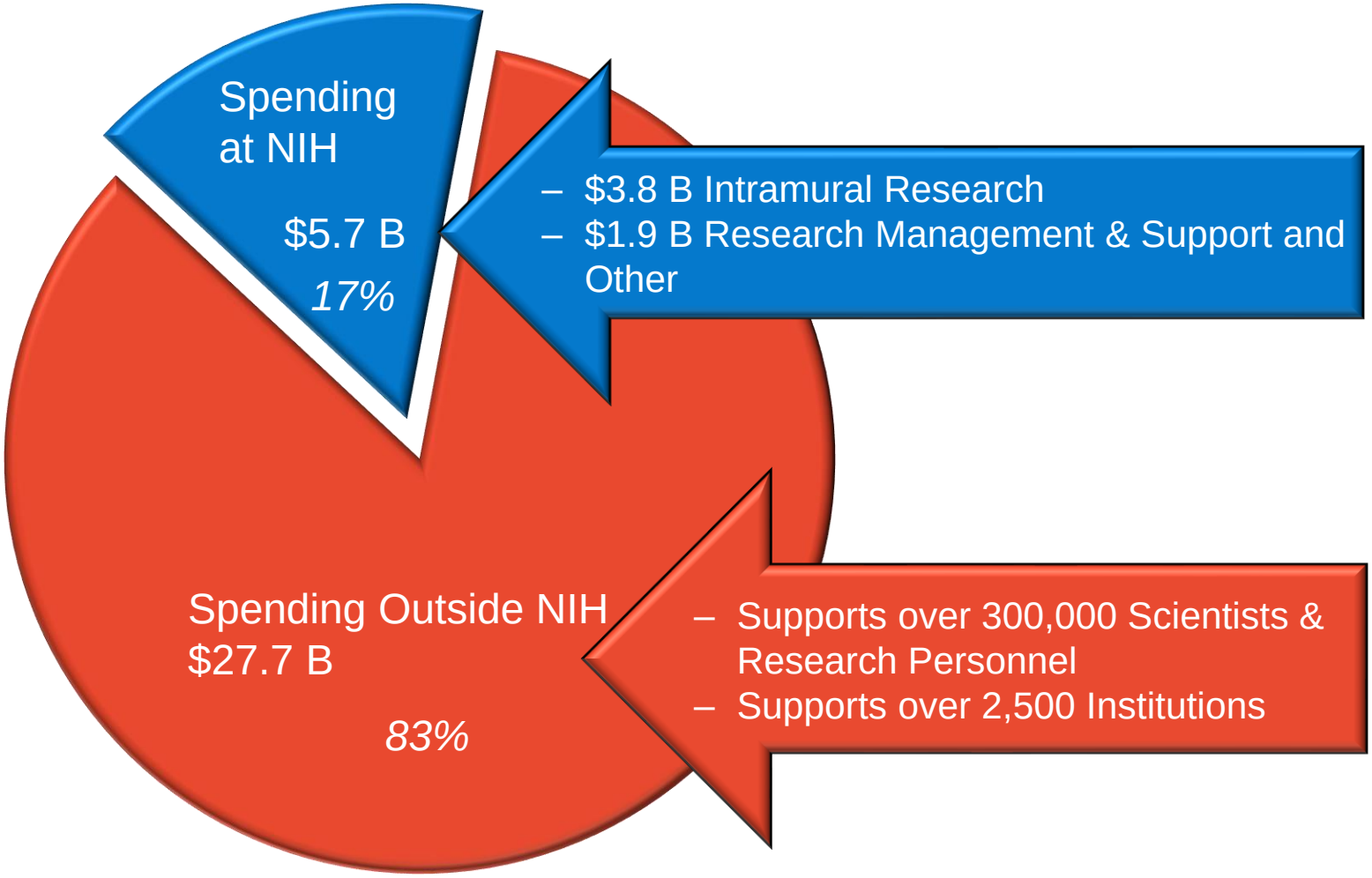
Former Director, Center for Scientific Review, NIH
Former Scientific Director, NIMH
Former Deputy Director, NIMH

NIH . . . A Great Mission Shared Across Science

NIH's mission is to seek fundamental knowledge about the nature and behavior of living systems and the application of that knowledge to enhance health, lengthen life, and reduce illness and disability.

