EVALUATION OF SCIENCE, RESOURCE ALLOCATION, AND SCIENTIFIC PERFORMANCE


Bogus Academic Business

• Increasing online “open access” journals
  – Objective: to allow everyone to read scientific papers
    (Traditional journals cannot be read without subscription.)
  – Reality: to charge scientists who are desperate enough to pay $100-2000, and accept papers almost unconditionally. Cost of publication is negligible since it’s online (no print).
Experiment

• Choose 300 suspicious journals

• Submit a fake paper to each journal
  – *Fake*: scientifically nonsensical, therefore, should be rejected if peer review occurs

• See what happens
  – Accepted or rejected?
  – How peer review works?
    • No review, serious, or superficial (only about format)

**Result**

- Accepted: 52%
- Superficial review: 38%
- Serious review: 10%
- No peer review (=Editor accepted): 52%
- Not yet result known: 16%
- Rejected: 32%
How could we really facilitate scientific productivity?

- How should scientists be evaluated?
- How should resources be allocated among scientists?
Pub Performance Measure

- The simplest one: count pub
  - This can be contaminated with bogus pubs.

- How do you know that a paper is excellent (if not reading the paper)?
  - You don’t have time to read all papers for evaluation.
  - You want to know if a paper is worth reading in the first place.

- Ask experts!!
  - Serious evaluation by experts
    - Scholarly awards (e.g., Nobel prize) (this infrequently happens)
  - Citation
    - Being cited means impact.
  - Journal reputation
    - A paper must be great if a prestigious journal accepted it.
      E.g., You may trust Nikkei or Financial Times but doubt tabloids (e.g., Sankei Sports, National Enquirer).
Journal Ranking

• Expert survey
  – Scholars in a certain field give scores to journals in the field (Serenko & Dohan 2011).

• Based on bibliometric information
  – SCImago Journal Rank
  – Eigenfactor
  – **Impact Factor**
    • Average #citation of a journal
      “The number of current citations to articles published in a specific journal in a two year period divided by the total number of articles published in the same journal in the corresponding two year period.” (Thomson Reuter)
    • Controversial but heavily used (Denrell 2012)
Impact Factor (IF)
JCR2012: Biochemistry & Molecular Biology, etc.

290 Journals
Median (IF) = 2.857

Density

0.0
0.1
0.15
0.2

0 10 20 30 40
IF

PLoS One
3.7

Science
31.0

Nature
38.6
Reaction of Scientists (1): Worship to Prestigious Journals

- Publishing in prestigious journals become a goal for scientists.

“The fact THAT a scientist publishes, HOW MUCH he publishes, and especially WHERE he publishes matters more than WHAT he is publishing.” (Holub 1991)

Reaction of Scientists (2):
Do research preferred by prestigious journals

• Novel study >> Replication study (Dewald 1986)
  – Replication studies are indispensable to verify reproducibility (Burman, Reed & Alm 2010).

• Research on Human >> Animal, Insect, Plant
  “Studies on humans are more valued than those on plants.” (Agri Dept, Yokohama City U)

• Positive result >> Negative result (Csada 1996, etc.)
  – i.e., Evidence of failing to support hypotheses
Reaction of Scientists (3): Publish as many as possible

(Aside from misconduct: e.g., duplicate pub)

• Do easy research
  – Incremental research rather than innovative one
  – Replication
    • “A previous study proved XXXX using mice, but no study is done with rats. Therefore, let’s prove XXXX using rats!!”

• Minimize effort per publication unit
  – Papers rather than books
  – Short papers rather than long papers
  – Fragmentation, or salami slicing (Broad 1981)
Many papers never cited

• %0-cited pub in economics constant over decades (Laband 2003)

• Cross-field comparison (Hamilton 1991; Wallace et al. 2009)

Many papers never read

Q: How often do you read the journal that you submitted your paper to?
IMPACT OF EVALUATION ON PRODUCTIVITY: CASE STUDY
Evaluation: Framework

• Objective
  – Funding allocation
  – Employment/promotion
  – Economic reward (salary, bonus)
  – Scholarly award

• Criteria
  – Quality vs. Quantity
  – Weight of research compared to other missions (teaching, etc.)

• Unit of evaluation
  – Individual, department, university, ...

