

U.S. Department of Health & Human Services



Center for
Scientific Review

Fairness of CSR Peer Review and Review of Sex as a Biological Variable

Richard Nakamura, Director

ORWH Advisory Council

April 4, 2017

Vertebrate Animal Welfare

Important Considerations

Simplified in 2016: **NOT-OD-16-006**

- Concise description of procedures involving vertebrate animals
- Justification that the species are appropriate for the research
- Adequacy of veterinary care
- Interventions for minimization of pain and distress
- Is method of euthanasia consistent with recommendation of the American Veterinary Medical Association guidelines?

https://grants.nih.gov/grants/olaw/vertebrate_animal_section.htm

NIH . . . A Great Mission Shared Across Science



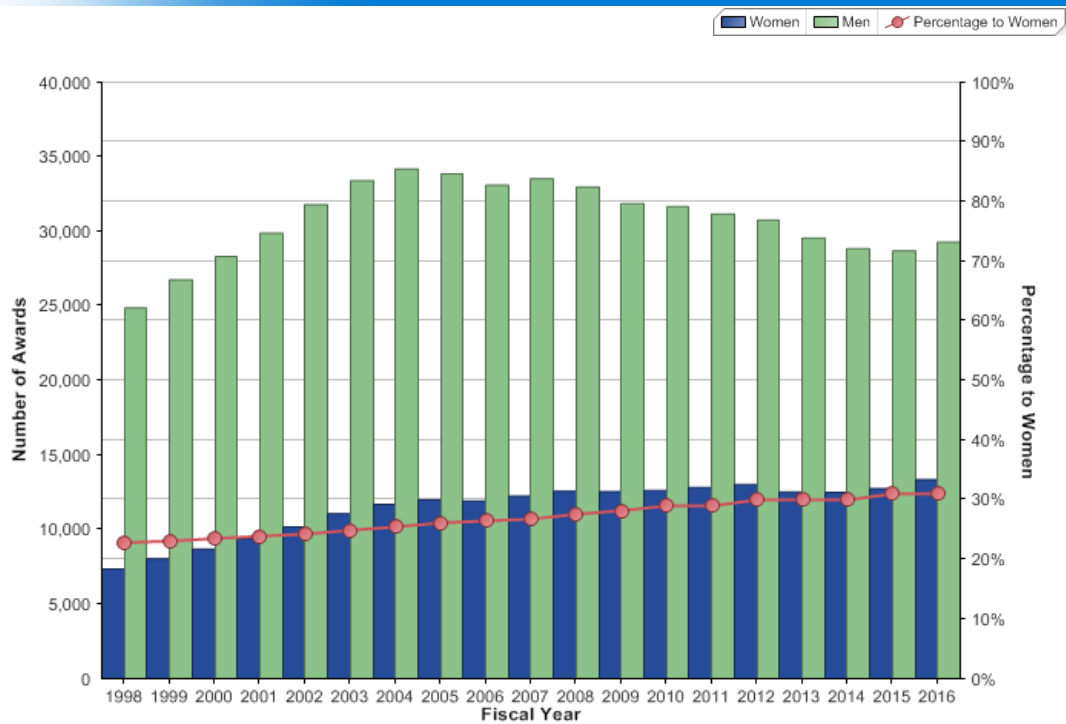
NIH's mission is to seek fundamental knowledge about the nature and behavior of living systems and the application of that knowledge to enhance health, lengthen life, and reduce illness and disability.

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

Gender Disparities?