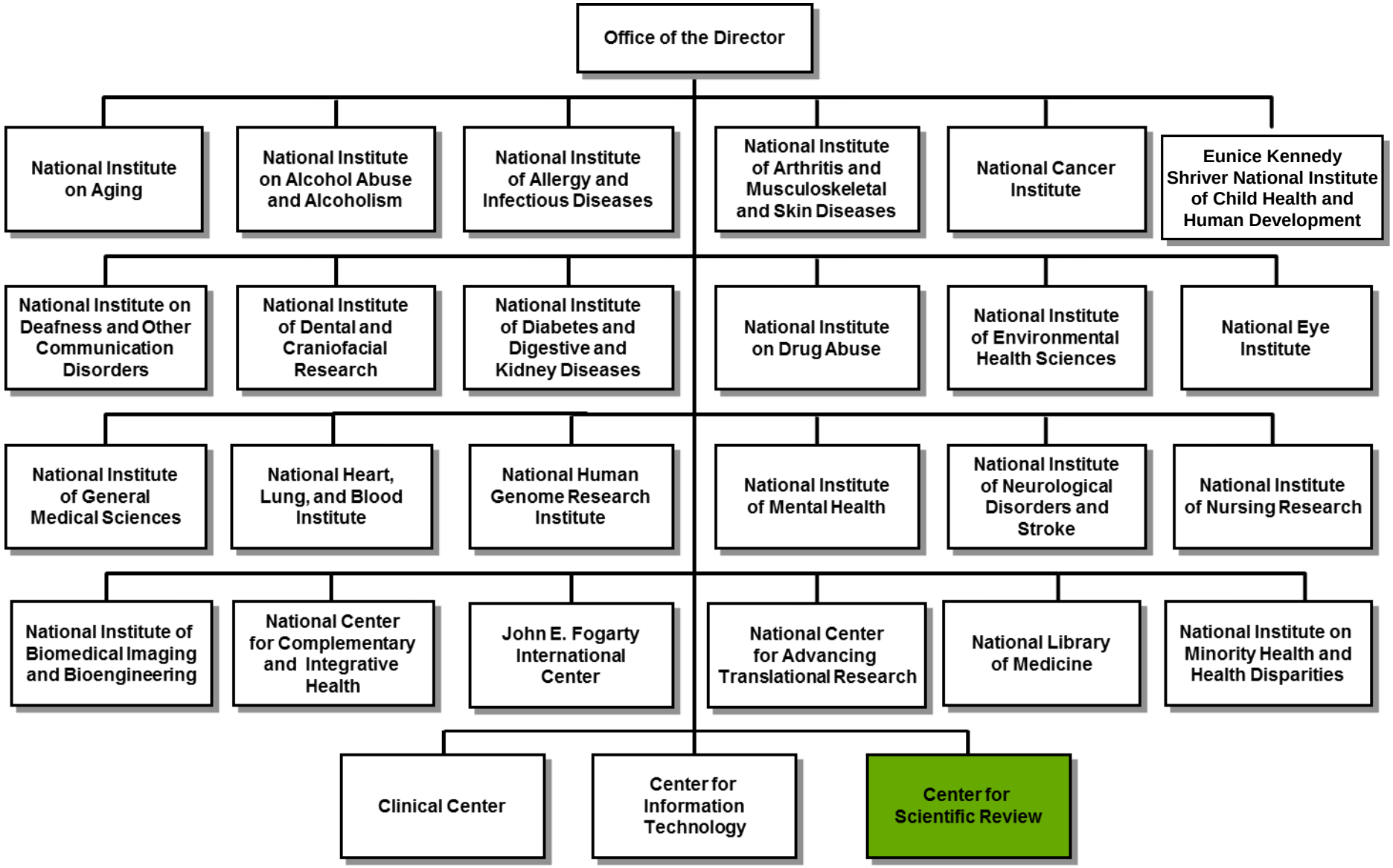
NIH achieves its mission largely through awarding research grants based upon peer review of applications from extramural scientists