• Methods
  – Indicator
  – Peer review
  – Informal peer reputation
Case 2: Funding in Australia

Reform of institutional funding system in the late 1980s

- **Objective**
  - Level 1: Allocate national research grants to universities
  - Level 2: Allocate the grants to researchers within a university

- **Unit of evaluation**
  - Level 1: University
  - Level 2: Individual scientist

- **Criteria**
  - Level 1/2: Publication output (#pub, oftentimes)

- **Method**
  - Level 1: indicator
  - Level 2: indicator + peer review

- **Result**
  - “Publication productivity has increased significantly in the last decade, but its impact has declined.” (Butler, 2003)

---

**Fig. 1. Australia’s share of publications and citations, and RCI in the SCI, 1981–1998.**

- World share in #Pub
- World share in #Cite
- #Cite/#Pub (in comparison to world average)

[UQ] Recruit brightest young researchers, and provide them with a strong resource base.

[UWA] introduce formula for distributing research funds, where #pub plays a major role.
Case 3: Incentive for high-impact publication


<table>
<thead>
<tr>
<th></th>
<th>#Submission to <em>Science</em></th>
<th>#Acceptance (#Publication)</th>
<th>%Acceptance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Individual cash bonus</strong></td>
<td>46%↑**</td>
<td>2%↓</td>
<td>39%↓*</td>
</tr>
<tr>
<td><strong>Institutional incentive</strong></td>
<td>24%↑**</td>
<td>21%↑</td>
<td>8%↓</td>
</tr>
<tr>
<td><strong>Individual career incentive</strong></td>
<td>12%↑**</td>
<td>34%↑**</td>
<td>10%↑</td>
</tr>
</tbody>
</table>

Note: **p<.01, *p<.05
Case 4: Formal/ininformal incentive in US

• Prior study: emphasis to *quality* (Hamermesh, 2012)
  – Impact on salary
    • #Cite↑ (67% increase by 5 ->95 percentile)
  – Impact on peer reputation (prize awarded, fellow of society, etc.)
    • #Cite↑
    • #Pub↓

• Study setting (Fanelli 2010)
  – Question: How pressure for quality affects pub behavior?
    • Does competition discourage negative reports?
  – All fields in US; comparison between states
Case 5: Japanese evaluation system/culture

• Quality metrics is highly emphasized these days.

   “The HQ emphasizes pubs in prestigious journals based on J ranking (i.e., Nature/Science). I believe that they should employ more brains, who can evaluate research based on contents.” (U Tokyo)

• Yet, quantity matters.
  – Condition for promotion
  – Condition for PhD graduation

• Once tenured, very weak incentive for pub
  – Seniority-based salary and life-time employment
Employment criteria (Life Sciences)

Q1: Any \#pub criteria in hiring?

Q2: If yes, how many pub needed to be FP?

Note: Survey of life sciences FP/AP in 2013
Q: How many pubs do PhD students have to publish for graduation?

Note: Survey of life sciences FP/AP in 2013
Q: Is the income of faculty members affected by pub performance? If so, to what extent?

Note: Survey of life sciences FP/AP in 2013
Portfolio of Impact Factor

Note: Biochemistry & Molecular Biology from Web of Science in 2010

Legend:
- Very low IF (<1)
- Low IF (<2)
- Middle IF (2-8)
- High IF (>8)
Transmission of Norms

I would rather scrap a paper than to publish in low-impact journals. *Junk papers* are not only unappreciated but also damage my reputation.” (U Tokyo, returned from US)

• Study setting
  – 400 Life science professors in Japan
  – Comparison of pub practices between
    • those who have stayed abroad (mostly in US) and
    • those who never stayed abroad
Q: Scientists should avoid publishing in journals that are rarely read, because it is a waste of resources. (5: Agree ↔ 1: Disagree)

** p < .01
Difference-in-Difference Analysis

Returnee vs. domestic

Returnee (Treatment)

Foreign stay (1 year +)

Domestic (Control)

5 years

PhD Completion

Now

5 years

5 years

Now

5 years

Copyright (C) 2015 Sotaro Shibayama

11/16/2015
Effect of Foreign Stay

Finding 1:
%Pub in Low-impact journals decreased.

Finding 2:
%Low-Impact pub (infrequently cited) did not change.
Effect of Foreign Stay

Finding 3:
%Pub in Low-impact journals decreased because #pub in low-IF journal decreased but not because #pub in high-IF journal increased.

