



NIH PUBLIC ACCESS

Author Manuscript

Int J Radiat Oncol Biol Phys. Author manuscript; available in PMC 2010 April 13.

Published in final edited form as:

Int J Radiat Oncol Biol Phys. 2010 March 1; 76(3 Suppl): S151–S154. doi:10.1016/j.ijrobp.2009.06.094.

IMPROVING NORMAL TISSUE COMPLICATION PROBABILITY MODELS: THE NEED TO ADOPT A “DATA-POOLING” CULTURE

Joseph O. Deasy, Ph.D.^{*}, Søren M. Bentzen, Ph.D.[†], Andrew Jackson, Ph.D.[‡], Randall K. Ten Haken, Ph.D.[§], Ellen D. Yorke, Ph.D.^{||}, Louis S. Constine, M.D.[‡], Ashish Sharma, Ph.D.[¶], and Lawrence B. Marks, M.D.^{**}

^{*}Department of Radiation Oncology, Mallinckrodt Institute of Radiology, Washington University School of Medicine, St. Louis, MO

[†]Department of Human Oncology, University of Wisconsin School of Medicine, Madison, WI

[‡]Department of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, NY

[§]Department of Radiation Oncology, University of Michigan, Ann Arbor, MI

^{||}Department of Radiation Oncology, University of Rochester Cancer Center, Rochester, NY

[¶]Center for Comprehensive Informatics, Emory University, Atlanta, GA

^{**}Department of Radiation Oncology, University of North Carolina, Chapel Hill, NC

Abstract

Clinical studies of the dependence of normal tissue response on dose-volume factors are often confusingly inconsistent, as the QUANTEC reviews demonstrate. A key opportunity to accelerate progress is to begin storing high-quality datasets in repositories. Using available technology, multiple repositories could be conveniently queried, without divulging protected health information, to identify relevant sources of data for further analysis. After obtaining institutional approvals, data could then be pooled, greatly enhancing the capability to construct predictive models that are more widely applicable and better powered to accurately identify key predictive factors (whether dosimetric, image-based, clinical, socioeconomic, or biological). Data pooling has already been carried out effectively in a few normal tissue complication probability studies and should become a common strategy.

Keywords

NTCP; Normal tissue complication probability models; Data sharing; Data reuse; Data pooling

THE CURRENT PACE OF PROGRESS IN NORMAL TISSUE COMPLICATION PROBABILITY MODELING: STEADY, BUT SLOW

Progress in radiation oncology should accelerate with an improved understanding of how treatment decisions affect outcomes. Key elements to improving predictive models of normal tissue toxicity include: discovering new factors (*e.g.*, dosimetric, anatomic, biological) that influence risk and validating evolving models with increasingly comprehensive datasets. In

Copyright © 2010 Elsevier Inc. Printed in the USA. All rights reserved

Reprint requests to: Joe Deasy, Ph.D., Professor, Department of Radiation Oncology, Washington University School of Medicine, 4921 Parkview Place, St. Louis, MO 63110. Tel: (314) 362-1420; Fax: (314) 362-8521; jdeasy@radonc.wustl.edu.

the face of biological complexity and measurement uncertainties, of course, models will never have perfect predictive power.

Despite a large number of dose-volume-outcome publications, made possible by the revolution in three-dimensional treatment planning, progress in normal tissue complication probability (NTCP) modeling to date has been modest. The QUANTEC reviews, though helpful, have demonstrated the limited accuracy of existing risk prediction models. New publications often take one step forward (identifying new factors or mathematical models), and one-half step back (raising issues of why different conclusions are reached compared to previous “solid” publications). A good example of this is radiation pneumonitis: some studies see a clear risk difference between upper and lower lung irradiation, some do not, and still others don’t look for the effect (see the review in this issue). Some of these discordant findings are likely due to a low statistical power resulting from a low absolute incidence of recorded toxicities. Nevertheless, there are also many examples in the QUANTEC reviews where large studies find contrasting results, including contrasting causative factors. Statistical power issues, as well as correct risk-factor identification issues, could potentially be reduced if we could pool data from much larger populations of patients.

Another example of unexplained variations between models is xerostomia. One carefully-studied dataset (1) shows little late salivary reduction at mean doses below ~40 Gy, whereas most other reports (see the xerostomia review in this issue) show some reduction in function at lower mean doses (15–20 Gy). This difference is probably related to dose–volume/location factors that are current unknown, but without actually doing a combined analysis based on the full dose distributions there is little chance of resolving the issue. Novel dose–volume metrics, that may require full access to the three-dimensional treatment planning data, are often proposed as a result of new research (*e.g.*, the weighted center of the dose distribution as a risk factor in radiation pneumonitis) (2). Another problem, not solved by data pooling is that many of the studies use different measured endpoints (*e.g.*, whole-mouth vs. single gland salivary function). The same principles hold true for tissue complication probability (TCP) modeling. Recent analyses of dose–volume factors affecting lung and prostate local control required the full dose distribution characteristics (3,4).