NIH Extramural & Intramural Funding
FY 2017 Enacted: $33.4 Billion



NIH Peer Review and Awards

Grant Applications Can Be Funded by One of 24 NIH Institutes or Centers



2-Level Review System for NIH Grants

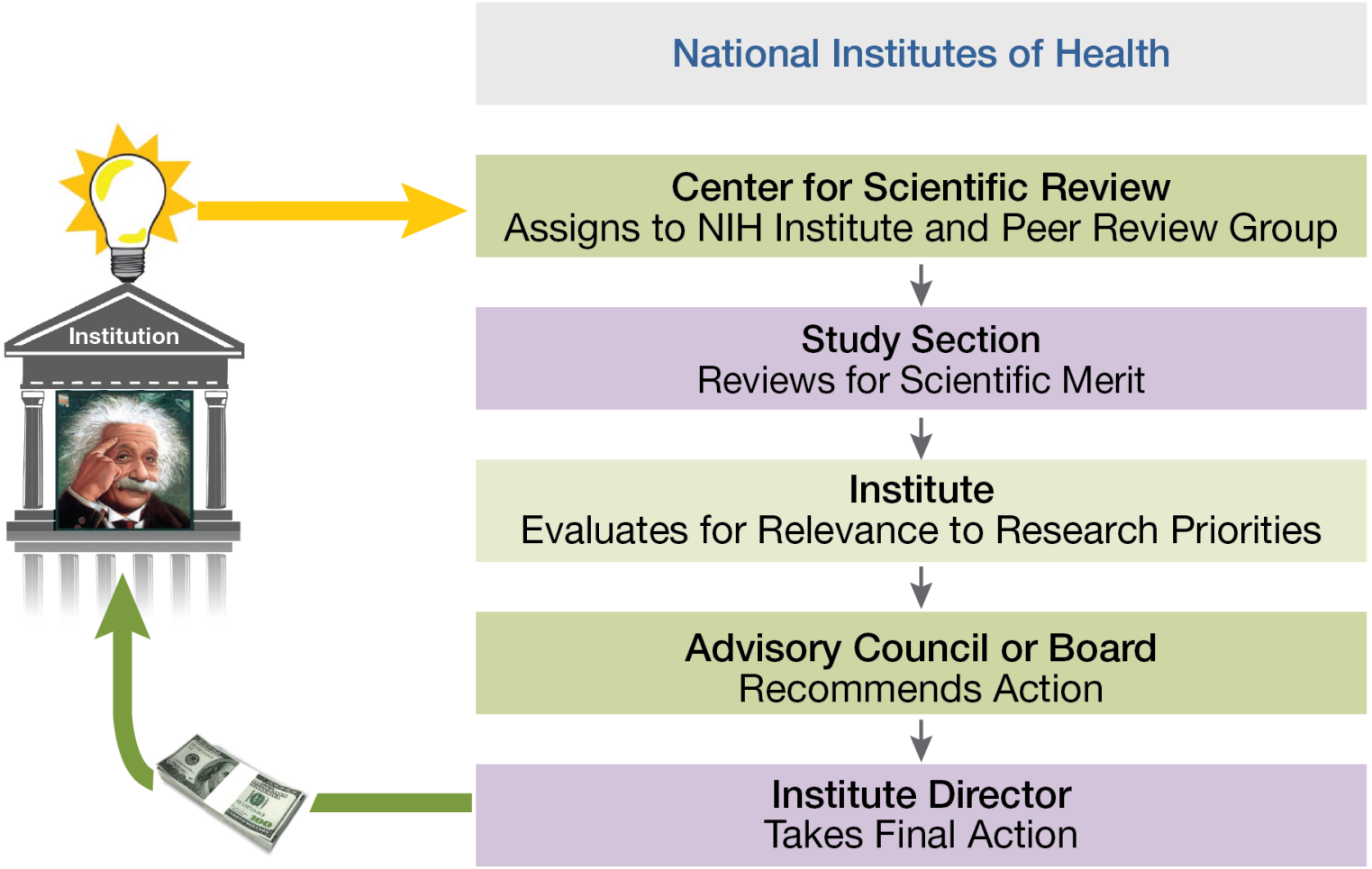
First Level of Review

Scientific Review Group
(Study Section) at CSR or IC

Second Level of Review

NIH Institute/Center Council

Peer Review and Funding of NIH Grant Applications



Role of the Scientific Review Group (SRG) or Study Section

A federal advisory group, consisting of independent scientists, that evaluates grant applications for NIH.

- Provides scientific impact scores and critiques of each grant application under the official supervision of a scientifically trained federal official
- The scores are rank ordered and percentiled and sent along with the application and summary statement to an institute for award consideration

