

# Exploring New Approaches to Optimizing Peer Review

*106<sup>th</sup> Meeting of the  
Advisory Committee  
to the Director  
June 14<sup>th</sup>, 2013*

Lawrence A. Tabak, DDS, PhD  
Deputy Director, NIH

Department of Health and Human Services



# Peer Review is Fundamental to the NIH Mission

- The NIH two-tier peer review system is the foundation on which the agency's funding of extramural research is based
- While this system is highly regarded throughout the world, it is vital that NIH continue to innovate and optimize the process by which grant applications are reviewed

# Continuous Review of Peer Review



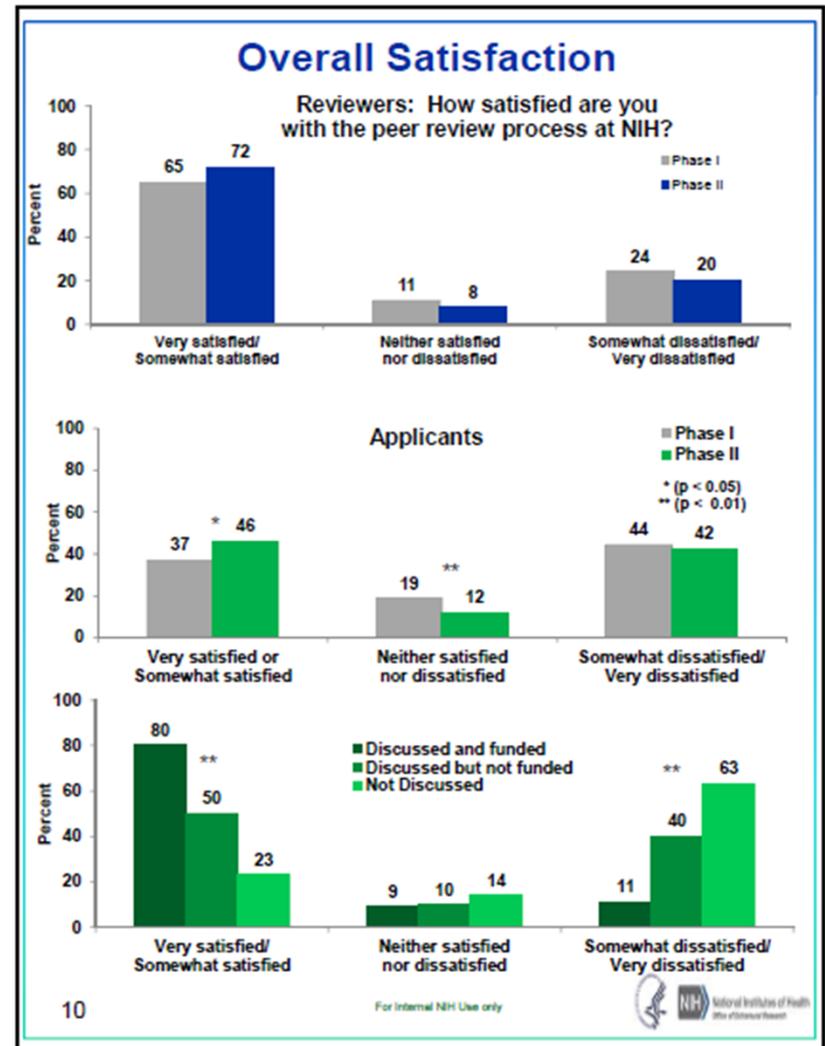
## Enhancing Peer Review Survey Results Report



Published May 2013



[review.nih.gov/docs/Enhancing\\_Peer\\_Review\\_Report\\_2012.pdf](http://review.nih.gov/docs/Enhancing_Peer_Review_Report_2012.pdf)



# Continuous Review of Peer Review

## **ACD Working Group on Diversity in the Biomedical Research Workforce**

NIH should establish a WG of the ACD comprised of experts in behavioral and social sciences and studies of diversity with a special focus on determining and combating real or perceived biases in the NIH peer review system (Recommendation #9)

NIH should first, pilot different forms of validated implicit bias/diversity awareness training for NIH scientific review officers and program officers to determine the most efficacious approaches. Once the best training approaches have been identified with NIH staff, pilot these programs with members of study sections to ascertain if their value is sustained. If they are, provide to all study section members (Recommendation #10)

# Additional Issues to Consider

- Concern has been expressed that the structure of CSR Integrated Review Groups (IRGs), taken together with our dependence on normalized percentiles across all IRGs, may lead to the funding of some applications that are not of highest “priority” (i.e. quality, novelty, and alignment with core mission of the IC or Agency).
- In theory, select pay, or high and low program relevance could be used to address this issue. However, historically, ICs generally adhere closely to percentile/score rankings.

# Additional Issues to Consider (cont.)

- Should a portion of NIH resources be redirected in a more systematic way to ensure that we consistently support the “best” opportunities?
  - To address this question, should we systematically evaluate the characteristics of study section “performance”?
- Our current IRG organization is driven by the nature and number of applications being submitted.
  - Should we be more proactive in attempting to identify emergent fields of science to ensure optimal review of the freshest ideas?

# New Approaches to Enhancing Peer Review

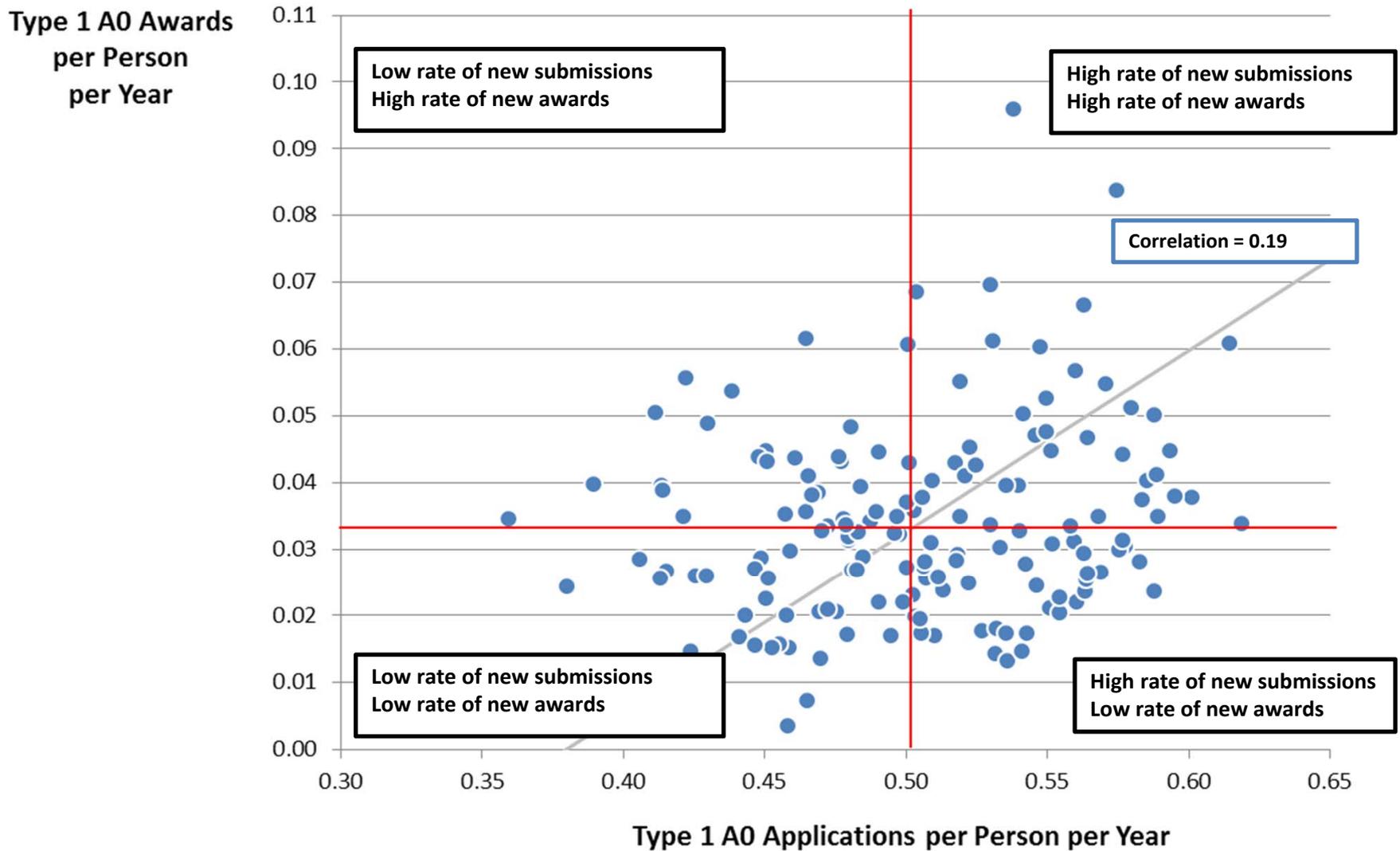
- A team consisting of leadership and staff from the Center for Scientific Review, the Division of Program Coordination, Planning, and Strategic Initiatives, and the Office of Extramural Research, was convened on behalf of the NIH Director in January 2013 to:
  - overseeing the development of methods to identify emergent, highly active, and areas of science that may have stagnated.
  - recommending approaches to couple the “state” of scientific fields to study section organization to yield a more optimized and dynamic system that is responsive to changes in scientific trends.