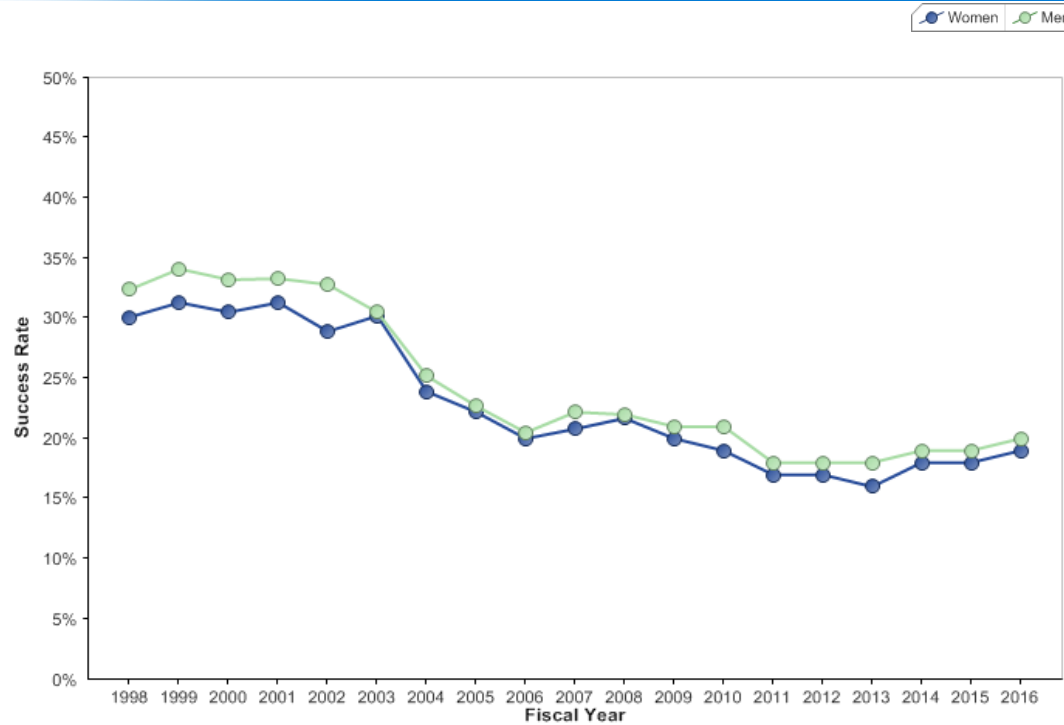


Research grants Awards, by gender





Research Project Grants Success rates, by gender

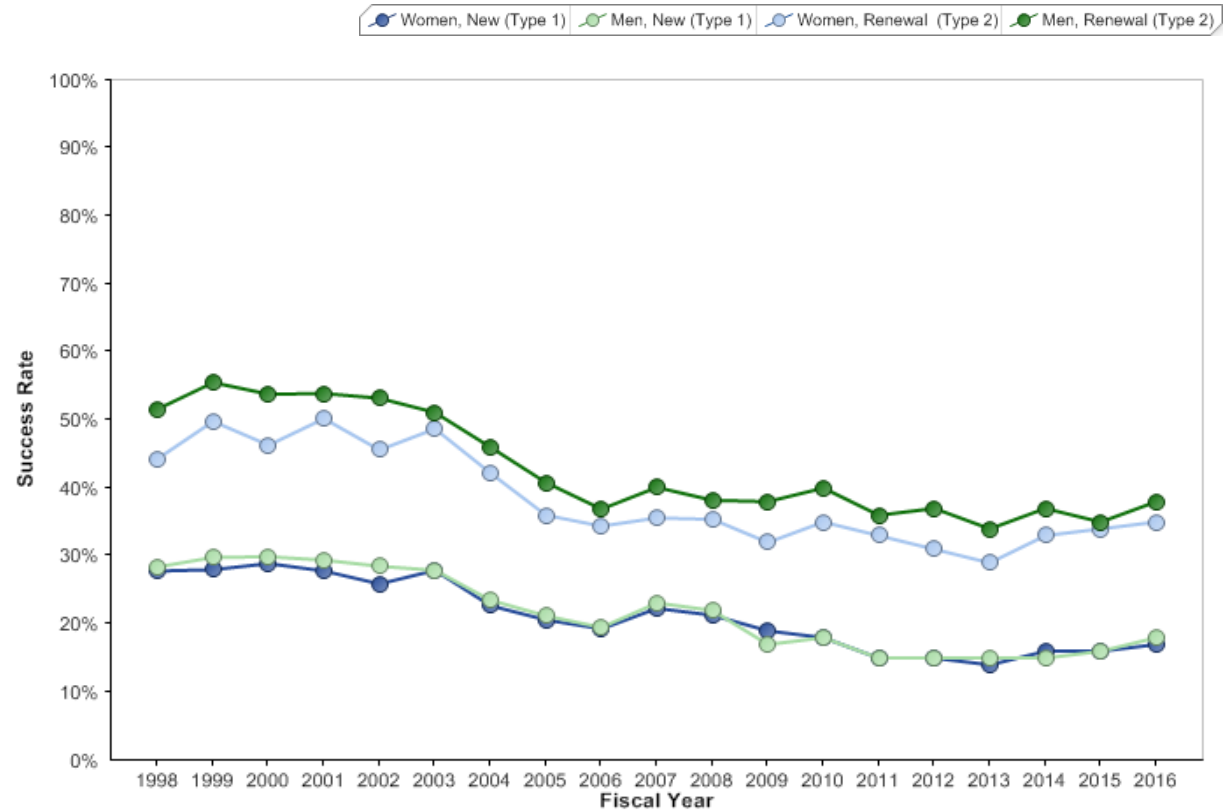


NIH National Institutes of Health
Office of Extramural Research

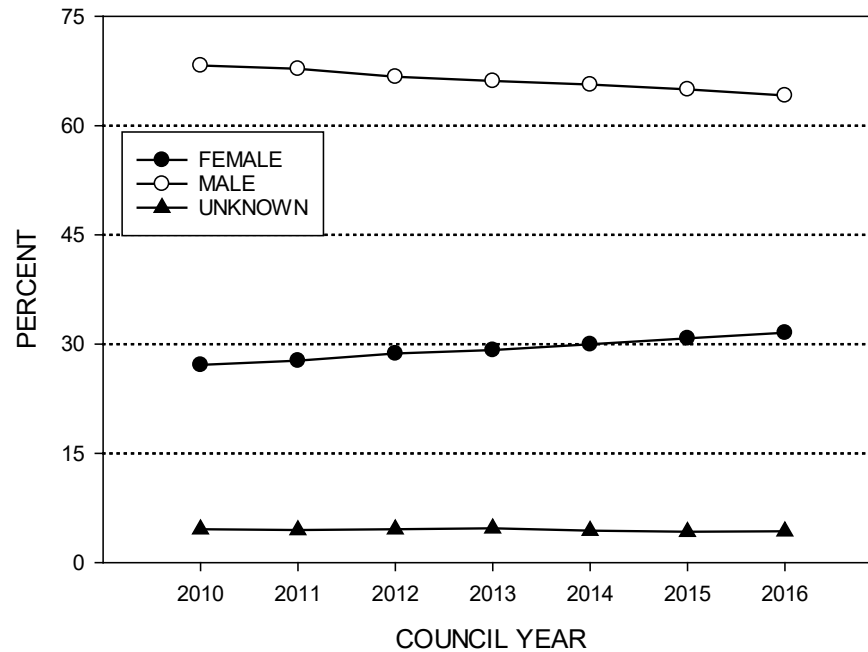
NIH Data Book – (<http://report.nih.gov/ndb/index.aspx>) Data provided by the Division of Statistical Analysis and Reporting Branch