The SRO Convenes the Study Section Meeting

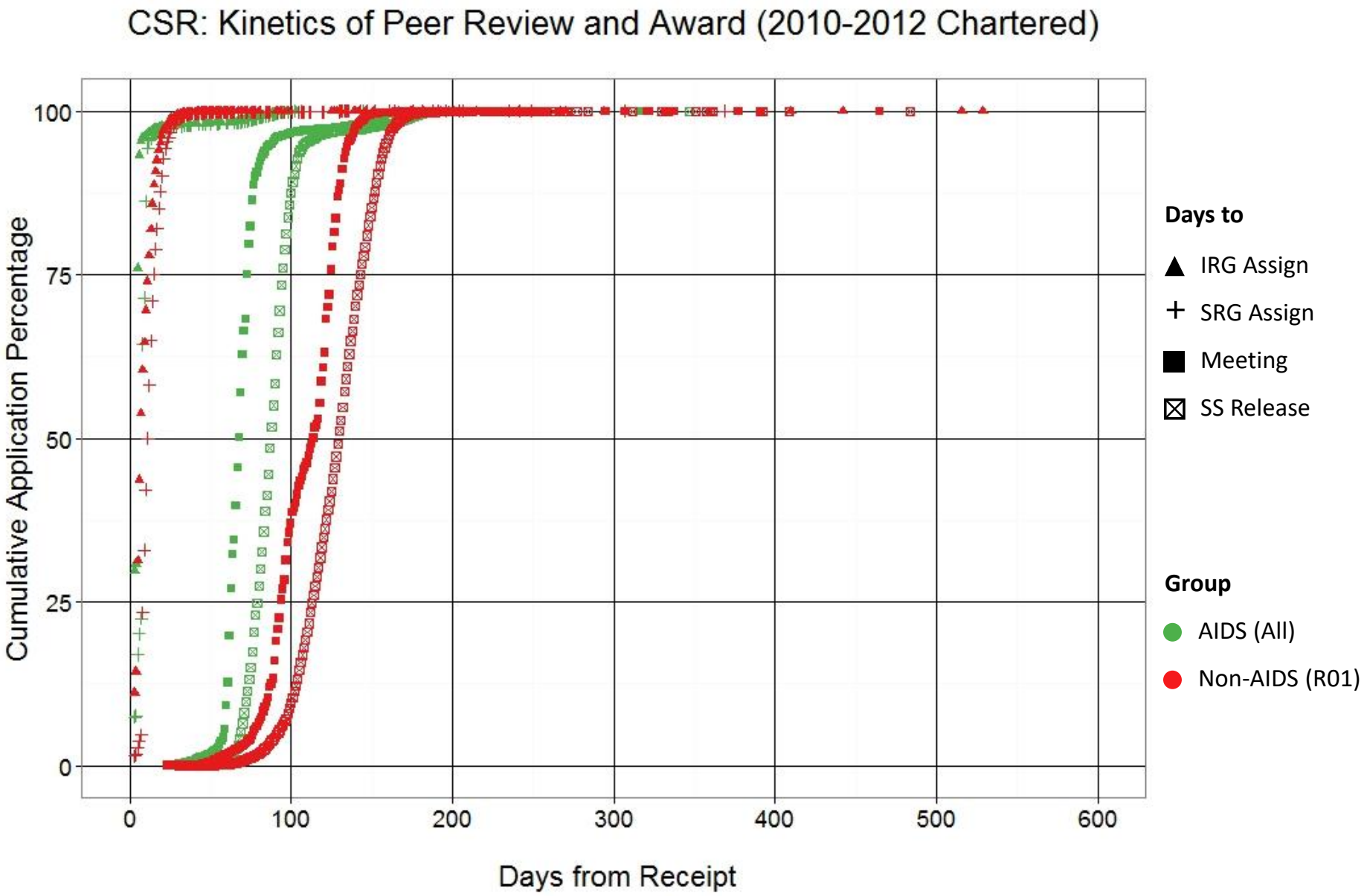
Studying the review process

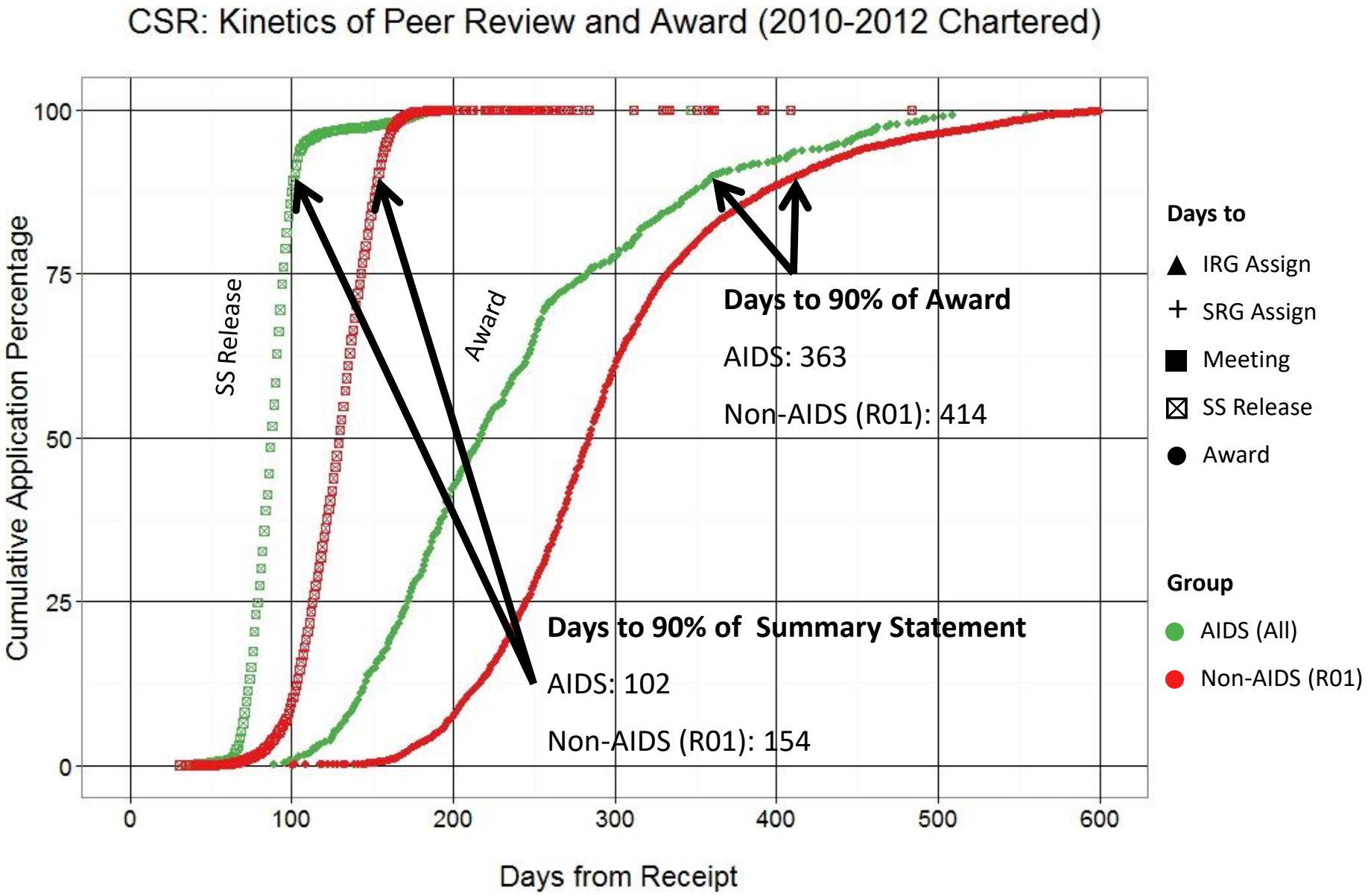
Improving Review: Goals, Measures, Tools

