Trust in Data, Integrity in Science

2021-September-21 @ 4:35 pm (EST)

Register for the data management sessions here: https://www.bio-itworldexpo.com/
Introducing Trust in Data and Integrity in Science

Indicators of trust

- Trust in Research
- Trust in Professionalism
- Trust in Reproducibility

Professionalism deep dive

Wrap up
How do we check the quality of science?
Quality Indicators

Research Professionalism Reproducibility
Research

Trust in the Research determines whether papers reviewed deserve the general specification of a ‘research’ article.

Indicators for trust in research
1. Study Objective
2. Section Headings Information
Professionalism

Trust in the Professionalism ascertains the legitimacy of the authors and whether their reporting adheres to established standards of professionalism.

**Indicators for trust in professionalism**

1. Authorship verification
2. Competing Interests
3. Ethical Approval
4. Funding Statement, Source, ID
Reproducibility

Trust in Reproducibility is centered around the elements of a paper which may facilitate a future researcher’s ability to achieve the same results when replicating the original study.

**Indicators for trust in reproducibility**
1. Analysis Software
2. Software Citations
3. Statistical Analysis Methods
4. Availability of unique/biological materials
5. Code Availability Statement
6. Data Availability Statement (DAS)
7. Data Location
Trust in Professionalism
Assumption: Author is a scientist

From US Law, A Scientific Expert Must...

1. **Possess knowledge** beyond that of a lay person

2. **Be qualified**
   - Education
   - Training
   - Experience
Imposter
not who they say they are
Impersonator takes the identity or relationship of another
Peacock researcher claiming expertise in another field
Retractions associated with Professionalism

Author Unresponsive
Complaints about the Author
Concerns/Issues About
Authorship
Ethical Violations by Author
Fake Peer Review
Forged Authorship
Lack of Approval from Author
Misconduct by Author
Paper Mill
Rogue Editor

Total Retractions due to Authorship Issues by Publication Year 2010-2019

Data extrapolated from the RetractionWatch database by LD McIntosh on 2021-Apr-11
Retractions tied to Professionalism
2010-2019
What's happening in preprints?

- No standard way of reporting or removing preprints
- 32 Retracted + 3 ‘Letters of Concern’ from Covid preprints
- Still exploring checks
- Automating checks
Scams

- Unknowing (real) Coauthors
- Fake Coauthors
- Unknowing Lead Author
- False Affiliations
- Fake References
- Fake Peer Review
- Fake Lead Author
Case Studies: Imposters, Impersonators, and Cats
Case Study #1
Imposter
An Imposter

Smith has authored upwards of twenty preprints in the last two years, all about COVID and often hosted on multiple preprint servers.

Smith presents as a polymath researcher involved in military special operations and experimental medicine.

Smith has only appeared as the lead author on their articles.

Not who they say they are
<table>
<thead>
<tr>
<th>Title</th>
<th>OSF</th>
<th>Academia.edu</th>
<th>Zenodo</th>
<th>figshare</th>
<th>Authorea</th>
<th>Thesis Commons</th>
<th>Open-ventio*</th>
<th>SSRN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Novel Coronavirus: Hypothesis of Treatment with SIRT1 Inhibitors</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Vulnerabilità sistemi ICS/SCADA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Coronavirus Treatment: eradicate Coronavirus by blocking replication, counteracting its defense system</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Tenovin-1 as potential Covid-19 treatment agent</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Eradicate Coronavirus by blocking replication, counteracting its defense system</td>
<td>XT</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>6 PROBLEMS OF CORONAVIRUS PATHOGENESIS AND MEDICATIONS ACTUALLY USED</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>7 Halting Coronavirus Replication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>8 Covid-19 Pfizer Vaccine: The Worst-Case Scenario</td>
<td>XT</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>9 BNT162b2 Vaccine: possible codons misreading, errors in protein synthesis and alternative splicing's anomalies</td>
<td>XT</td>
<td>XT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

XT= Indicates both an original paper and translated version were placed on the same platform.
*Openventio is a suspected predatory publisher.
Smith’s ‘Ambiguous’ Network

1. Odd subject areas
2. Gaps in time

Publishing times allow for some rough disambiguation, along with reviewing this author’s ORCID, institutional claims, we removed extra publications.
Smith’s ‘Disambiguated’ Network

• After reviewing papers we could access with the same name, we were able to filter out some real researchers by the same name.