Inter-report comparisons of NTCP modeling results are often problematic due to differences in patient cohorts, treatment techniques, complication reporting methods, analysis methods, and models tested (Fig. 1). Currently, it is virtually impossible to revisit previously published, high-quality datasets to assess such issues raised in the later studies. Reports can be made more comparable by using only standard models, but this would impede progress: the “best” dose–volume model is never known *a priori* and the most common currently used models (*e.g.*, the Lyman-Kutcher-Burman model) are clearly oversimplified. Hence, modeling results are often not reconciled, and under our current way of doing things, the resulting literature is destined to continue to be a collection of conflicting and incomplete results.

DATA SHARING TODAY

Unfortunately, the typical method of dealing with radiotherapy data used for NTCP or TCP analyses today is illustrated in Fig. 2. The “trash can,” of course, is typically a set of computer tapes or disks that simply gather dust. In many cases storage media have deteriorated over time or the device for reading these may no longer be available. This occurs despite a commitment on the part of the NIH to emphasize the need for effective data sharing. Reports of the use of existing databases reused for NTCP modeling projects are limited.

THE CASE FOR DATA POOLING

An excellent illustration of the value of data sharing is the pooled analysis of pneumonitis data and dose–volume histograms for 540 patients from five institutions in the United States and Europe published by Kwa *et al.* more than 10 years ago (5). This study allowed the comparison of alternative NTCP models fitted to the dataset and provided a powerful argument for mean lung dose as a simple and useful metric for this clinical endpoint. The study also carefully analyzed some of the issues relating to the pooling of data from multiple studies and showed how institutional differences in the incidence of pneumonitis remained even in the final model. Searching for other causative factors will often require full access to the treatment planning data, however.

NTCP models are meant to have a wide range of applicability. Ideally, then, NTCP modeling should include datasets gathered for a wide range of treatment techniques. This was shown to be crucial, for example, in the radiation pneumonitis example of Bradley *et al.*, which combined three-dimensional institutional and Radiation Therapy Oncology Group (RTOG) clinical trial 93-11 data to arrive at a single model that well-described pneumonitis risk for many types of irradiation patterns (6). Combining datasets would normally be expected to increase the generalizability/applicability of the resulting model to a wider range of patient and dose distribution characteristics, and would shed light on important differences between datasets. Data pooling could also provide crucial statistical power in probing causative factors for rare, yet life-threatening, complications. Of course, the value of data pooling depends critically on data quality: adding unreliable data to reliable data is not progress. Data collection quality assurance and adherence to objective standards remain key issues.

Difficulties analyzing and modeling multidimensional datasets are ubiquitous in science, and are not limited to radiation oncology. Other areas in science are shifting to curation, reconciliation, and accrual of valuable datasets (this is in addition to the traditional work product of published individual reports). Many editorials, reports, and letters have detailed the expected advantages of, and obstacles to (7), moving to a “data sharing culture” (8,9). Advantages of data sharing further include the ability to replicate results, refine methods (10), address new questions with the same data (11), and make research more efficient (12).

THE NEED FOR EVOLVING MODELS

The clinical scenario being studied is never static: “baseline” treatments, adjuvant systemic therapies, RT dose delivery methods (*e.g.*, stereotactic body radiotherapy, proton therapy), and desired doses/targets/fractionation schemes, each evolve. Thus, relevant databases will need to contain broad clinical data, and dose–volume models will need to be continuously updated and adapted to new treatment conditions. Accessible NTCP/TCP data repositories should include key institutional datasets data from Phase III clinical trials representing the systematic study of the relationships between idealized data from prospective trials—often conducted in centers of excellence. However, such repositories should also contain data recorded during routine conditions to make the results as widely applicable as possible.

IS RADIATION ONCOLOGY DATA SHARING FEASIBLE?

Yes. The Image-guided Therapy Center at Washington University stores clinical trial treatment planning data for all the recent RTOG three-dimensional–based trials. Datasets from the prostate cancer three-dimensional conformal radiotherapy trial 94-06 and the lung cancer dose escalation trial 93-11 have been exported and used for extensive NTCP modeling (6,13). Although technically cumbersome, data sharing has moved forward, even for these complicated treatment planning datasets.