# Possible Quantitative Approaches Being Explored

- Analysis of study section “inputs”
  - Examine the number of new applications, the number of new awards, and the relationship between the two for different study sections, while controlling for their different sizes

# Application and Award Rates

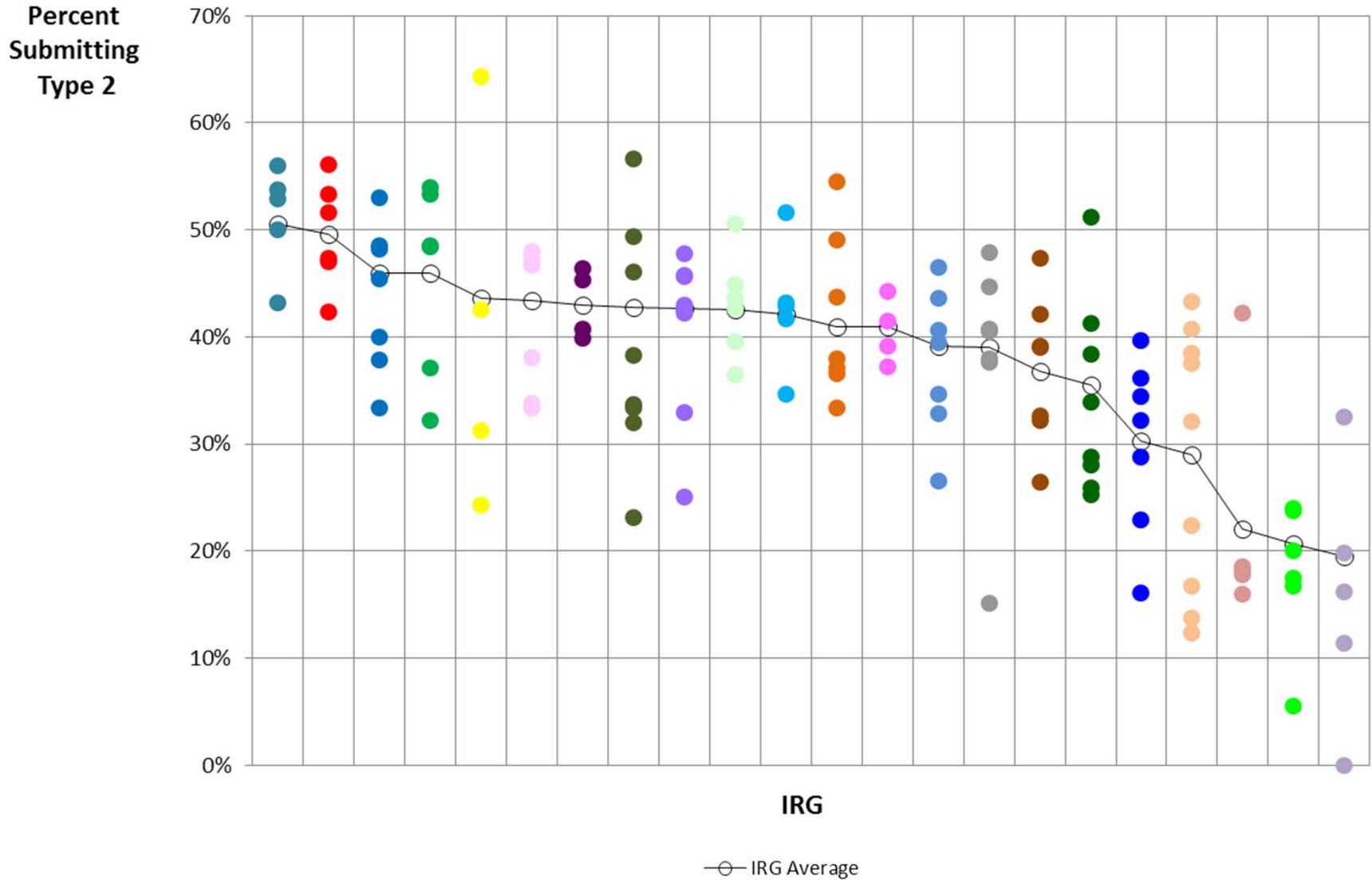
## Unsolicited Type 1 R01s (A0s)



# Possible Quantitative Approaches Being Explored

- Analysis of study section “inputs”
  - Examine the number of new applications, the number of new awards, and the relationship between the two for different study sections, while controlling for their different sizes
  - Percent of Awardees who submit competing renewals by IRG

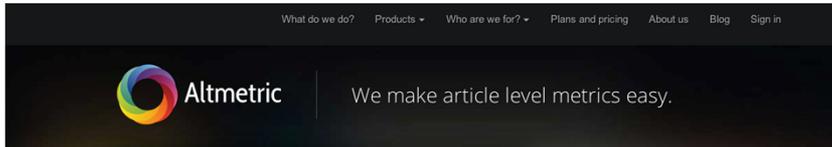
# Percent of Awardees Who Submit Competing Renewal (FY 2004-2007 Awardees)



# Possible Quantitative Approaches Being Explored

- Analysis of study section “inputs”
- Tracking indicators of emergent fields
  - “Word bursts” in literature, applications – which precede widespread adoption could indicate a new research area
  - The appearance of new investigators in applications to the study section
  - Citation analysis of applications –emerging areas tend to cite interdisciplinary references
  - “Altmetrics”

# “Altmetrics”



1. Add bookmarklet to your bookmarks toolbar
2. Visit any paper
3. Get article level metrics with a single click



## THE CHRONICLE OF HIGHER EDUCATION

June 8, 2013

Steven B. Roberts's 103-page tenure package features the usual long-as-your-arm list of peer-reviewed publications. But Mr. Roberts, an assistant professor at the University of Washington who studies the effects of environmental change on shellfish, chose to add something less typical to his dossier: evidence of his research's impact online.

He listed how many people viewed his laboratory's blog posts, tweeted about his research group's findings, viewed his data sets on a site called Figshare, downloaded slides of his presentations from SlideShare, and otherwise talked about his lab's work on social-media platforms. In his bibliography, whenever he had the data, he detailed not only how many citations each paper received but how many times it had been downloaded or viewed online. The strategy was part of "an attempt to quantify online science outreach," he explained in his promotion package.