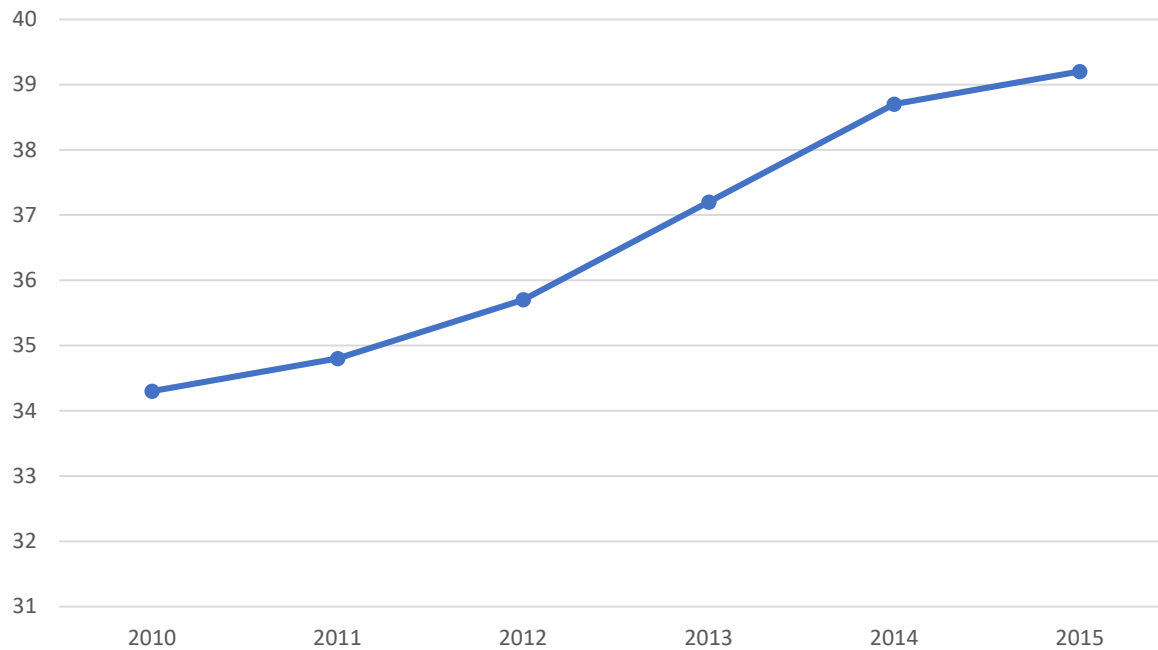
R01-Equivalent grants Success rates, by gender and type of application



Percent of All CSR Reviewers (Chartered and Ad Hoc) by Gender and Council Year



Percent of Female Chartered Members - All CSR



Comment on Implicit Bias

An Anonymizing Study



Anonymization Experiments – Basic Assumptions

- Racial disparities in grant funding exists (Ginther et al): AAs award rates much lower than Whites.
- At least 3 reviewers evaluate an application and their 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

ACD Rx's: CSR conduct studies using anonymization as a quality control check of our peer review process.

Specific Aims

To determine if masking personally identifiable information from grant applications changes the differences in final scores :

1. for Black and White applicants. (Primary aim)
2. for Male and Female applicants.
3. for Established and Early Stage Investigator applicants.
4. for applicants from more research intensive and less research intensive institutions.

Anonymizing Experiments – Design

1200 previously reviewed applications in 2014 – 2015 [400 AA, 400 Whites matched by science area, score, gender, degree, institution (NIH research \$) and seniority] & 400 Whites randomly selected

Application Formats

A. Original R01 Application	B. Full Anonymization	
With Investigator and Institution Information	No applicant or institution information provided using entire application (Information will be redacted)	Test for differences in scores between Original and Full Anonymization

Debriefing Aims

Debriefing Reviewer

- Can Reviewer identify Investigator(s), lab, institution?
- Can Reviewer identify Race, Gender or Seniority of PI?
- Can Reviewer rate the grantsmanship of the application?

Review of Rigor and Reproducibility

Growing Challenge: Ensuring the Rigor and Reproducibility of Science

Vox SCIENCE & HEALTH

Too many mice are sacrificed for seriously flawed studies

Updated by Julia Belluz on July 28, 2016, 10:20 a.m. ET @juliaoftoronto julia.belluz@

community and beyond in several

PERSPECTIVE

JBMR®

Reproducibility of Results in Preclinical Studies: A Perspective From the Bone Field

Stavros C Manolagas¹ and Henry M Kronenberg²

RESEARCH ARTICLE

PSYCHOLOGY

Estimating the reproducibility of psychological science

Open Science Collaboration^{††}

The Economist

World politics Business & finance Economics Science & technology Culture

Unreliable research

Trouble at the lab

Scientists like to think of science as self-correcting. To an alarming degree, it is not

Like 11k Tweet 1,227

Science

Evaluating replicability of laboratory experiments in economics

Colin F. Camerer,^{1,††} Anna Dreber,^{2†} Eskil Forsell,^{2†} Teck-Hua Ho,^{3,††} Jürgen Huber,^{4†} Magnus Johannesson,^{2†} Michael Kirchler,^{5,6†} Johan Almenberg,⁷ Adam Altmejd,² Taizan Chan,⁸ Emma Heikensten,² Felix Holzmeister,² Taisuke Imai,¹ Siri Isaksson,² Gideon Nave,¹ Thomas Pfeiffer,^{9,10} Michael Razen,² Hang Wu[†]

THE WALL STREET JOURNAL.

OPINION | COMMENTARY

Getting the Bogus Studies Out of Science

Government funding should provide more incentives for replicating research

By ADAM MARCUS And IVAN ORANSKY

Why animal research needs to improve

Many of the studies that use animals to model human diseases are too small and too prone to bias to be trusted, says Malcolm Macleod.