• The result is this ‘Disambiguated Network’, with Smith’s only coauthors coming from a parsing error and NA’s introduced by coercion.
Smith: Conclusion

• Smith case raises a variety of red flags when publication timeframes, co-authorship, history, or titles are examined.

• Still, Smith has/had:
  ○ An ORCID
  ○ Unique researcher IDs on multiple platforms
  ○ Cited by real researchers from one paper
Case Study #2
Impersonation
Impersonation of Collaboration

Impersonation is the use of another person’s identity. We are generalizing this to also include a fictitious collaboration.

**Situation:**
Real researchers are added as coauthors to a paper to appear more reputable/garner a benefit for the lead author (getting the author’s name associated with more prominent people in some field, etc)
Farzana Ahmed, the Impersonator

- Dr. Ahmed wrote about diagnosing dengue infection in children in 2019.
- Paper was retracted after the four coauthors reported it to the journal.
- Identical paper with a slightly altered title was published in an alleged predatory journal - a new coauthor replaced the four coauthors.
Diagnostic value of ferritin for the severity of dengue infection in children

Farzana Ahmed
Pediatric Intensive Care Unit, United Hospital Ltd, Dhaka, Bangladesh.

Nargis Ara Begum
Pediatric Intensive Care Unit, United Hospital Ltd, Dhaka, Bangladesh.

Md. Moshir Rahman
Pediatric Intensive Care Unit, United Hospital Ltd, Dhaka, Bangladesh.

Md. Salim Shakur
Pediatric Intensive Care Unit, United Hospital Ltd, Dhaka, Bangladesh.

Shabrina Sharmin
Pediatric Intensive Care Unit, United Hospital Ltd, Dhaka, Bangladesh.

Published
2019-10-03

How to Cite

More Citation Formats
Farzana Ahmed - Original Paper

- Republishing aside, these coauthors make for interesting case studies for our process.
  - They have multiple name-parts, most have other authors by the same name, and some make for challenging disambiguation.

<table>
<thead>
<tr>
<th>Name</th>
<th>Number of Publications</th>
<th>Number of Name Variations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moshiur Rahman</td>
<td>368</td>
<td>36</td>
</tr>
<tr>
<td>Nargis Ara Begum</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Salim Shakur</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>Shabrina Sharmin</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Serum ferritin as an early predictor of the severity of dengue infection in children

Farzana Ahmed¹ and Aftab Yousuf Raj²

¹Head of the Department of Paediatrics, Marks Medical College Hospital, Dhaka, Bangladesh
²Department of Neonatology and Paediatrics, Square Hospitals Ltd., Dhaka, Bangladesh

Corresponding author: Farzana Ahmed, Head of the Department of Paediatrics, Marks Medical College Hospital, Dhaka, Bangladesh, E-mail: farzanaahmedpaed@gmail.com

Received date: May 07, 2020; Accepted date: May 22, 2020; Published date: May 29, 2020

Copyright: © 2020 Ahmed F, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
Farzana Ahmed - New Article

- New Coauthor has very short publication history and has only written with Dr. Ahmed.
- Correspondence to Aftab Yousuf Raj is directed to: doctor2427211@yahoo.com

Journal of Pediatrics and Infants

Emerging Multi Drug Resistant Bacterial Strain in Neonatal Intensive (NICU) in a Developing Country-Bangladesh

Abstract

**Introduction:** Neonatal sepsis is one of the major causes for mortality and morbidity among newborn. Bacterial pathogens and antibiotic sensitivity patterns are different in hospitals of each country. The objective of the present study was to identify the spectrum of bacterial isolates causing septicemia in neonates and to determine their antimicrobial sensitivity pattern.

**Results:** Out of 171 babies with positive septic screening, 82 were culture positive samples, gram positive and gram negative organisms were 28% and 72% respectively. *Acinetobacter* remained the predominant isolate (n=25, 30.5%) followed by *Coagulase negative Staphylococcus* (N=16, 19.5%). Most of the gram positive isolates exhibited higher resistance to penicillin, vancomycin, and tetracycline.