*By Jennifer Howard*

<http://www.altmetric.com>

# Possible Quantitative Approaches Being Explored

- Analysis of study section “inputs”
- Tracking indicators of emergent fields
- Analysis of the study section “outputs”
  - Bibliometric history of publications (patents), normalized by field of science, attributed to funded applications that were reviewed by an IRG
    - Plot of citations/year vs. journal impact factor as a function of time
      - May reveal if the “performance” of a study section is changing as a function of time

# Welcome to a Lab Meeting



JORGE CHAM © 2006

[www.phdcomics.com](http://www.phdcomics.com)

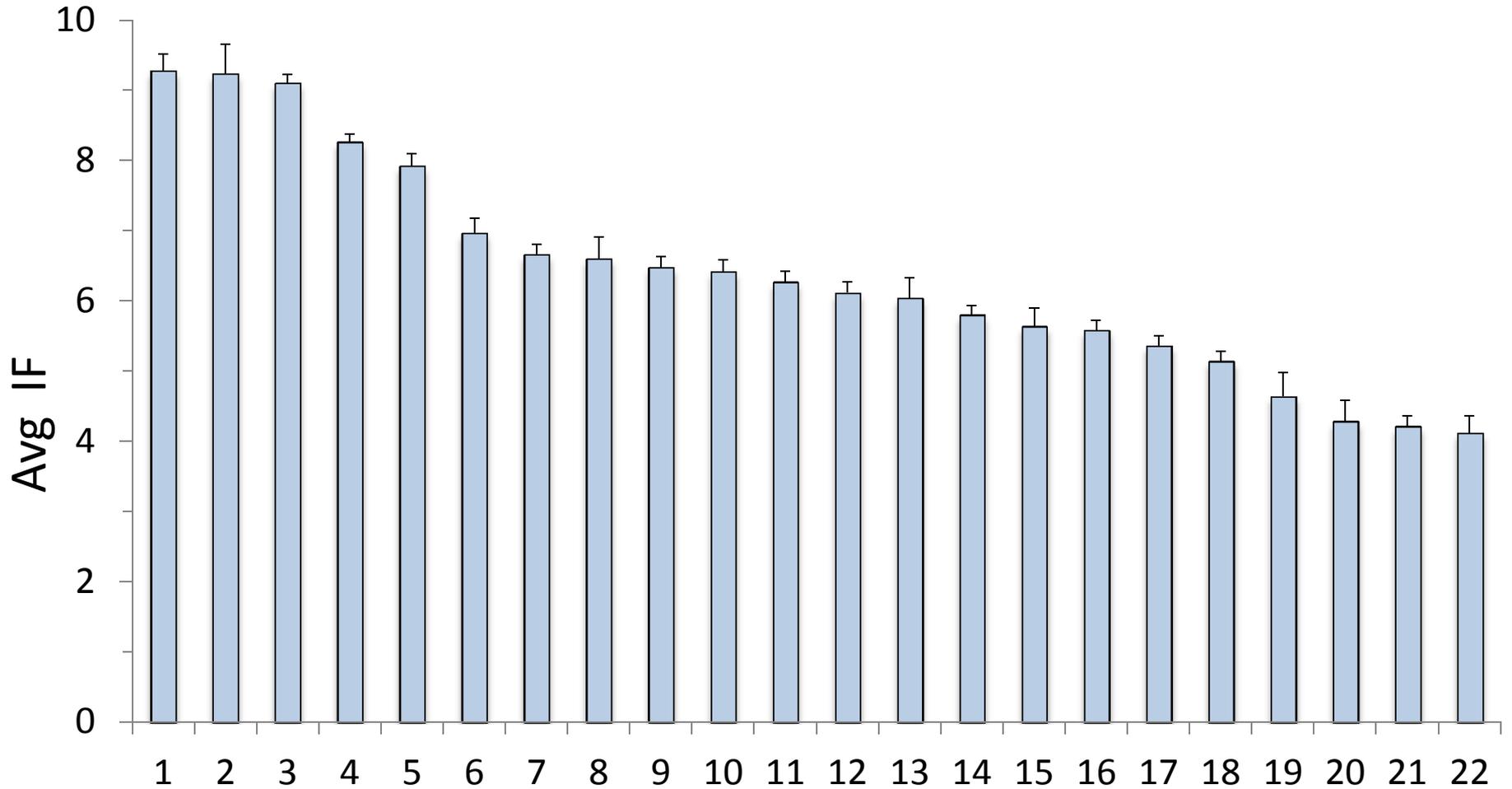
Bibliometric output\* of R01s  
reviewed in CSR standing study sections  
2007-2011

## Overview

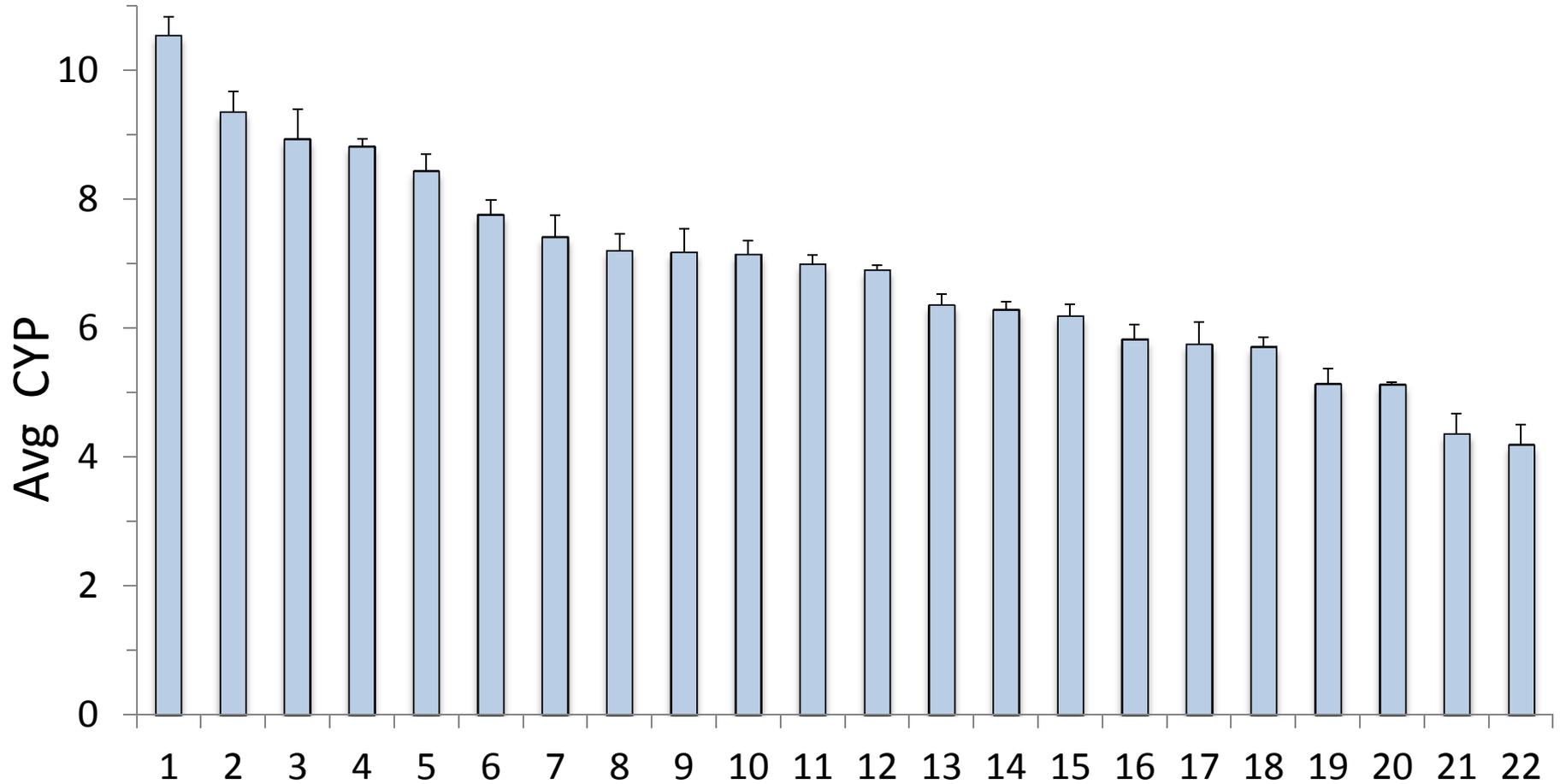
Yr	SSs	R01s	Journals	Pubs	Pubs/R01	Avg IF	Cites/Yr/Pub
2007	137	6879	2338	17520	2.55	6.10	6.81
2008	141	9810	2494	25677	2.62	6.33	7.15
2009	151	11814	2644	32947	2.79	6.43	7.27
2010	155	13068	2748	38419	2.94	6.58	7.13
2011	159	11841	2604	34814	2.94	6.98	6.69
Average:			2,566	29,875	2.77	6.49	7.01