Believe it or not: how much can we rely on published data on potential drug targets?

Florian Prinz, Thomas Schlange and Khusru Asadullah

TECHNICAL COMMENT

PSYCHOLOGY

Comment on "Estimating the reproducibility of psychological science"

Daniel T. Gilbert,^{1,††} Gary King,¹ Stephen Pettigrew,¹ Timothy D. Wilson²

False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant

Raise standards for preclinical cancer research

C. Glenn Begley and Lee M. Ellis propose how methods, publications and incentives must change if patients are to benefit.

Reforming Science: Methodological and Cultural Reforms

“Mission Control, we have a problem.”

Rigor and Reproducibility in Research

To support the **highest quality science, public accountability, and social responsibility in the conduct of science**, NIH's Rigor and Reproducibility efforts are intended to clarify expectations and highlight attention to four areas that may need more explicit attention by applicants and reviewers:

- Scientific premise
- Scientific rigor
- Consideration of relevant biological variables, such as sex
- Authentication of key biological and/or chemical resources

Role of reviewers: Assess the scientific merit of each application according to the review criteria, which include consideration of scientific premise, rigor, and consideration of relevant biological variables, and the adequacy of the authentication of key biological and/or chemical resources as an administrative issue. Evaluations should be based on current best practices in the field.

What are some of the problems?

Insufficient Reporting of Methodological Approaches is Evident for Pre-Clinical Studies

Table 3. Prevalence of selected quality characteristics in other experimental models

	Number of publications	Randomisation (%)	Blinded assessment of outcome (%)	Sample-size calculation (%)
Transgenic stroke studies	157	n/a	3	0
Stroke pathophysiology studies	166	5	18	0
Parkinson's disease	118	12	15	0
Multiple sclerosis	183	2	11	0

Trends Neurosci 2007; 30: 433-439

Small, Underpowered Studies

- Small, underpowered studies suffer from:
 - More false-negatives
 - More false-positives
 - Reduced positive predictive value – the probability that a positive research finding reflects a true effect

Selective Reporting

- Of data, subjects and experiments

Raise Community Awareness

COMMENT

612 | NATURE | VOL 505 | 30 JANUARY 2014

NIH plans to enhance reproducibility

Francis S. Collins and Lawrence A. Tabak discuss initiatives that the US National Institutes of Health is exploring to restore the self-correcting nature of preclinical research.

“Efforts by the NIH alone will not be sufficient to effect real change in this unhealthy environment.”

A growing chorus of concern, from scientists and laypeople, contends that the complex system for ensuring the reproducibility of biomedical research is failing and is in need of restructuring^{1, 2}. As leaders of the US National Institutes of Health (NIH), we share this concern and here explore some of the significant interventions that we are planning.

Science has long been regarded as ‘self-correcting’, given that it is founded on the replication of prior work. Over the long term, that principle remains true. In the shorter term, it balances that one who have been hobbled by today’s findings.

Let’s be clear: there is no evidence that reproducibility is also. In 2011, the Office of the US Department of Health Services pursued. Even if this report the actual problem

NIH National Institutes of Health
Turning Discovery Into Health

Health Information | Grants & Funding | News & Events | Research & Training | Institutes at NIH | About NIH

RIGOR AND REPRODUCIBILITY

Rigor and Reproducibility
Principles and Guidelines
Publications
Training
Meetings and Workshops
Expanded Guidelines
Application Instructions

Rigor and Reproducibility
Two of the cornerstones of science advancement are rigor in designing and performing scientific research and the ability to reproduce biomedical research findings. The application of rigor ensures robust and unbiased experimental design, methodology, analysis, interpretation, and reporting of results. When a result can be reproduced by multiple scientists, it validates the original results and readiness to progress to the next phase of research. This is especially important for clinical trials in humans, which are built on studies that have demonstrated a particular effect or outcome.

Email Updates
To sign up for updates please enter your e-mail address.

Related Links
Letter from Dr. Stephen J. Katz: An Update on the NIH Initiative to Enhance Research Rigor and Reproducibility
Nature commentary on the Promise and Peril of Chemical Probes

Contact Us
Please send email to help@odf@nih.gov.

PERSPECTIVES



CELL BIOLOGY
Fixing problems with cell lines
Technologies and policies can improve authentication

By Jan S. Lambeth¹, Frederick C. Collins²,
Amalinder Upadhyay-Schwartz³

COMMENT



NIH to balance sex in cell and animal studies

Janine A. Clayton and Francis S. Collins unveil policies to ensure that preclinical research funded by the US National Institutes of Health considers females and males.

- Workshops with PhRMA and Journal Editors
- **Over 135 journals** endorsed the principles

<http://www.nih.gov/science/reproducibility/principles-guidelines.htm>

Science

JCB

jbc THE JOURNAL OF
BIOLOGICAL CHEMISTRY

nature

Role for Individual Scientists

What you can do:

- Stimulate discussion among societies/organizations
- Increase transparency
- Promote training in experimental design
- Encourage data and material sharing
- Publish refutations and negative results

Role of Peer Review



Reviewing Rigor and Transparency of Research: RPG Applications

	Applies to which applications?	Where will I find it in the application?	Where do I include it in my critique?	Addition to review criteria	Affect overall impact score?
Scientific Premise	All	Research Strategy (Significance)	Significance	Is there a strong scientific premise for the project?	Yes
Scientific Rigor	All	Research Strategy (Approach)	Approach	Are there strategies to ensure a robust and unbiased approach?	Yes
Consideration of Relevant Biological Variables, Such as Sex	Projects with vertebrate animals and/or human subjects	Research Strategy (Approach)	Approach	Are adequate plans to address relevant biological variables, such as sex, included for studies in vertebrate animals or human subjects?	Yes
Authentication of Key Biological and/or Chemical Resources	Project involving key biological and/or chemical resources	New Attachment	Additional review considerations	Comment on plans for identifying and ensuring validity of resources.	No



Scientific Premise: Guidance for Reviewers

GOAL: Ensure that the underlying **scientific foundation** of the project—concepts, previous work, and data (when relevant)—is sound.