In the last few years, CSR has developed multiple approaches to examine the quality of review via the CSR Research Unit:

- Strategy for Quality Measurement
- CSR Quick Feedback Surveys
- Evaluation of Review Alternatives
- Ranking/scoring Studies
- Assessing Fairness and Reliability in Peer Review

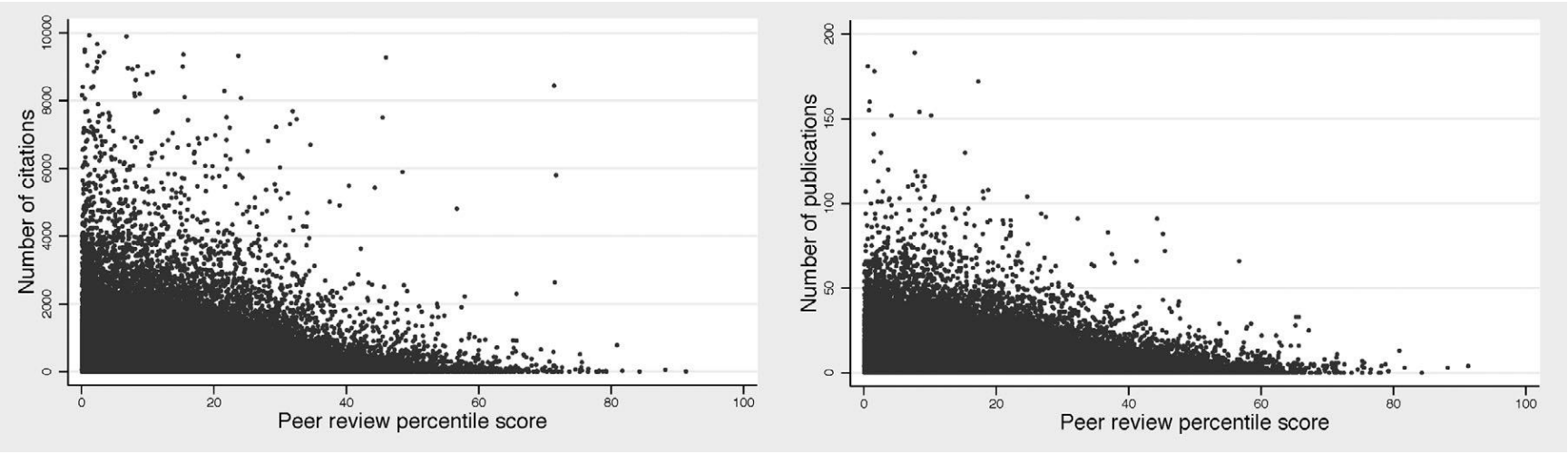
Efficiency: Speed of review


CSR: Kinetics of Peer Review and Award (2010-2012 Chartered)
Cumulative Application Percentage
100
75
50
25
0
0
100
200
300
400
500
600
Days from Receipt
Days to
IRG Assign
SRG Assign
Meeting
SS Release
Group
AIDS (All)
Non-AIDS (R01)