\*A total of 149,377 publications were analyzed

# IRG Average Impact Factor (IF) 2007-2011

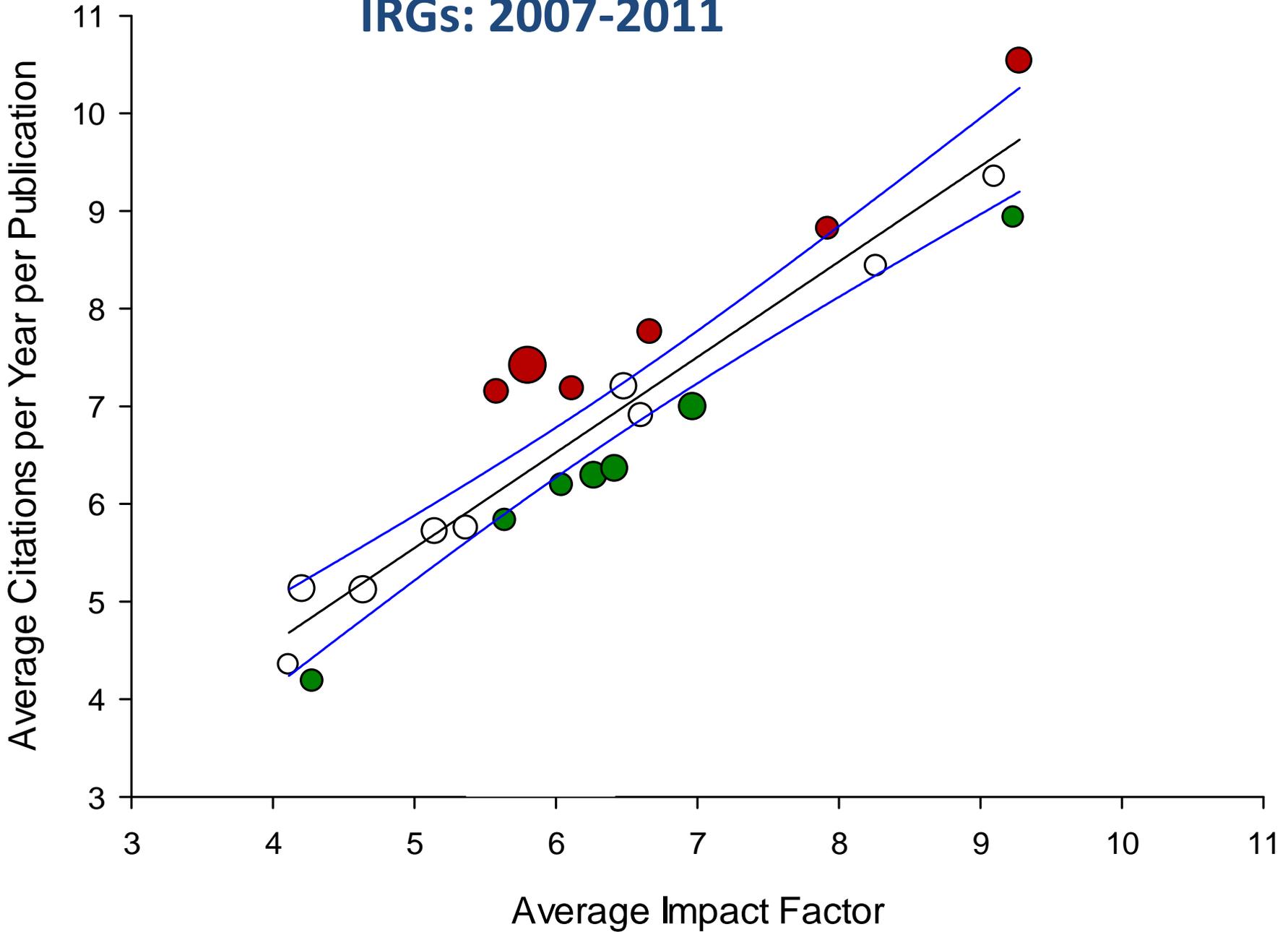


# IRG Average Citations/Year/Publication (CYP) 2007-2011

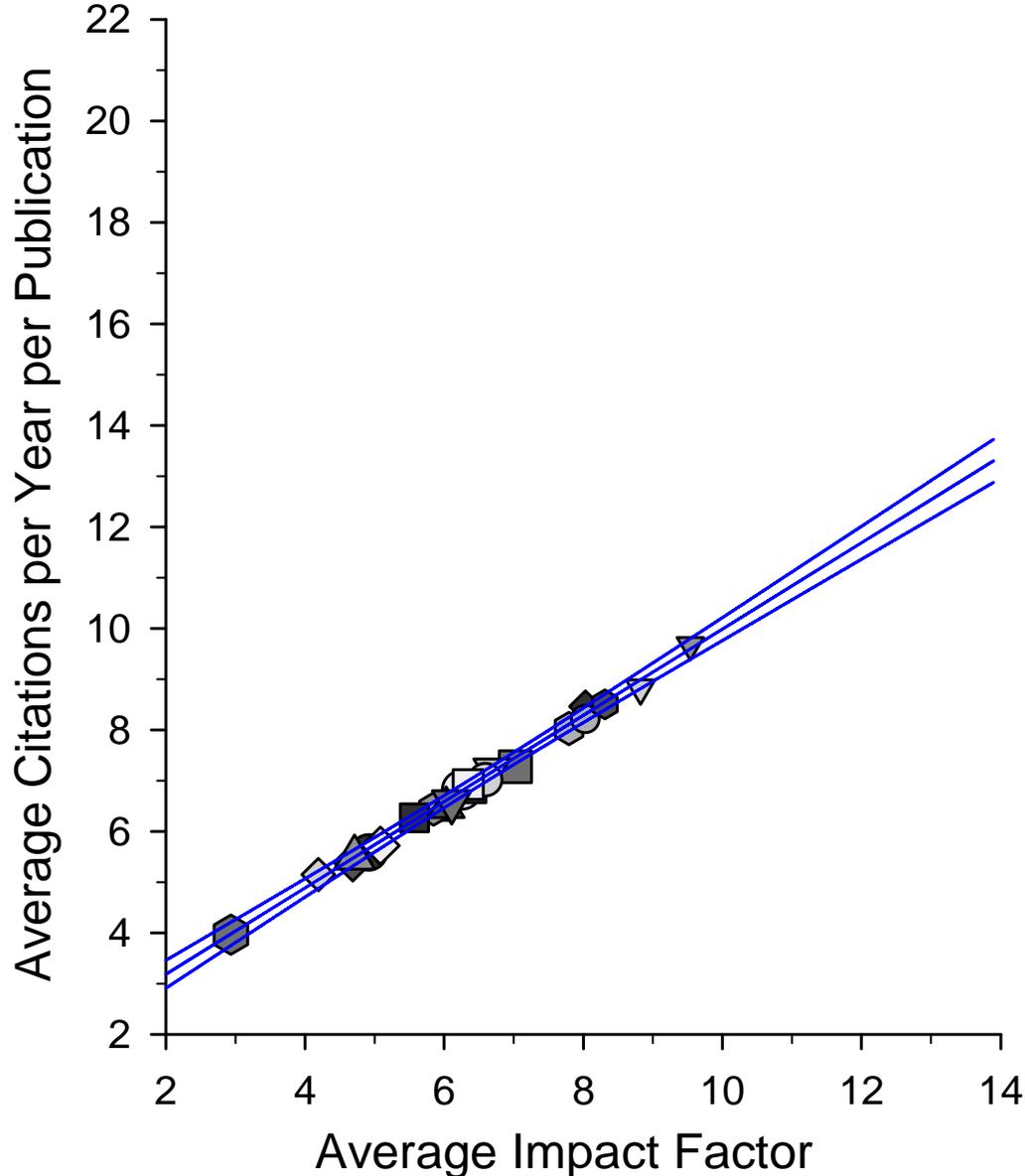