- Pertains to the strength of the scientific foundation of **evidence/data** that increase possibility of high impact for the project

Premise should not be confused with hypothesis or significance

Premise

Specifically, have you:

- Provided sufficient justification for the proposed work?
- Cited appropriate work and/or preliminary data?
- Appropriately identified strengths and weaknesses in prior work in the field?
- Proposed to fill a significant gap in the field?

Scientific Rigor: Guidance for Reviewers

GOAL: Ensure a strict application of scientific method that supports robust and unbiased design, analysis, interpretation, and reporting of results, and sufficient information for the study to be assessed and reproduced. Give careful consideration to the methods and issues that matter in your field.

Scientific Rigor:

Possible considerations, if appropriate for the scientific field and research question, include plans for:

- determining group sizes
- analyzing results
- reducing bias
- ensuring independent and blinded measurements
- improving precision and reducing variability
- inclusion or exclusion of research subjects



Sex as a Biological Variable:

Guidance for Reviewers

NIH expectations for **reviewers**:

- As part of the Consideration of Relevant Biological Variables, assess whether the plans to address sex as a biological variable are adequate (for studies in vertebrate animals or human subjects).
- If the study involves only one sex, is this justified scientifically?
- Assess within the context of the research question and current scientific knowledge.



Plan for Resource Authentication: Guidance for Reviewers

GOAL: Ensure processes are in place to identify and regularly validate key resources used in their research and avoid unreliable research as a result of misidentified or contaminated resources.

- Researchers are expected to authenticate key biological and/or chemical resources used in their research, to ensure that the resources are genuine.
- New Review Consideration
- Reviewers rate as acceptable/unacceptable (provide brief explanation if unacceptable)
- Does not affect criterion scores or overall impact score

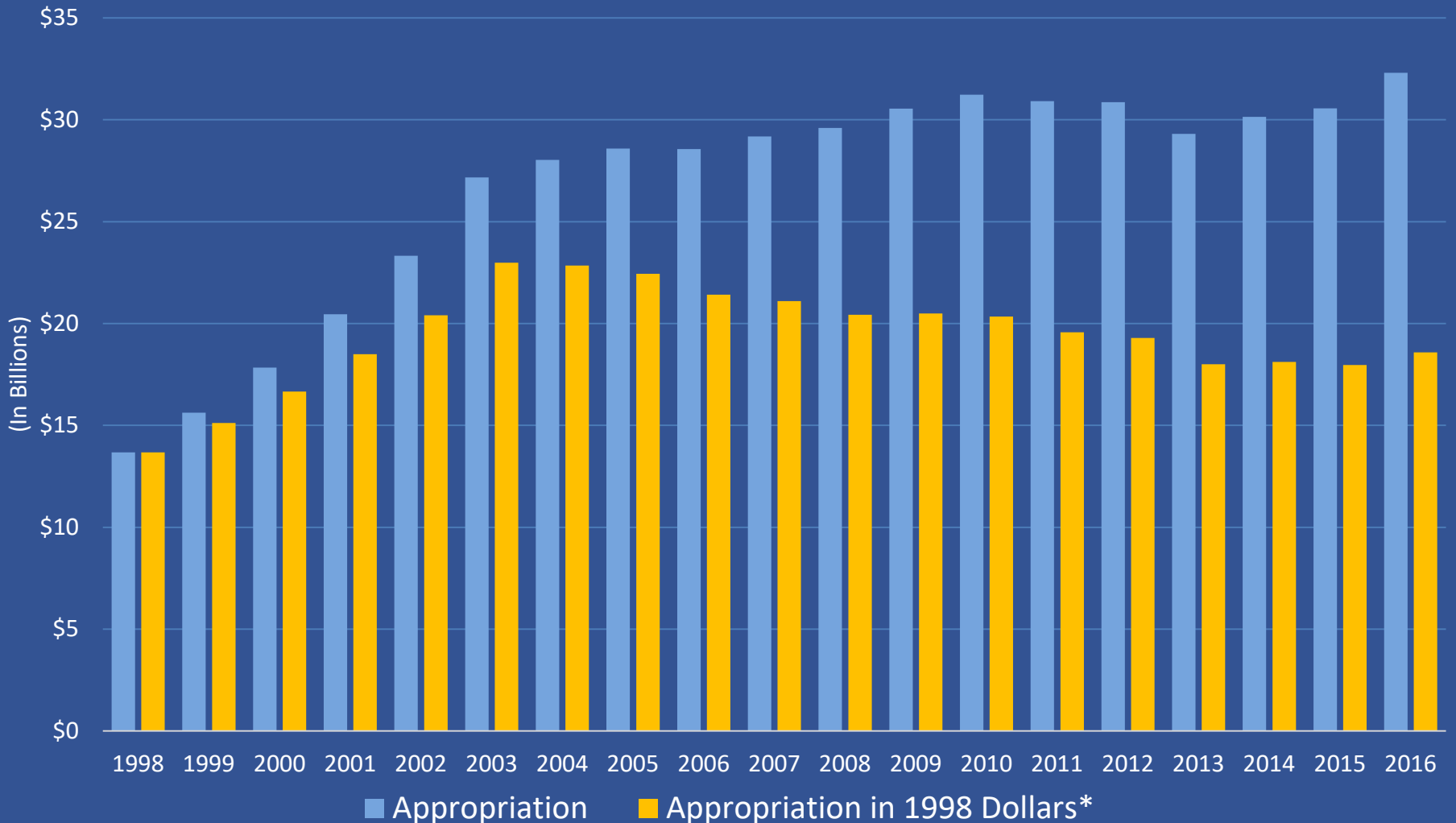


Related review issues:

- Different research fields may have different best practices for and reach different conclusions about scientific premise and rigor. Use the words. Assess based on best practices in the field.
- Page limits have not changed
- Cost of larger subject populations
- Good science can emerge from different styles
- More background investigation of premise
- The reviewer is the judge of premise and rigor
- Exploratory studies are still allowed
- Significance and potential for impact are still important

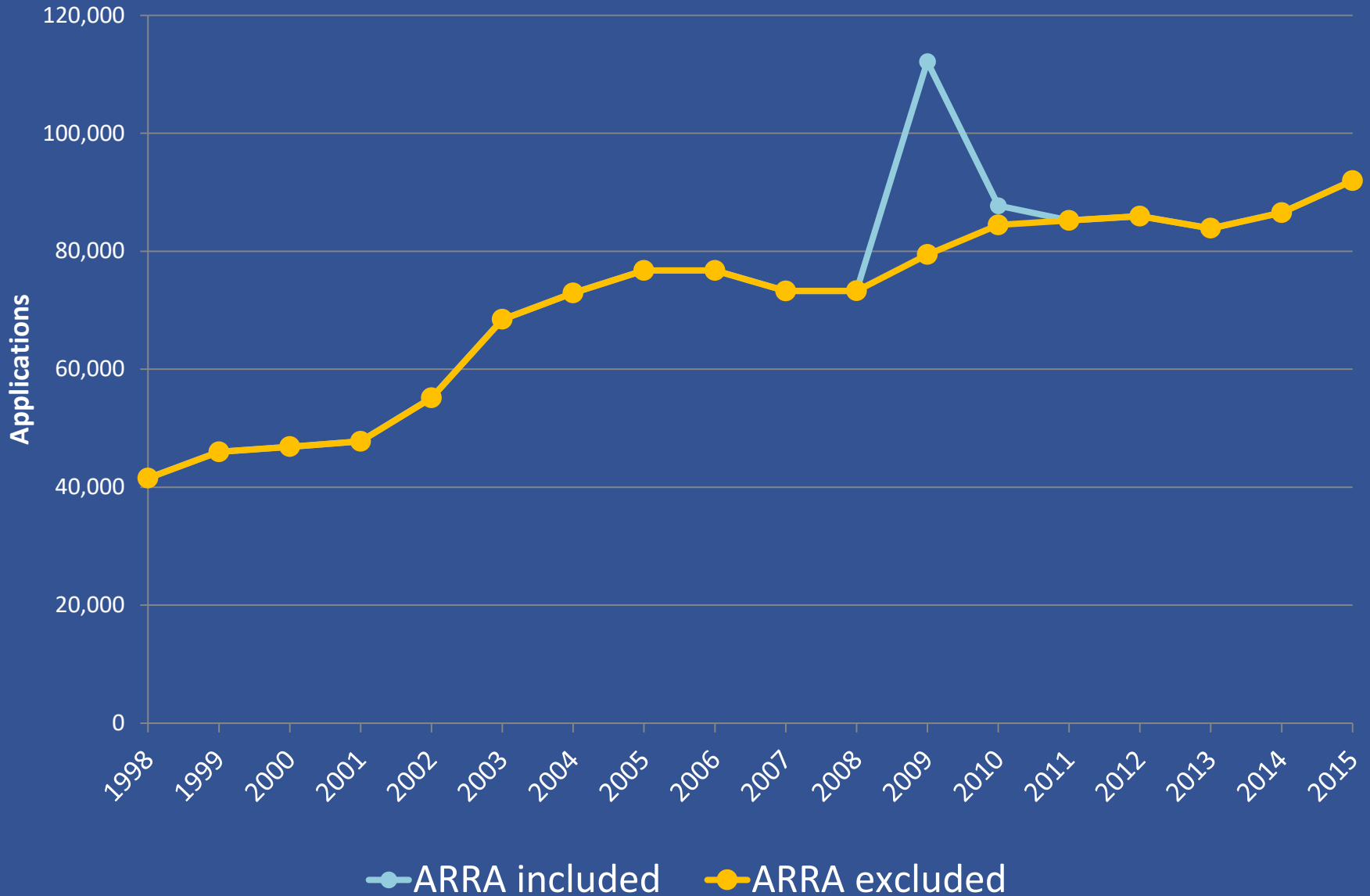
The Future of Science and Technology

NIH Appropriation in Nominal Dollars and Constant 1998 Dollars FY1998 – FY2016



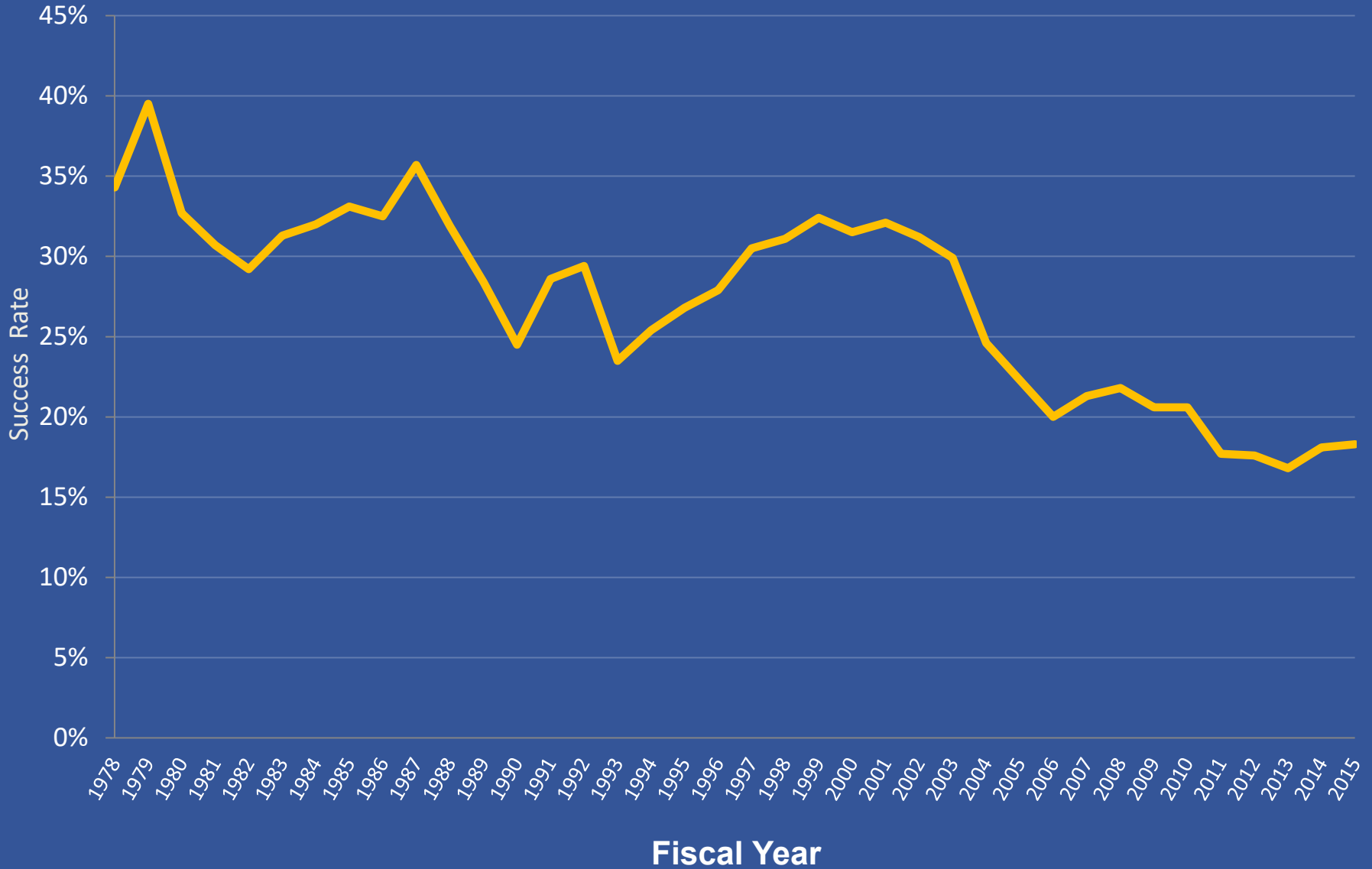
*FY 2016 BRDPI estimate based on FY 2015 actual