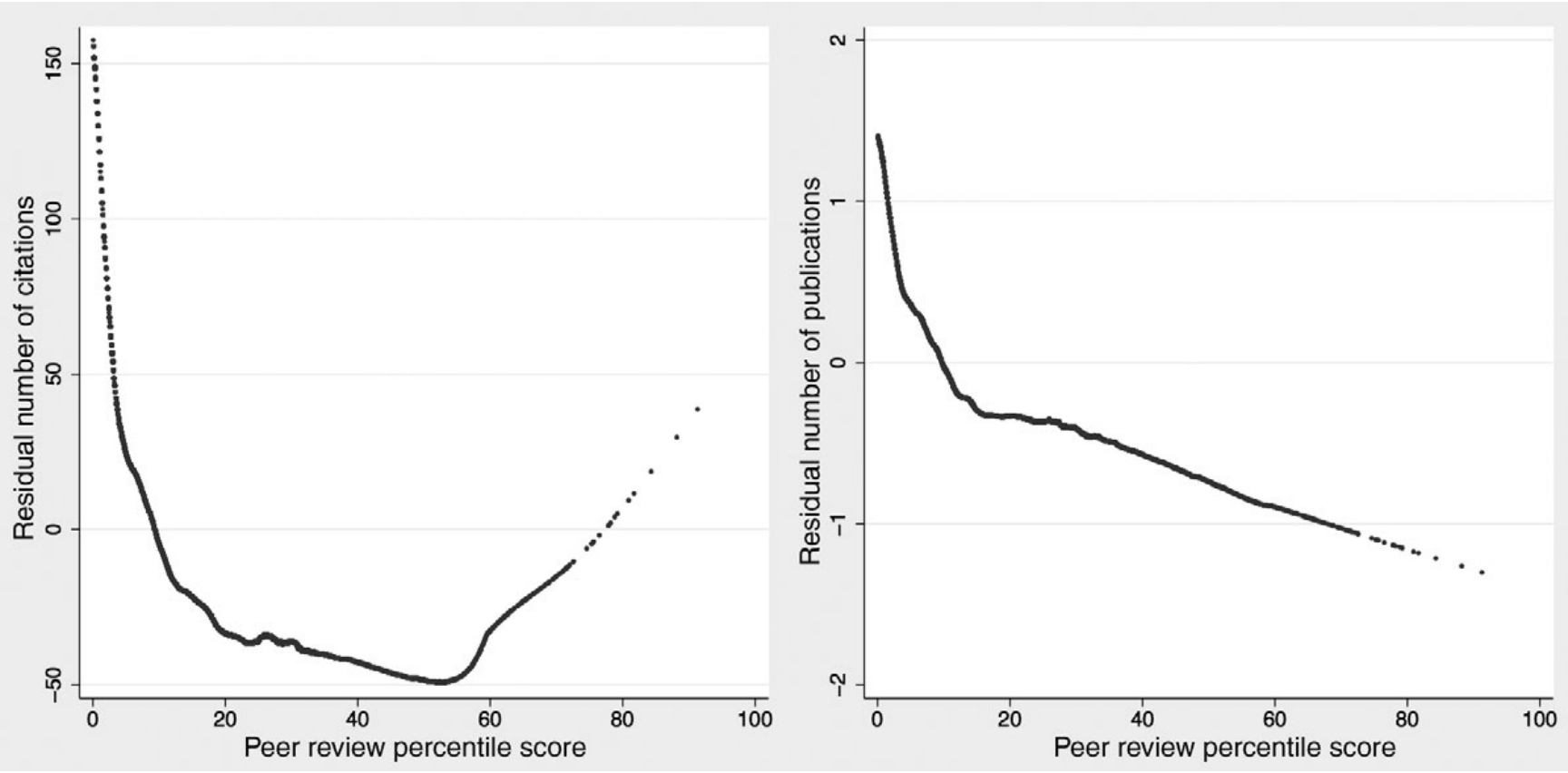

CSR: Kinetics of Peer Review and Award (2010-2012 Chartered)
Cumulative Application Percentage
100
75
50
25
0
0
100
200
300
400
500
600
Days from Receipt
SS Release
Award
Days to 90% of Award
AIDS: 363
Non-AIDS (R01): 414
Days to 90% of Summary Statement
AIDS: 102
Non-AIDS (R01): 154
Days to
IRG Assign
SRG Assign
Meeting
SS Release
Award
Group
AIDS (All)
Non-AIDS (R01)