# Observed vs Expected Citation Rate

## IRGs: 2007-2011



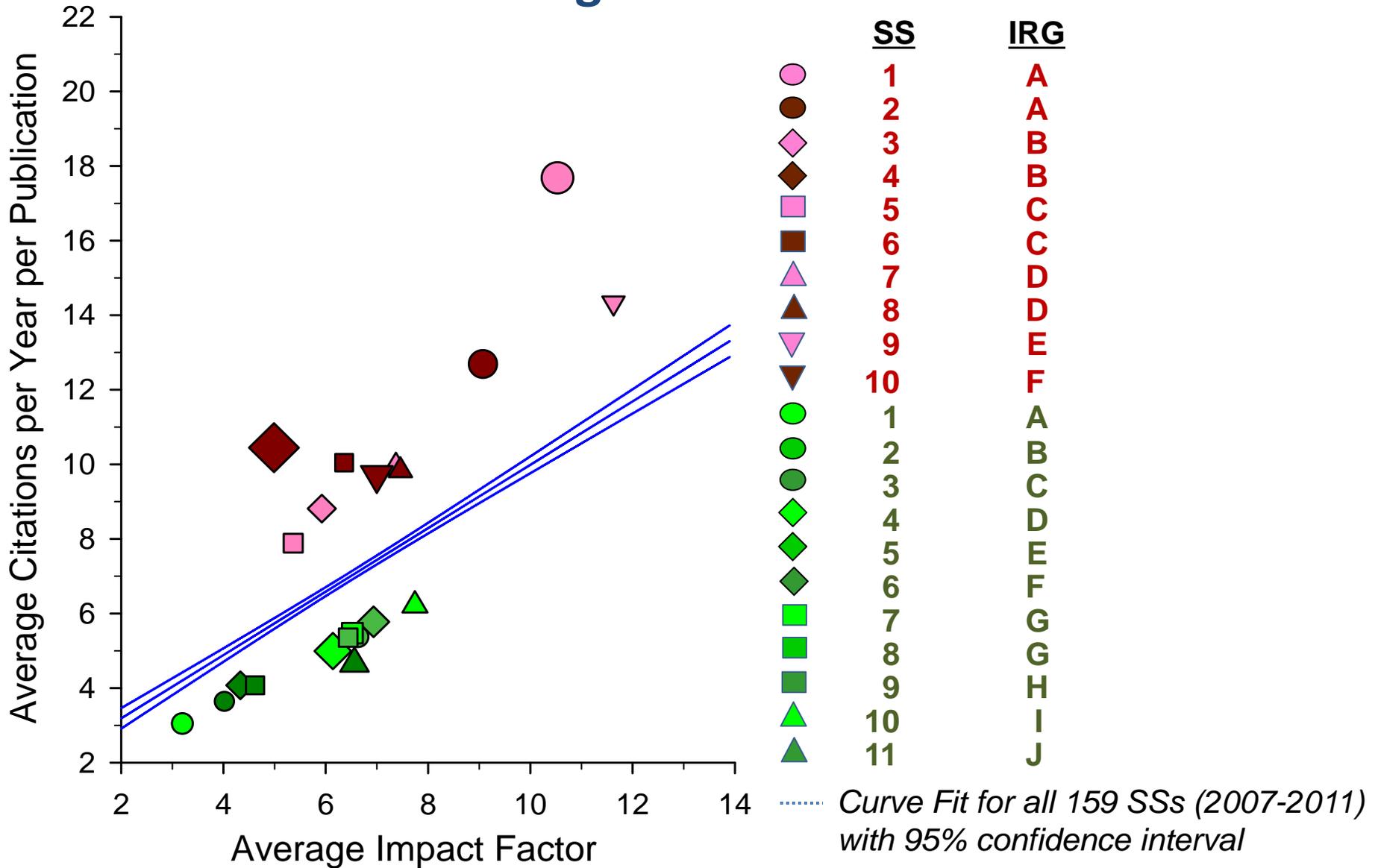
# “Lukewarm” Study Sections 2007-2011



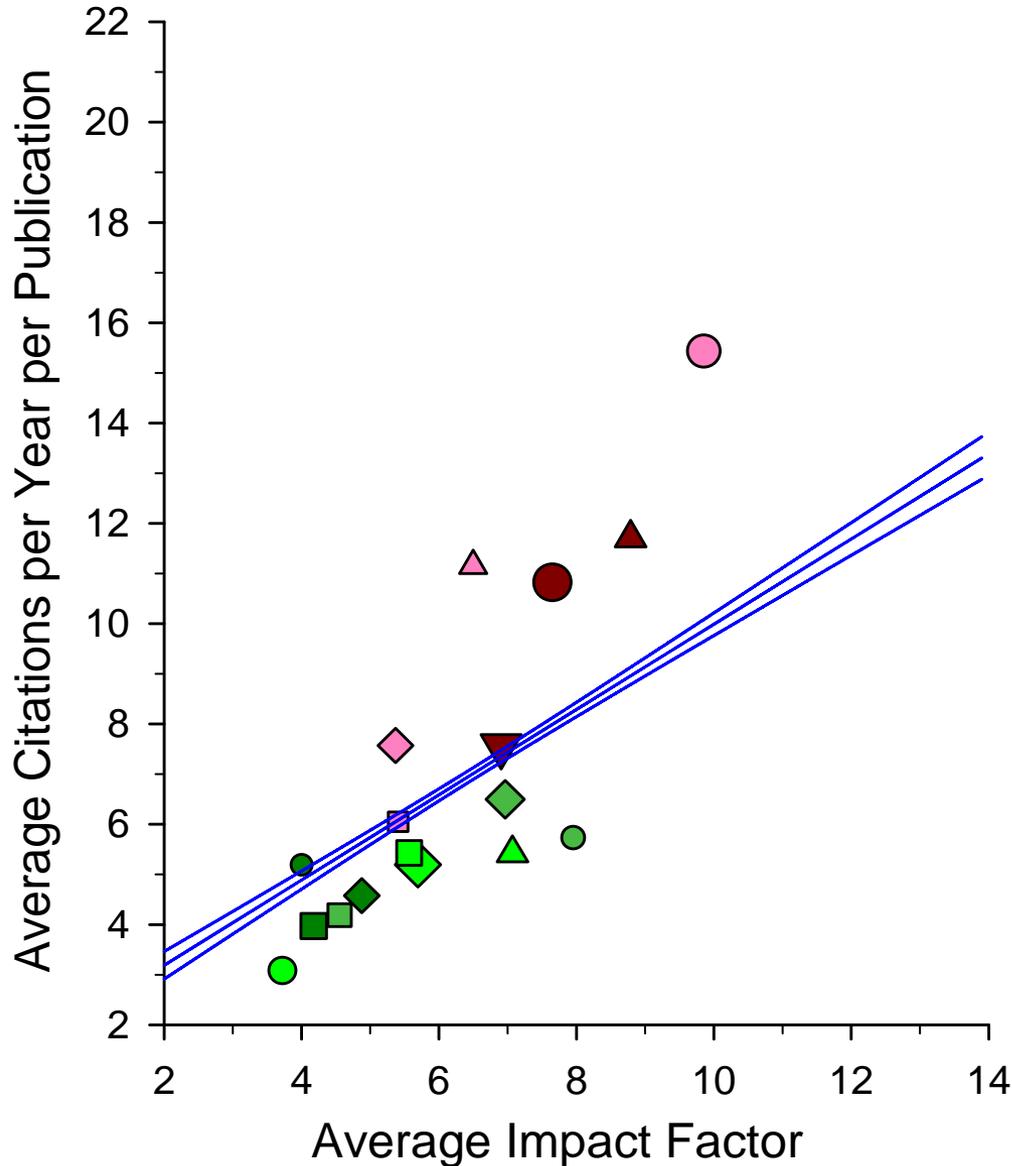
SS	IRG
1	A
2	A
3	A
4	B
5	B
6	C
7	D
8	D
9	D
10	E
11	F
12	F
13	G
14	G
15	H
16	D
17	I
18	J
19	K
20	L
21	K
22	M
23	N