Impact of Journal

<table>
<thead>
<tr>
<th>Not Low-IF (IF&gt;2)</th>
<th>Low-IF (IF&lt;2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td>Blue</td>
</tr>
</tbody>
</table>
Summary

• Scientific production is incentivized by various means
  – not only by formal evaluation system (employment, funding, salary...),
  – but also by implicit culture.

• These differ by field, by institution, by country, and by time.

• Scientists seem to react to them quite sensitively, not necessarily as intended.
RESOURCE ALLOCATION
US result (NIH)

222 NIH grants: 22 researchers

A whopping 200 scientists received six or more grants each from the US National Institutes of Health (NIH) in 2007, according to data analysed by Nature. One principal investigator was awarded 32 grants, the data reveal, and many others got eight or nine.

The amounts awarded to some of these grandee grantees held some surprises too. Robert Sherwin of Yale University in Connecticut received eight grants totalling US$14.5 million last year for his research into diabetes; Harold Varmus, president of the Memorial Sloan-Kettering Cancer Center, New York, received grants of $13 million for work on cancer; and cell-death researcher John Reed of the California-based Burnham Institute for Medical Research received nearly $11 million in 11 grants.

The data that Nature analysed include all

<table>
<thead>
<tr>
<th>Principal investigator</th>
<th>Institution</th>
<th>Research type</th>
<th>of grants</th>
<th>(US$1,000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andrew Robertson</td>
<td>Keystone Symposia</td>
<td>Conference organizer</td>
<td>32</td>
<td>4,900</td>
</tr>
<tr>
<td>Terri Grodzicker</td>
<td>Cold Spring Harbor Laboratory</td>
<td>Conference organizer</td>
<td>16</td>
<td>8,690</td>
</tr>
<tr>
<td>Stan Vermund</td>
<td>Family Health International</td>
<td>HIV-prevention trials</td>
<td>11</td>
<td>24,132</td>
</tr>
<tr>
<td>John Reed</td>
<td>Burnham Institute for Medical Research</td>
<td>Cell death</td>
<td>11</td>
<td>10,868</td>
</tr>
<tr>
<td>Jeffrey Murray</td>
<td>University of Iowa</td>
<td>Birth defects</td>
<td>11</td>
<td>7,060</td>
</tr>
<tr>
<td>Joseph McCune</td>
<td>University of California, San Francisco</td>
<td>Translational science</td>
<td>9</td>
<td>25,396</td>
</tr>
<tr>
<td>Bert O’Malley</td>
<td>Baylor College of Medicine</td>
<td>Reproductive biology</td>
<td>10</td>
<td>17,149</td>
</tr>
<tr>
<td>David Rawlings</td>
<td>Children’s Hospital, Seattle</td>
<td>Gene therapy</td>
<td>15</td>
<td>22,590</td>
</tr>
<tr>
<td>David Allison</td>
<td>University of Alabama at Birmingham</td>
<td>Obesity</td>
<td>8</td>
<td>55,529</td>
</tr>
<tr>
<td>David Stewart</td>
<td>Cold Spring Harbor Laboratory</td>
<td>Conference organizer</td>
<td>9</td>
<td>22,590</td>
</tr>
<tr>
<td>Robert Sherwin</td>
<td>Yale University</td>
<td>Diabetes</td>
<td>8</td>
<td>14,550</td>
</tr>
<tr>
<td>Harold Varmus</td>
<td>Memorial Sloan-Kettering Cancer Center</td>
<td>Cancer</td>
<td>8</td>
<td>13,119</td>
</tr>
<tr>
<td>Pamela Davis</td>
<td>Case Western Reserve University</td>
<td>Cystic fibrosis</td>
<td>8</td>
<td>12,518</td>
</tr>
</tbody>
</table>

Budget Structure of Japanese Universities

**Budget source of national universities**

- Competitive grant: 17%
- Own income (tuition, etc.): 44%
- Block grant: 39%

**Source of research budget for national university faculty**

- Government: 65%
- Affiliated Univ: 19%
- Industry: 7%
- Foundation: 7%
- Other source: 2%

Funding Allocation in Japan
Grants-in-Aid (科研費)

Top 1% professors received 21% of total budget.

Top 10% professors received 60% of total budget.

