Development and testing of a database of NIH research funding of AAPM members

Working group for the development of a research database (WGDRD)

Working group for funding and grantmanship (WGFG)

WHY BUILD A RESEARCH DATABASE?
DATA MINING OF NIH RECORDS
TRENDS IN FUNDING OF AAPM MEMBERS
DISCUSSION & CONCLUSIONS

Background

"The mission of the American Association of Physicists in Medicine is to advance the science, education and professional practice of Medical Physics"

Critical metric for scientific activity: research funding

How to measure success?
The initial project is to pursue the creation of (and then continued support for) an AAPM-managed database of federally funded research in Medical Physics (NIH, NSF, DOE, DOD, etc.).

Of what use is a research database?

1. Understand the magnitude and breadth of AAPM member research activities
2. Determine trends of overall funding, and trends within specific research areas
3. Identify funding opportunities where members have been successful in the past
4. Track the success of AAPM members to specific research initiatives of the NIH
5. Understand demographics of successful researchers, identify areas with disparities
6. Identify speakers for the AAPM annual meeting and chapter meetings
7. Select reviewers for the AAPM annual meeting
8. Lobby grant-funding bodies for increased consideration of med. phys. applications
9. Perform strategic planning

WHY BUILD A RESEARCH DATABASE?
DATA MINING OF NIH RECORDS
TRENDS IN FUNDING OF AAPM MEMBERS
DISCUSSION & CONCLUSIONS
Essentially, the goal of this project was to find some way to link these two data bases together.

Input data

<table>
<thead>
<tr>
<th>Name</th>
<th>Date modified</th>
<th>Name</th>
<th>Date modified</th>
</tr>
</thead>
</table>

NIH data: downloaded from NIH Exporter
AAPM data: supplied by AAPM

Unique Identifier: PI_ID
Unique Identifier: Status_ID

Logical testing

<table>
<thead>
<tr>
<th>Test</th>
<th>True if:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Last name</td>
<td>string match</td>
</tr>
<tr>
<td>First name</td>
<td>string match</td>
</tr>
<tr>
<td>First initial</td>
<td>string match</td>
</tr>
<tr>
<td>Institution</td>
<td>Any word from institutes gives string match after common words* are removed</td>
</tr>
<tr>
<td>Funding body</td>
<td>Funding body is one of: NCI, NHLBI,NIBIB, NINDS</td>
</tr>
</tbody>
</table>

* e.g. "medical" "university" "the" "of" "physics" etc.

Inclusion Criteria

<table>
<thead>
<tr>
<th>Inclusion instance</th>
<th>Last Name</th>
<th>First Name</th>
<th>First Initial</th>
<th>Institute</th>
<th>Funding Body</th>
<th>Manual review?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>2</td>
<td>T</td>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>3</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
</tbody>
</table>
Manual review

Records are flagged for manual review if:

1. They were included via inclusion criteria three
2. The same PI_ID was matched to more than one person
3. Different PI_IDs are matched to one person
4. The last name is a ‘danger name’

False positives and negatives

- **False Positives:**
  - 100 PI_ID numbers selected at random, and manually reviewed
  - Repeated throughout development algorithm

- **False Negatives:**
  - Compared to a list of 99 member supplied PI_ID numbers

WHY BUILD A RESEARCH DATABASE?
DATA MINING OF NIH RECORDS
TRENDS IN FUNDING OF AAPM MEMBERS
DISCUSSION & CONCLUSIONS
False positives and negatives

- **False Positives:** 4% (95% CI: 1-10%)
  - 100 PI_ID numbers selected at random, and manually reviewed
  - Repeated throughout development algorithm

- **False Negatives:** 10% (95% CI: 5-18%)
  - Compared to a list of 99 member supplied PI_ID numbers

Number of grants held by AAPM members

- Yearly grants held by AAPM members from 1985 to 2015.

Total Funds awarded to AAPM members

- Total funds awarded to AAPM members from 2000 to 2015.
Total Funds awarded to AAPM members

- Total Funds
- Funding of AAPM members
- Inflation adjusted funding
- 2015 funding

Funding pool

Inflation adjusted funds per member

- Dollars per member
- Number of members
- Number of AAPM members
WHY BUILD A RESEARCH DATABASE?
DATA MINING OF NIH RECORDS
TRENDS IN FUNDING OF AAPM MEMBERS

DISCUSSION & CONCLUSIONS

Trends in AAPM funding

› Funding in real terms was lower in 2015 than in 2003
› Steady growth of AAPM in same period: heavy dilution of funds/member
› Not unique to AAPM
› Other sources of funding (Has industry picked up the slack?)

Funded members: Please provide your PI_ID!

To find your PI_ID number:
If you have an eRA Commons Account, please log into your account and click on “Personal Profile” -> “Name and ID”. Enter the eRA Person ID below. (For internal AAPM data collection purposes only.)

To enter it in AAPM website:

Members
- Change Membership Type
- Directory Online
- Directory Print

Specialty
- Funding
- Student Type
- Photo
- Certification

eRA Person ID: [numeric field only]
Conclusions

› Algorithm created to mine NIH data for AAPM member records
› Low false positive and negative rates (4%, 10%)
› Serious decline in research capacity of AAPM on a per member basis
› Database will be made available for download to AAPM members
› Database and code to be maintained by members of the WGFG

“Having algorithm and having rhythm are 2 different things.”