

# DATA MANAGEMENT in MODERN ERA

---

**Rajagopal Ramesh, Ph.D.**

**Professor**

**Department of Pathology**

**Associate Director, Cancer Research Training and  
Education Coordination (CRTEC)**

**OU Health Stephenson Cancer Center**

**University of Oklahoma Health Sciences Center**

**Email: [rajagopal-ramesh@ouhsc.edu](mailto:rajagopal-ramesh@ouhsc.edu)**



A Cancer Center Designated by the  
National Cancer Institute



# DATA INTEGRITY FAILURE IN THE NEWS !

nature

•12 December 2023

**More than 10,000 research papers were retracted in 2023 — a new record**

The number of articles being retracted rose sharply this year. Integrity experts say that this is only the tip of the iceberg.

Science News

**A prestigious cancer institute is correcting dozens of papers and retracting others after a blogger cried foul**

The criticism spotlights how artificial intelligence is playing a growing role in catching sloppy or dubious science.

**U.S. Education News**

**The Feds Want More Oversight of Scientific Research. Universities Are Fighting Back.**

Research institutions are pushing back against proposed changes to misconduct, plagiarism investigations

**Dana-Farber Cancer Institute Researchers Accused of Manipulating Data**

**A Science Sleuth Accuses a Harvard Medical School Neuroscientist of Research Misconduct**

Researchers await the outcome of an ongoing investigation into dozens of instances of alleged image problems spanning 29 publications over a period of 23 years.

•Science Insider

**Stanford president to step down** despite probe exonerating him of research misconduct  
University investigation finds data manipulation by others in Marc Tessier-Lavigne's lab and says he should have corrected work more "decisively"

•Nature, 20 July 2022

**Exclusive: investigators found plagiarism and data falsification in work from prominent cancer lab**

Ohio State University investigations identified misconduct by two scientists in lab of high-profile cancer researcher Carlo Croce. The university has cleared Croce of misconduct, but disciplined him over management problems.



# DATA INTEGRITY

---

**“Integrity is what you do when no one is watching; it's doing the right thing all the time, even when it may work to your disadvantage”**

Tony Dungy (*Hall of Fame, NFL coach*)

**Definition:** Process of ensuring that any data generated is accurate, complete, consistent, and valid throughout its lifecycle.

**Significance:** Enables individuals/entities to test and reproduce the publicly available data and validate the findings.

Serves as a background platform to advance an individuals, group or organization’s decision-making process and thus **saves time, effort and money.**



# DATA INTEGRITY: Who is RESPONSIBLE?

**Principal Investigator** - Director in-charge of directing the scientific research and receiving grant funding (local, state, federal, foundations) is responsible for oversight and ensuring all study results/data are documented, accurate, and reproducible.

**Laboratory Members** - **Trainees** (graduate students, post-doctoral fellows, medical school students, clinical fellows/residents)  
**Staff** (research technician, lab managers, program managers)

**Collaborators** - Play an important role if they are actively involved in generating data to the project.

Also responsible for checking data provided by the PI of the project at lab meetings, abstract or research publication submission.

**Institution** – Offer responsible conduct of research (RCR) course

- Ensure individual conducting any kind of research take the RCR course and subsequently take refresher course (once in 2-3 years)



# POTENTIAL FACTORS for FAILING DATA INTEGRITY

## Multifactorial

- 1. Principal Investigator** – Pressure from peers/department/institute - number of publications, impact factor, grant submission, promotion, tenure, awards, honor, etc., among others.  
Too busy! Not committing adequate time to monitor the study design, study results, interpretation, application of appropriate statistics and ensuring reproducibility; no validating data that is too good to be true!
- 2. Research students/fellows/staff** – Pressure from PI and colleagues – publication, grant submission, career advancement, faculty position, awards, honor, etc., among others.
- 3. Institution** – number of grant awards (indirect costs), national recognition, pressure to raise the bar for promotion and tenure, salary and incentives, laboratory space and resources, etc.



# DATA SHARING and COMPLIANCE

## Timelines for Data Sharing

**The time of an associated publication:** Scientific data underlying peer-reviewed journal articles should be made accessible no later than the date on which the article is first made available in print or electronic format.

**The end of the performance period:** Scientific data underlying findings not disseminated through peer-reviewed journal articles should be shared by the end of the performance period unless the grant enters into a no-cost extension. If a no cost extension is permitted, then the recipient should share the data by the end of the extended performance period. In addition, researchers should be aware that some preprint servers may require the sharing of data upon preprint posting, and repositories storing data may similarly require public release of data upon preprint posting.

## Data Preservation

Researchers are encouraged to consider relevant requirements and expectations (e.g., data repository policies, award record retention requirements, journal policies) as guidance for the minimum time frame that scientific data should be made available, which researchers may extend.

**Source:** [sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/data-sharing-approaches](https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/data-sharing-approaches)