Number of Applications Received by Fiscal Year



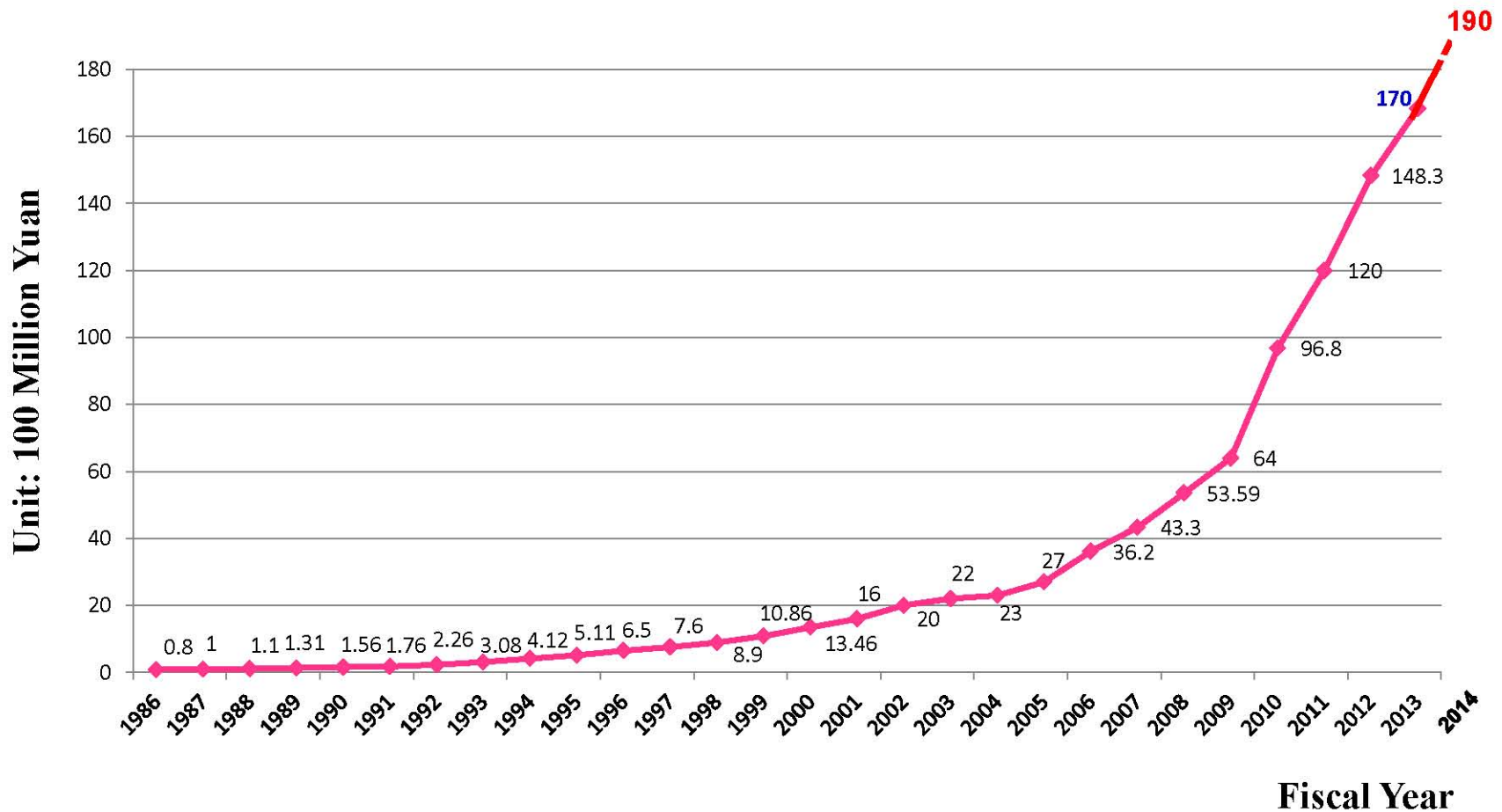
Grant Success Rates

FY 1978-2015

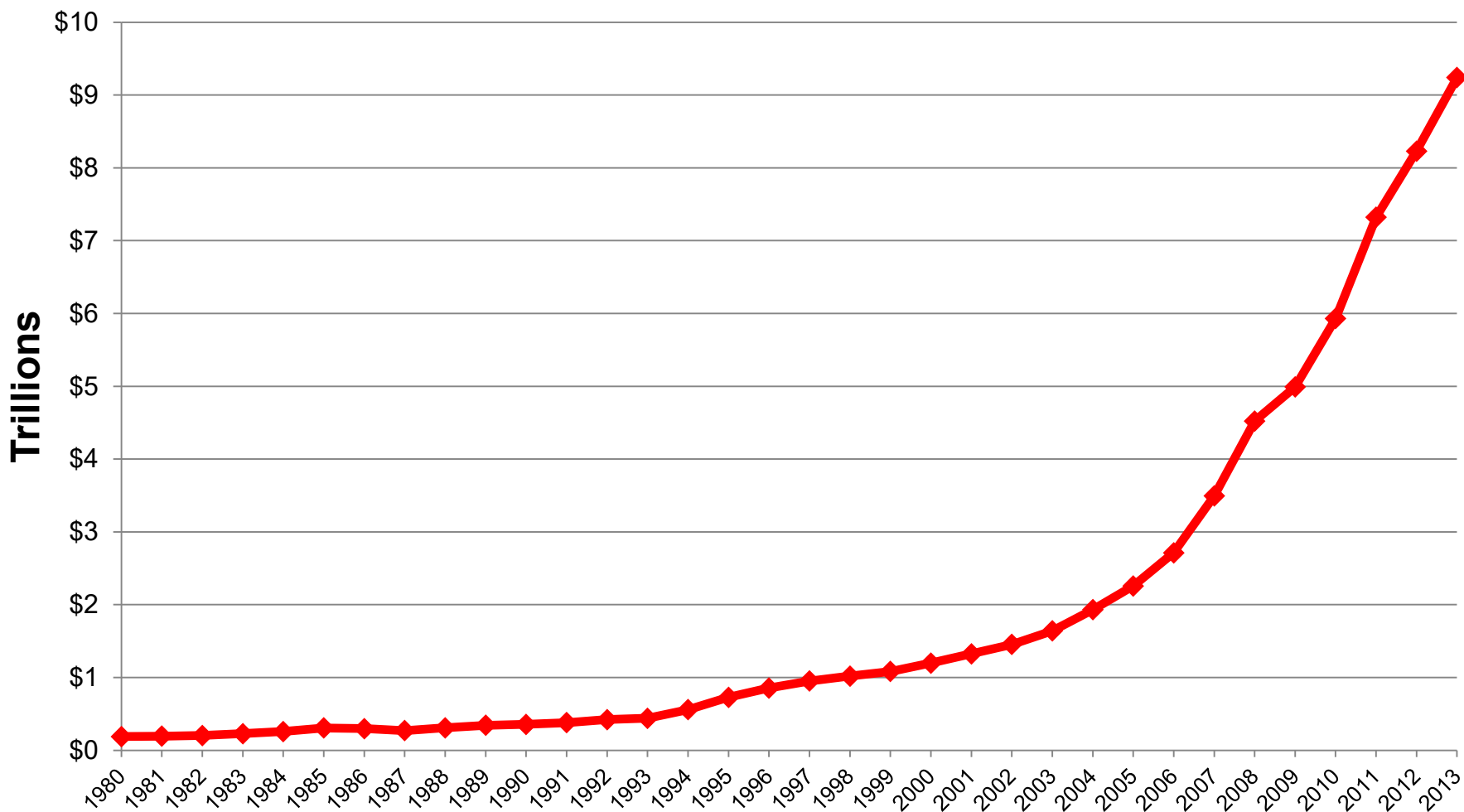


Budget for 1986-2014

The total budget for 2014 is ¥19Billion (~\$ 3.05Billion), an increase by 11.7% over the year 2013.



China GDP in Current US Dollars, 1980-2013



<http://data.worldbank.org/country/china>

Discussion