..... Curve Fit for all 159 SSs (2007-2011)  
with 95% confidence interval

# “Hot” and “Cold” Study Sections 2007-2011 Average



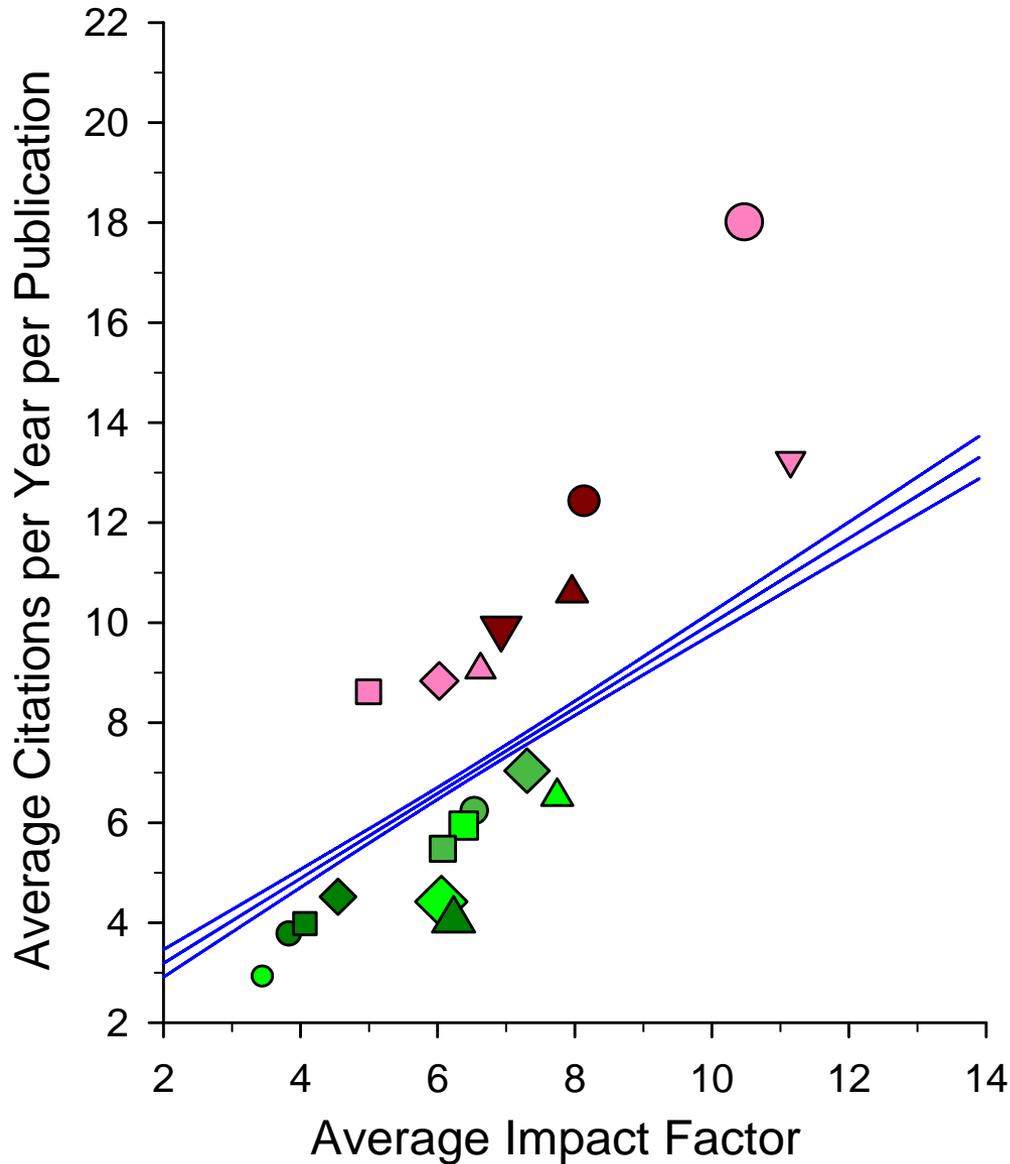
# “Hot” and “Cold” Study Sections 2007



<u>SS</u>	<u>IRG</u>
1	A
2	A
3	B
5	C
7	D
8	D
10	F
1	A
2	B
3	C
4	D
5	E
6	F
7	G
8	G
9	H
10	I

..... Curve Fit for all 159 SSs (2007-2011)  
with 95% confidence interval

# “Hot” and “Cold” Study Sections 2008



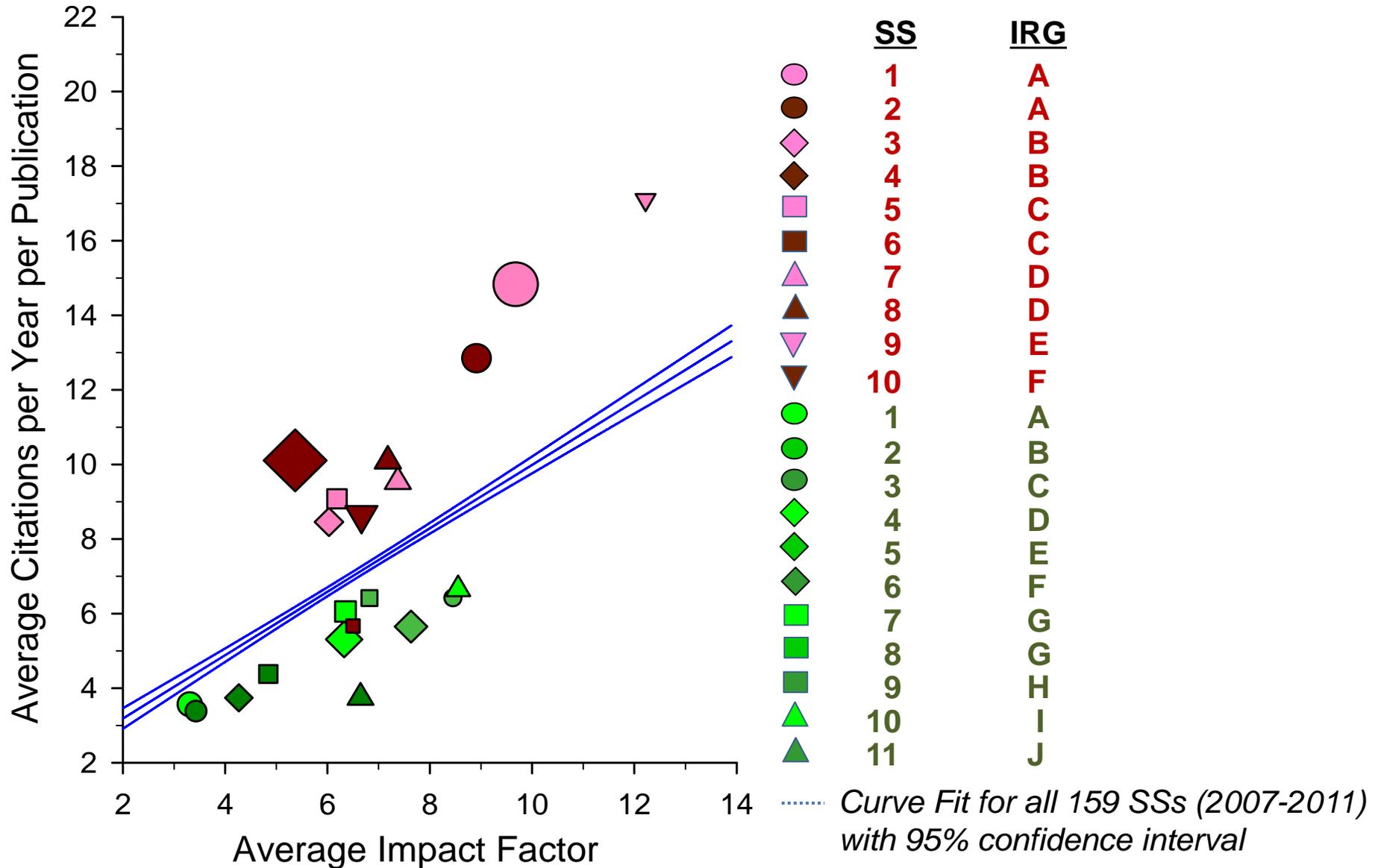
**SS**

- 1
- 2
- 3
- 5
- 7
- 8
- 9
- 10
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11