Judging Performance by Outcome

Positive relationship of score to outcomes

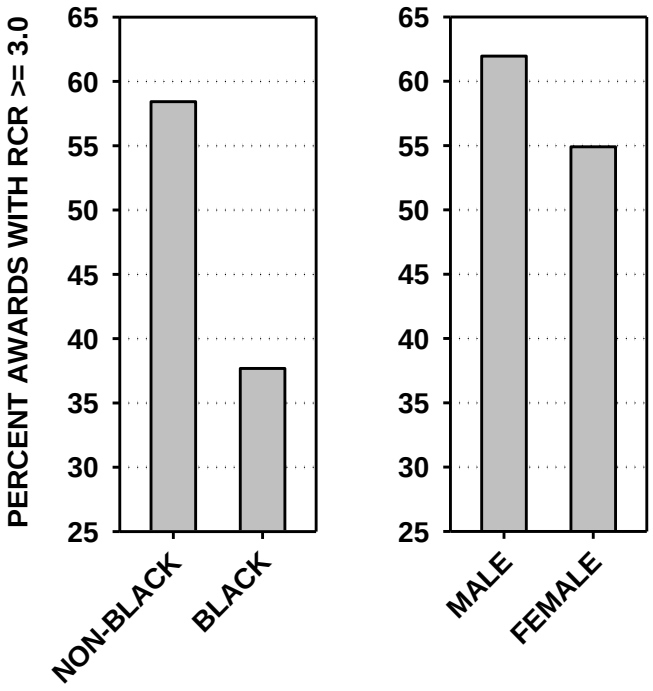


Best rated application are the most productive



Issue of bibliometrics

Women and Black/AA Appear to be Less Productive by Bibliometric Measures



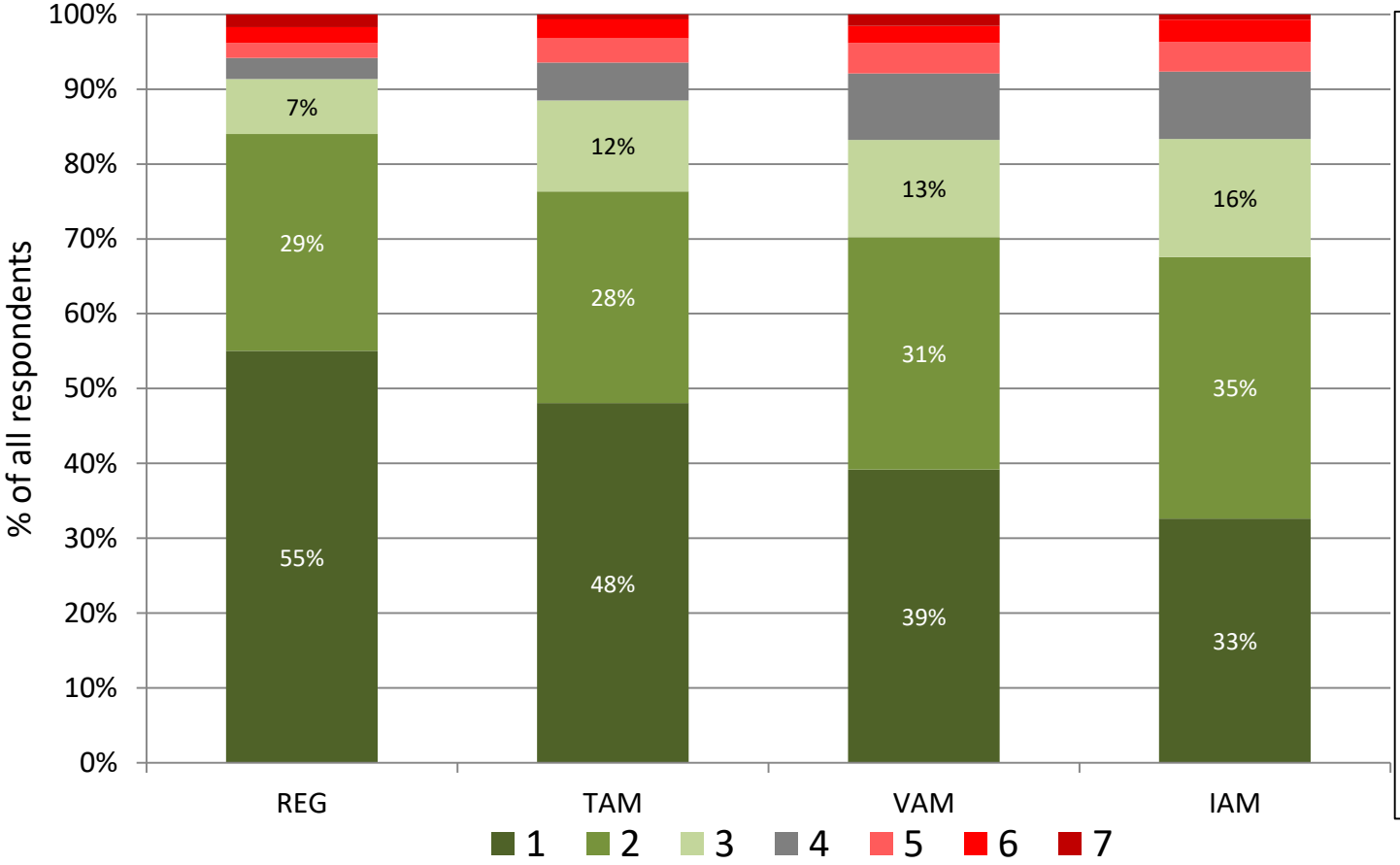
Analysis of Type 1 R01 awards in FY 2010 (N=3,900). Percentage of 'Productive' awards (awards with at least one publication with an RCR ≥ 3.0. Black (*N*=69), Non-black (*N*=3612) Not-Reported (*N*=219, not shown above), χ^2 (2, *N* = 3900) = 12.82, p = 0.0016. Male (*N*=2392), Female (*N*=969) Unknown (*N*=539, not shown above), χ^2 (2, *N* = 3900) = 57.072, p < 0.0001.

Manipulating citation-based measures produces the wrong incentives

1. Publish many papers
2. Publish methods papers
3. Cite as many grants as possible on each paper
4. Use a low threshold for authorship and trade for such favors
5. Encourage students to cite ones papers; specify a preferred paper
6. Discourage citations outside one's "school of thought"

Scientists as Judges

Comparison of Quality of Discussion by Review Format



REG, TAM, VAM - The nature of the scientific discussions supported the ability of the panel to evaluate the applications being reviewed.

IAM - Was the quality of discussions appropriate to evaluate the impact/scientific merit of the applications?