Article Information

- **Article Type:** Research
- **Article Number:** JHI115
- **Received Date:** 11 March 2019
- **Accepted Date:** 22 March 2019
- **Published Date:** 28 March 2019

1. Fellow Newborn Medicine (Canada) Consultant, Pediatrics and Neonatology, Square Hospitals Ltd Bangladesh
2. Associate Professor & Head of the Department of Pediatrics, Mark Medical College Hospital Bangladesh

Aftab Yousuf Raj 1
Farzana Ahmed 2

*Corresponding author: Aftab Yousuf Raj, MBBS, MD (Pediatrics), MD (Neonatology). Fellow Newborn Medicine (Canada) Consultant, Pediatrics and Neonatology, Square Hospitals Ltd, Dhaka, Bangladesh. Tel: 01720342787. Email: doctor2427211@yahoo.com*
Farzana Ahmed - New Article

- Disambiguation is less complicated with the new article.
- Email address domains, former retractions, changing coauthors, and author networks are all potential red flags here.
- This case in particular is informative in our efforts to eliminate possible sources of bias in our algorithms (names, prominence, country of origin, etc).

<table>
<thead>
<tr>
<th>Name</th>
<th>Number of Publications</th>
<th>Number of Name Variations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Farzana Ahmed</td>
<td>39</td>
<td>4</td>
</tr>
<tr>
<td>Aftab Yousuf Raj</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>
Case Study #3
False Result
False Result

- Multiple scenarios with false results in checking authorship.
- One example of that comes from a humorous physicist out of UC Berkeley.
False Result

For this example we are considering Dr. Armstrong to be two authors, one physicist interested in neural networks and one who may as well write for the satirical news outlet *The Onion*. 
False Result

• Dr. Armstrong has nearly 40 publications.
• Almost all of them are real research articles, conducted with reputable coauthors.
• A few are preprints that are meant as a joke.
False Result

- A few are preprints that are meant as a joke
  - Dr. Armstrong is not trying to trick anyone - after an email discussion
  - Joke preprints are submitted around April Fools Day and do not have DOIs.
  - Dr. Armstrong’s real work almost always features established coauthors.
  - These joke papers are easily distinguishable with a fairly basic algorithm.
  - Scientists should have fun, and Dr. Armstrong is a great example of that which also happens to mesh well with Ripeta’s culture.
  - However, due to cultural contexts, these jokes may not be understood.
Scams

• Unknowing (real) Coauthors
• Fake Coauthors
• Unknowing Lead Author

• False Affiliations
• Fake References
• Fake Peer Review
• Fake Lead Author
Automating Authors Checks + Checking for Biases

**Step 1:**
Automated checks of named authors to confirm they are a researcher

**Step 2:**
Semi-automated check if names are not confirmed (check for biases)

**Step 3:**
Names and information reported to appropriate channels
Best practices for research integrity: Moving beyond bibliometrics
Trusting the Individual, Organizations, and Process

Ways to promote trust in your own work or organization (non-exhaustive list):

- Be transparent about your process
- Ensure data quality control
- Allow time for trust to form
- Do not advocate for a specific outcome, instead present findings and potential consequences for actions
- Include the basic quality indicators outlined for trust in research, reproducibility, and professionalism.

Tusting the Individual, Organizations, and Process

Checks to evaluate trustworthiness (non-exhaustive list):

- **Authors**: do they have previous publications? Where are they publishing their work? Do they have a social media presence? Do they use a lot of self-citations? Do they have any history of retractions?

- **Journal**: Have you read anything from this journal before? Is it reputable?

- **Funders**: Who is funding this research? What is their mission?

- **Ethics**: has the study gone through all of the necessary channels to ensure ethical research?

Source: NatureIndex [study](https://www.nature.com/articles/d41586-021-00959-3). Accessed 2021-September-13
Thank You!

Dr. Leslie McIntosh
CEO & Founder
Don't Be a Stranger

Follow us for timely industry updates, helpful blog posts, and more!

@RipetaReview

@Ripeta

www.ripeta.com
info@ripeta.com
Building Trust in Science