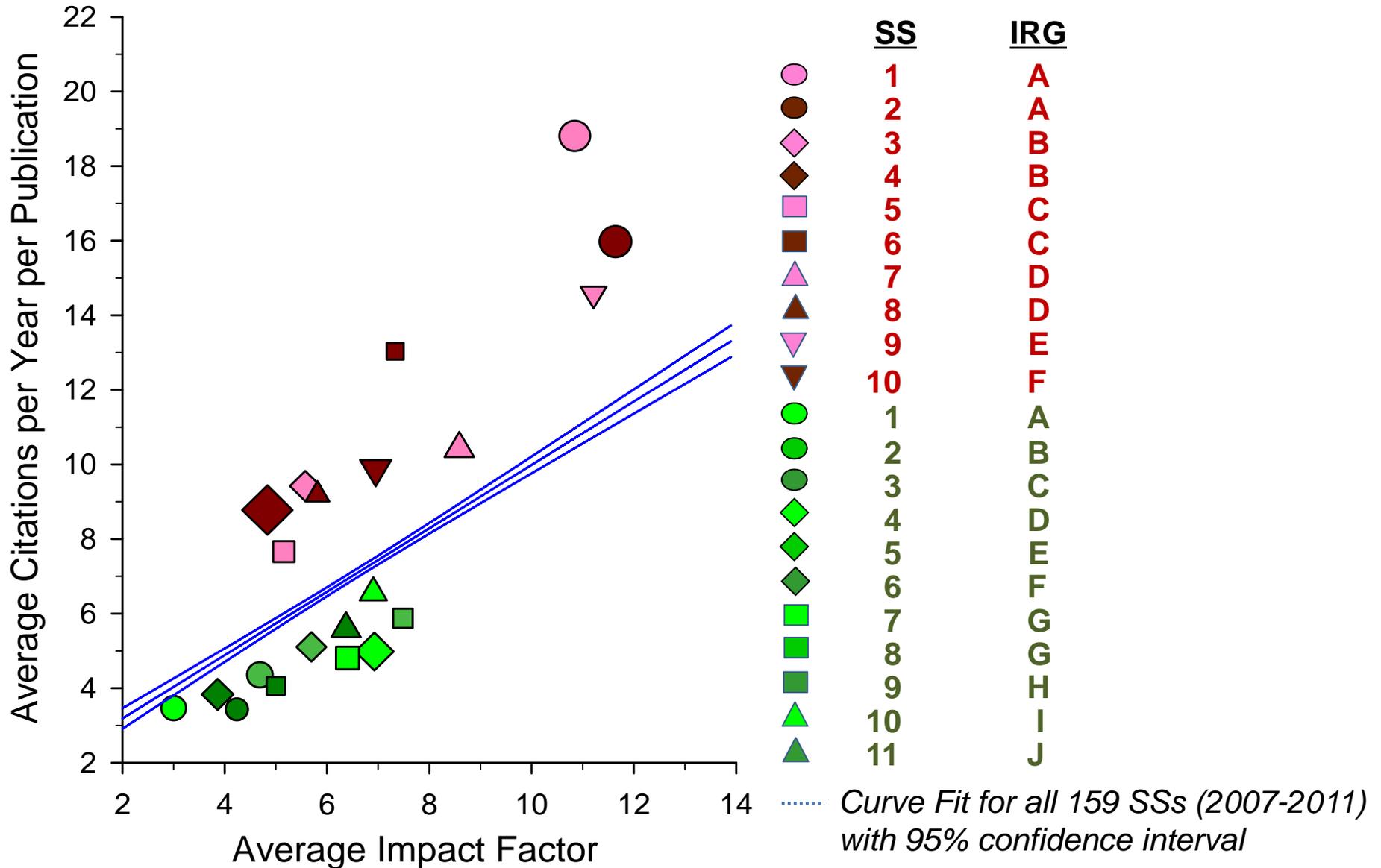
**IRG**

- A
- A
- B
- C
- D
- D
- E
- F
- A
- B
- C
- D
- E
- F
- G
- G
- H
- I
- J

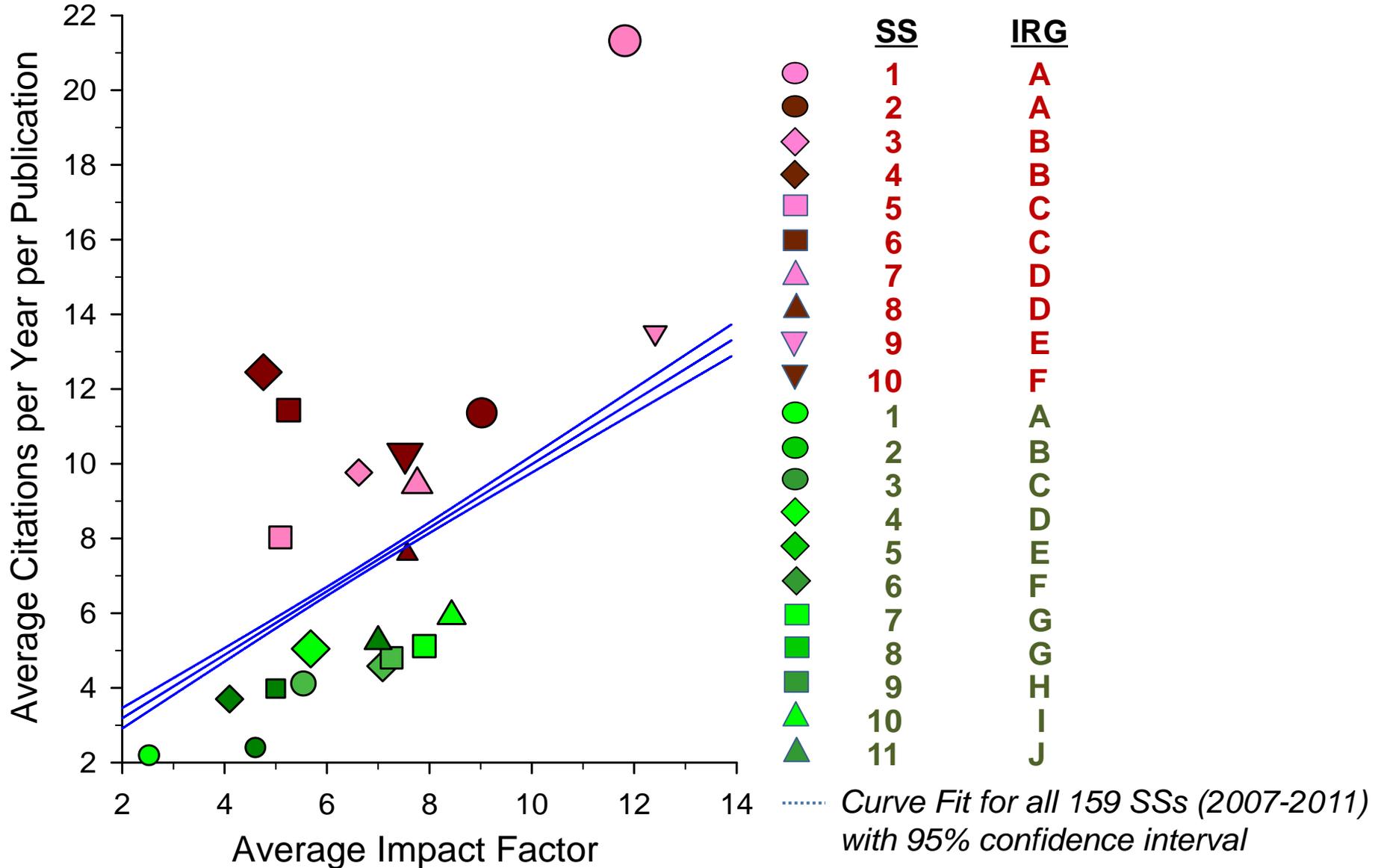
# “Hot” and “Cold” Study Sections 2009



# “Hot” and “Cold” Study Sections 2010



# “Hot” and “Cold” Study Sections 2011



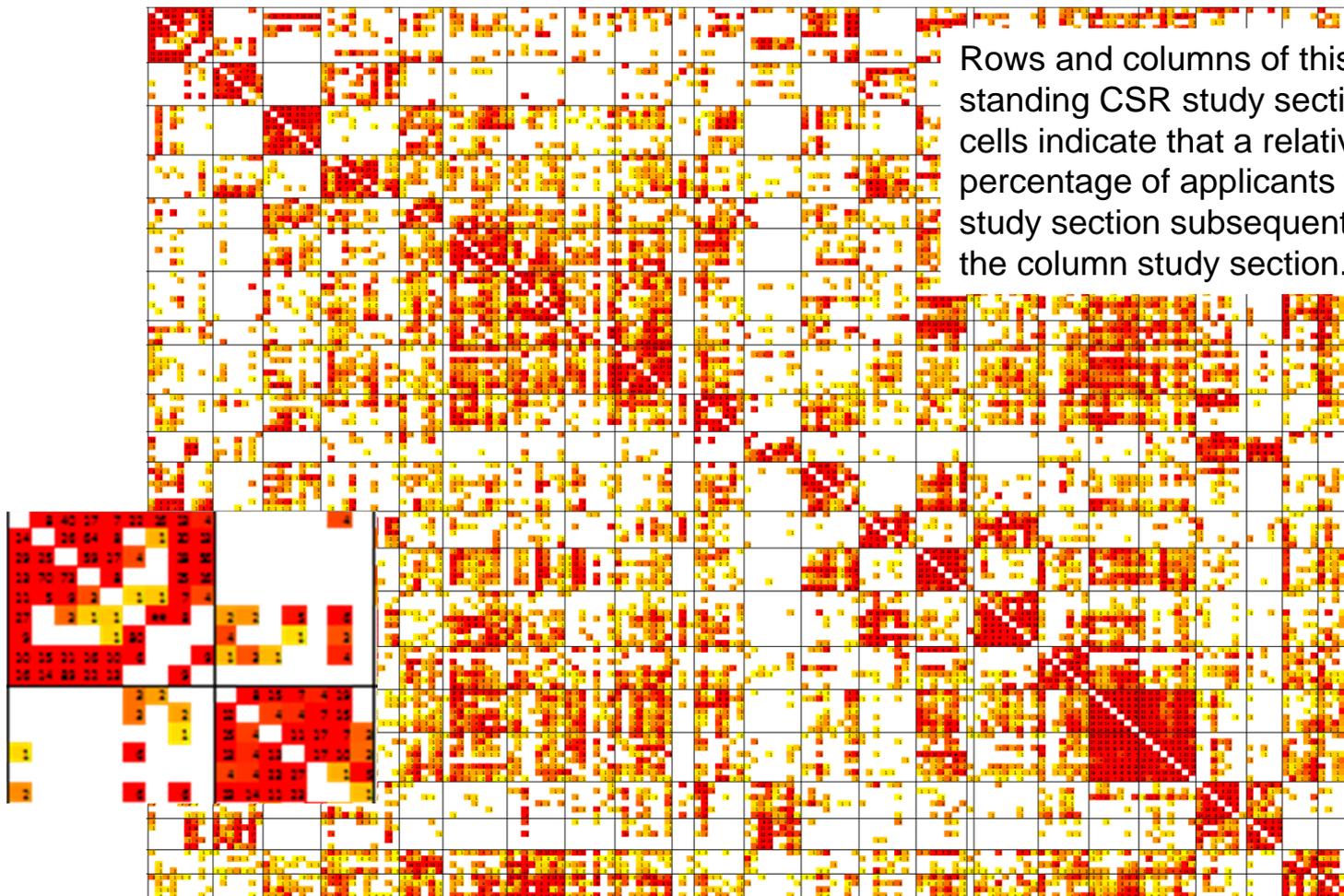


# Possible Quantitative Approaches Being Explored

- Analysis of study section “inputs”
- Tracking indicators of emergent fields
- Analysis of the study section “outputs”
- Analysis of study section “uniqueness”
  - Scientific similarity among applications reviewed by a study section
    - Fingerprints of applications
    - Reviewer citation patterns
    - Assignments and assignment requests
    - Applicant publications

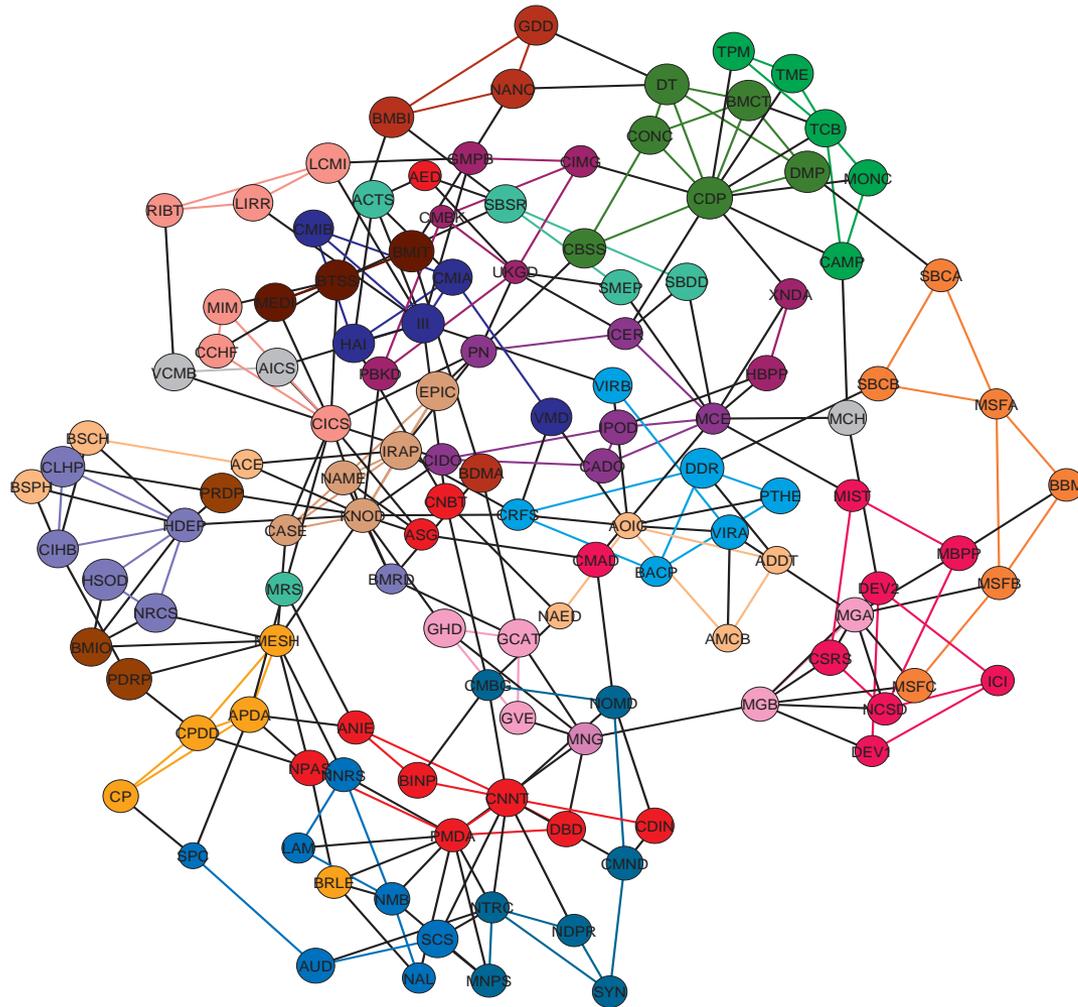
# Study Section Connectedness

## PIs Applying to Multiple Study Sections



Rows and columns of this matrix are standing CSR study sections. Red cells indicate that a relatively high percentage of applicants to the row study section subsequently apply to the column study section.

# A system Level Graph Representation of IRGs



- AARR
- BBBP
- BCMB
- BDCN
- BST
- CB
- CVRS
- DKUS
- EMNR
- ETTN
- GGG
- HDM
- IDM
- IFCN
- IMM
- MDCN
- MOSS
- OBT
- OTC
- PSE
- RPHB
- SBIB
- VH
- Other





# Limitations of quantitative approaches

- Analysis of review group outputs may be suggestive but clearly this will not provide sufficient insight into the nature and quality of the science we are supporting, nor will it be sufficient to identify emergent fields.
- Any quantitative analyses must be supplemented with expert qualitative input.