Regular Meetings (REG) - Response rate = 50%, n = 7,094, 267 meetings
Teleconference Meetings (TAM) - Response rate = 40%, n = 1,083, 115 meetings
Virtual Assisted Meetings (VAM) - Response rate = 40%, n = 978, 61 meetings
Internet Assisted Meetings (IAM) - Response rate = 40%, n = 1122, 67 meetings

Likert type scale where 1 = very strongly agree to 7 = very strongly disagree

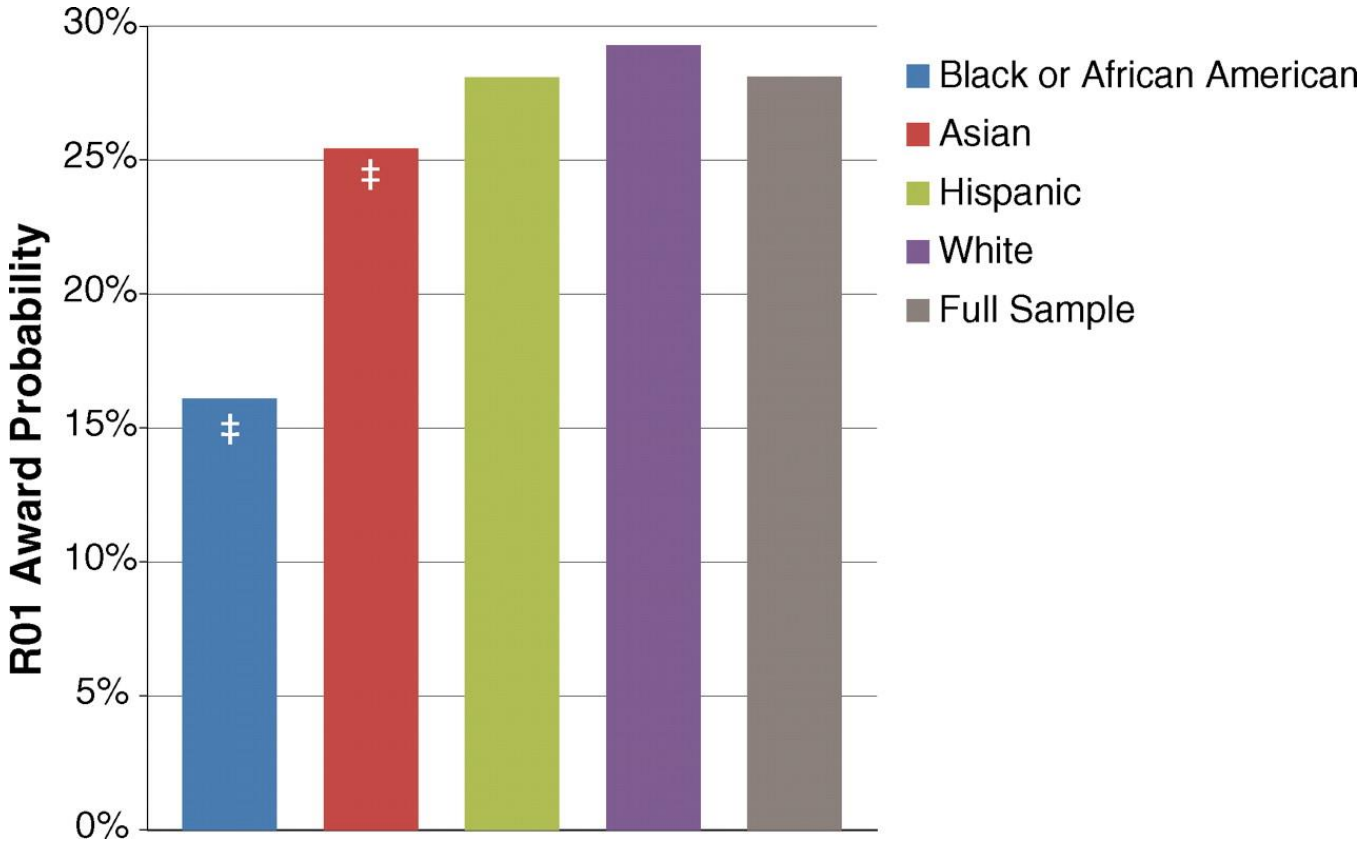
Fairness of Review

Fairness of Review

- In general, review is pretty fair
- Stage of career
- Gender
- Field
- Race/Ethnicity
- Reviewers

Concerns of Reviewer Bias

Probability of NIH R01 award by race and ethnicity, FY 2000 to FY 2006 (N = 83,188)



D K Ginther et al. Science 2011;333:1015-1019

Science
AAAS

The Hunt for Bias

An Anonymizing Study

Anonymization Experiments – Basic Assumptions

- Racial disparities in grant funding exists (Ginther et al): AAs award rates much lower than Whites. Other biases are suspected.
- Average preliminary overall impact scores account for variance in final scores that account for award disparity.
- The major hypotheses for score disparity are:
  - Reviewer bias and/or
  - Quality of application submission

Specific Aims

1. **To determine if masking PII information from grant applications reduces the differences in final scores for Black and White applicants.**
2. To determine if this reduces the differences in final scores for Male and Female applicants. (secondary)
3. To determine if this reduces the differences in final scores for Established and ESI applicants. (secondary)
4. To determine if this reduces the differences in final scores for applicants from more research intensive and less research intensive institutions. (secondary)

Discussion