The American College of Radiology's web-accessible data warehouse and the National Biomedical Imaging Archive (14), for example, are powerful data repositories that will support reuse of imaging datasets. As part of an ongoing data sharing program, RTOG clinical trial 0522 data and American College of Radiology Imaging Network (ACRIN) data from a companion trial are being exported to the National Cancer Institute Imaging Archive database, including both Fluorine-18-deoxyglucose-positron emission tomography (FDG-PET) and digital imaging and communications in medicine (DICOM-RT) data objects. Washington University has similarly developed a web-based database architecture to store myriad datasets, including fully archived treatment planning datasets with open-source viewing tools. Publicly accessible datasets do not necessarily need to be shared in a single massive data warehouse. Data repositories can potentially be decentralized so long as common data formats are used and the appropriate middleware is available.

In addition, software is now freely available that allows conversions into a vendor-neutral archive file from a wide range of commercial and academic treatment planning systems (15). The same software system (computational environment for radiotherapy research) has extensive tools to review and analyze this format on personal computers.

DISCOVERABLE DATABASES

The inability to be able to query for the existence of an archive can often be just as big a roadblock as it is to obtain access to the data itself. We envision that institutional databases could become discoverable for the sake of contributing to universal queries with anonymized summaries of available data. In this way (as a fictitious example), a resident at Medical School A could make a universal query, and learn that Medical School B had 86 cases relevant to radiation proctitis after proton therapy, and that the cases included dose distributions, outcomes (with 13 responders), and computed tomography images. At this point, no protected health information of detailed data values would be shared, only aggregate "inventory" summaries of available data.

Such a system, linking together discoverable databases, would be the key motivator to move to the next step and begin the relatively more cumbersome process of contacting the appropriate researchers and beginning the institutional review board process.

The National Cancer Institute's caBIG program has made significant progress in this area through the use of an architecture paradigm wherein individual datasets are kept under institutional control yet made discoverable (16,17). A remote user can query multiple datasets across institutions. These "grid services" are also very sensitive to the security needs of multi-institutional access. After users have discovered the appropriate datasets, the data custodians/investigators could be contacted about obtaining research approval. The grid infrastructure would then provide the researcher with a single "virtual" archive, customized using their credentials and privileges, from which they can get their data. The caGrid effort thus provides one potential solution to the problem of a universal query of datasets that could be pooled.

OBSTACLES TO DATA POOLING

There are significant administrative and professional issues involved in regulating access to data while still expediting research. Biomedical on the one hand, cooperative trial groups need to protect the opportunity of researchers/physicians who generated the data to get a "first crack" at data analysis and publication, whereas the larger societal need is to maximize the benefit of the entire process for scientific/clinical advancement. A compromise is needed. We suggest that cooperative groups grant wider access to anonymized data after some set time after publication of the original envisioned analyses. Further citations and data reuse would only increase the value of the original investment in the clinical trial.

Despite our enthusiasm, it needs to be said that data pooling is not a panacea to NTCP modeling problems, especially because differences in datasets can obscure the looked-for effects. In addition to the added cost of keeping data in usable formats, other obstacles to fully leveraging archived data include variations in: outcome reporting and analysis (RTOG vs. Common Terminology Criteria for Adverse Events scale vs. institutional scales), the prevalence of known or unknown comorbidities, dose calculation algorithms (older methods vs. convolution or Monte Carlo), setup procedures, contouring policies, etc. (See reviews by Bentzen *et al.* and Jackson *et al.* in this issue).

Some limitations of current NTCP studies can be overcome by a greater reliance on collaboration, for example as part of multi-institutional trials. Nonetheless, making such data discoverable and available for pooled analyses is still superior to one-time use, especially considering the probability that the models are likely to change.

RECOMMENDATIONS

Apart from the practical issues discussed previously, progress in making the archiving of key datasets the norm in radiation oncology research will require cooperation from major grant-funded stakeholders, including individual investigators, cooperative groups, and journal editors. Despite these (surmountable) obstacles, creating a culture of data reuse in the radiation oncology research community will have long-term advantages and would very likely lead to improved radiation oncology outcomes. A complete shift to making data archiving a requirement for publishing is arguably problematic because of technical costs. To maximize the use and relevance of clinical trials data, cooperative groups may want to adopt a policy of anonymizing clinical trials data and making the data publicly accessible after a reasonable delay. This delay would enable publication of all the investigator-driven, planned studies. Digital data, unlike tissue specimens, do not become depleted with use, and so less access management would be required. For these reasons, we encourage the establishment of key databanks of linked treatment planning, imaging, and outcomes data.

Costs associated with data curation/quality assurance, anonymization, access policy, data verification, data format conversion, and data extraction of course need to be considered by funding agencies. Potentially, discoverable databases could be conveniently queried for linked lists of data including outcomes, images, dose, tissue samples, and previously conducted bio-array results. Much of the software and associated standards required to do this has been developed.