Note: Full/Associate professors in all fields based on KAKEN 2001-2005.
Inequality measure

• Gini coefficient
  \[ = \frac{A}{(A+B)} \]
  – 0.685 (all U)
  – 0.746 (Top 7 U)
Inequality of Income Distribution

Source: World Development Indicators: Distribution of income or consumption The World Bank (2014)
Grant Size and Count

Grant size getting larger

Multiple funding increased with some recent remedy

Note: price adjusted to the year 2000 level.
Trend of Inequality

Note: Full/Associate professors in all fields based on KAKEN 2001-2005.
Problem of “Concentration & Focus”

• A single PI cannot do meaningful contribution to too many projects.
• Diversity of research is compromised.
• Decreasing returns: 1 pub/$1,000 does not guarantee 10 pub/$10,000.
• Evaluation depends on past productivity.
  → Initial winners continue winning.
  → Dynamics is compromised.
• In evaluation of productivity, output could be measured, but input is often difficult to know systematically. (various funding sources, etc.)
  → Initially marginal difference is augmented exponentially.
Funding distribution is more skewed than pub distribution.

Note: For full and associate professors in all universities, the distributions of funding and publications in 2001–2005 are illustrated. I counted the number of publications resulting from the grants, and summed them up for each PI. When more than one PI is assigned to one grant, the number of publications is divided by the number of PIs.
Budget and Pub Impact

Too big a budget leads to a greater number of mediocre pubs.

Too big a budget leads to a smaller number of high-impact pubs.
Top 10% universities produce 60% of all PhDs. Top 7 produce 38%.

The skewness has been increasing since the 1990s.
Problem of “Concentration & Focus” (in terms of PhD allocation)

- Limited training effort
  
  “If I accepted more students, I would have to assign them improvised or mediocre projects. ... mediocre publications would be inevitable in a larger lab..” (Kyoto U)

- Excessive competition among students
  
  “Some PIs in top universities accept many students and produce numerous papers. Students in such big labs have to compete with one another inside the small lab, and some of them inevitably end up losers and become demotivated. I believe that even those losers would be stars and much more productive in a less competitive lab.” (Kyoto U)
Human Resource and Pub Impact

Many PhD students reduces the number of mediocre pubs.

Too many PhD students leads to a smaller number of high-impact pubs.
Summary

• Resource (money and HR) tends to be concentrated on limited researchers/institutions/fields
  – partly due to intrinsic force in academic society (i.e., Matthew effect)
  – but also because of government’s policies
    • “Concentration & focus” is fashionable?
    • Easier to focus rather than to find optimal distribution.

• Excessively skewed resource allocation compromises productivity for various reasons.

• In particular, unfair resource allocation can deteriorate the sense of community (Martinson, 2006; 2010).
SUPPLEMENTARY
Boundary-crossing research

NGSeasy: a next generation sequencing pipeline in Docker containers [version 1; referees: 3 approved with reservations]

Amos A Folarin, Richard JB Dobson, Stephen J Newhouse

Abstract

Motivation: Bioinformatic pipelines often use large numbers of components with substantial configuration and maintenance burden that remains a significant research challenge. Our aim is to define a new paradigm and best practices for dev pipelines encapsulated in Docker containers (lightweight virtualization), with sequencing (NGS) workflows. This approach provides several advantages: portability, versioning, and reproducibility. Using the NGSeasy pipeline, a tool for building and deploying pipelines, we have demonstrated the feasibility of this approach.

Major

- One concern is the install instructions in the article. Specifically:

  sudo make INTSALLDIR="/media/scratch" all
  sudo make install

  I am wary of using 'sudo' to install. I know that using tools like 'apt-get' require 'sudo' however for most bioinformatics software I prefer to install in my user directory simply to avoid any possible security problems. I took a look at the Makefile and I believe that sudo is not necessarily required to install, only the the INSTALLDIR and TARGET_BIN are owned by the user. Also there is typo here in 'install'.

- The project doesn't include any NGS tools related to assembly or transcriptomics. Though not stated specifically, the tools and data described here lead me to believe this project is focused around clinical applications and human genomics. If that is the case perhaps this should be clarified in the article and the title.

- The Docker security issue at the end feels tagged-on. This, I think, is a pressing concern that prevents many people from using Docker on HPC machines as opposed to on-demand computing such as AWS. This is the case at the JGI.
5,000 scientists read papers and rate them as 'Good', 'Very Good' or 'Exceptional'. F1000Prime uses the individual scores to calculate the total scores for each article, which are used to rank the articles in each discipline.