# Examples of Qualitative Analyses

- Aggregate the top 15% of all proposals within an IRG and subject this group of applications to a secondary ranking by a group of external experts.
  - Link secondary ranking with study section to ascertain if the distribution is random; if it is not consider “weighting” study sections by their rank.
- Conduct a series of NIH-wide IRG reviews to analyze the structure and compare highly scored applications across SRGs

# IRG Evaluation Schedule

- July 9, 2012: Biological Chemistry and Macromolecular Biophysics
- July 26, 2012 Infectious Diseases and Microbiology
- July (last week), 2012 Population Studies and Epidemiology

# Questions for Evaluation Panel

- Is the organization of the IRG consistent with its scientific domain?
- Do overlaps and gaps exist?
- Are emerging trends in science being accounted for?
- Are the rosters of the SRG's appropriate?
- Are the best applications in each SRG comparable in terms of expected quality of research?

# Examples of Qualitative Analyses (cont.)

- Conduct a series of NIH-wide portfolio reviews to compare the qualitative measures to quantitative assessments by experts
  - It would be easier to compare study section “performance” within a single field.
  - Any comparisons across fields would be complicated by a “value judgment” about the relative importance and/or alignment with NIH mission of one field versus the other.
  - However, regardless of which one field is selected, it will beg the question, why was that field was chosen?

# Thanks To:

- Jim Anderson (DPCPSI)
- George Chacko (CSR)
- Della Hann (OER)
- Kathy Hudson (OD)
- Richard Nakamura (CSR)
- Jim Onken (OER)
- Sally Rockey (OER)
- George Santangelo (OSB, DPCPSI)

# Discussion