In contrast to a data pooling culture, our current “data loss” paradigm leaves NTCP and TCP modeling in a Sisyphean cycle: rolling a slightly different boulder up a slightly different hill, over and over. We need to escape from this trap and gain the advantages of data sharing. Converting to a data reuse paradigm will require changes in operation by cooperative groups, institutional investigators, and journal. Informatics support will be crucial. Many groups across the sciences have already called for a wholesale change in process that supports data reuse. Some scientific journals, especially in genomics-related areas, already require submission of supporting datasets with research articles. However, adoption of policies to “publish the data” has been much slower in fields dealing with sensitive health information.

If the radiation oncology world were to adopt a data reuse policy, progress toward improved NTCP models (and other types of treatment effect comparisons) would accelerate, new factors relevant to outcomes would be identified, and the road block to consensus would be surmountable. It is only by making published datasets available for ongoing combined analyses that we can hope to produce powerful and validated models of quantitative normal tissue effects in the clinic.

Acknowledgments

Development of this statement was partially supported by NIH grant CA85181 (JOD).

REFERENCES

1. Li Y, Taylor JM, Ten Haken RK, et al. The impact of dose on parotid salivary recovery in head and neck cancer patients treated with radiation therapy. *Int J Radiat Oncol Biol Phys* 2007;67:660–669. [PubMed: 17141973]
2. Hope AJ, Lindsay PE, El Naqa I, et al. Clinical, dosimetric, and location-related factors to predict local control in non-small cell lung cancer. *Int J Radiat Oncol Biol Phys* 2005;63:S231.
3. Mu Y, Hope AJ, Lindsay P, et al. Statistical modeling of tumor control probability for non-small-cell lung cancer radiotherapy. *Int J Radiat Oncol Biol Phys* 2008;72:S448.
4. van Herk M, Witte MG, Heemsbergen WD, et al. Relation between dose outside the prostate next term and failure free survival in the Dutch previous term prostate next term cancer trial. *Int J Radiat Oncol Biol Phys* 2008;72:S65.
5. Kwa SL, Lebesque JV, Theuws JC, et al. Radiation pneumonitis as a function of mean lung dose: An analysis of pooled data of 540 patients. *Int J Radiat Oncol Biol Phys* 1998;42:1–9. [PubMed: 9747813]
6. Bradley JD, Hope A, El Naqa I, et al. A nomogram to predict radiation pneumonitis, derived from a combined analysis of RTOG 9311 and institutional data. *Int J Radiat Oncol Biol Phys* 2007;69:985–992. [PubMed: 17689035]
7. Colditz GA. Constraints on data sharing experience from the Nurses' Health Study. *Epidemiology* 2009;20:169–171. [PubMed: 19234411]
8. Piwowar HA, Chapman W. Envisioning a biomedical data reuse registry. *AMIA Annu Symp Proc* 2008:1097. [PubMed: 18998885]
9. Parr CS, Cummings MP. Data sharing in ecology and evolution. *Trends Ecol Evol* 2005;20:362–363. [PubMed: 16701396]
10. Hernan MA, Wilcox AJ. Epidemiology, data sharing, and the challenge of scientific replication. *Epidemiology* 2009;20:167–168. [PubMed: 19234410]
11. Van Horn JD, Ishai A. Mapping the human brain: New insights from fMRI data sharing. *Neuroinformatics* 2007;5:146–153. [PubMed: 17917125]
12. Koslow SH. Sharing primary data: A threat or asset to discovery? *Nat Rev Neurosci* 2002;3:311–313. [PubMed: 11967561]
13. Tucker SL, Dong L, Bosch WR, et al. Fit of a generalized Lyman normal-tissue complication probability (NTCP) model to Grade ≥ 2 late rectal toxicity data from patients treated on protocol RTOG 94-06 next term. *Int J Radiat Oncol Biol Phys* 2007;69:S8–S9. [PubMed: 17848302]
14. National Biomedical Imaging Archive. [Accessed March 24, 2009]. Available online at: <https://cabig.nci.nih.gov/tools/NCIA>
15. Deasy JO, Blanco AI, Clark VH. CERR: A computational environment for radiotherapy research. *Med Phys* 2003;30:979–985. [PubMed: 12773007]
16. Saltz J, Kurc T, Hastings S, et al. e-Science, caGrid, and translational biomedical research. *Computer* 2008;41:58–64.
17. Oster S, Langella S, Hastings S, et al. caGrid 1.0: An enterprise grid infrastructure for biomedical research. *J Am Med Inform Assoc* 2008;15:138–149. [PubMed: 18096909]

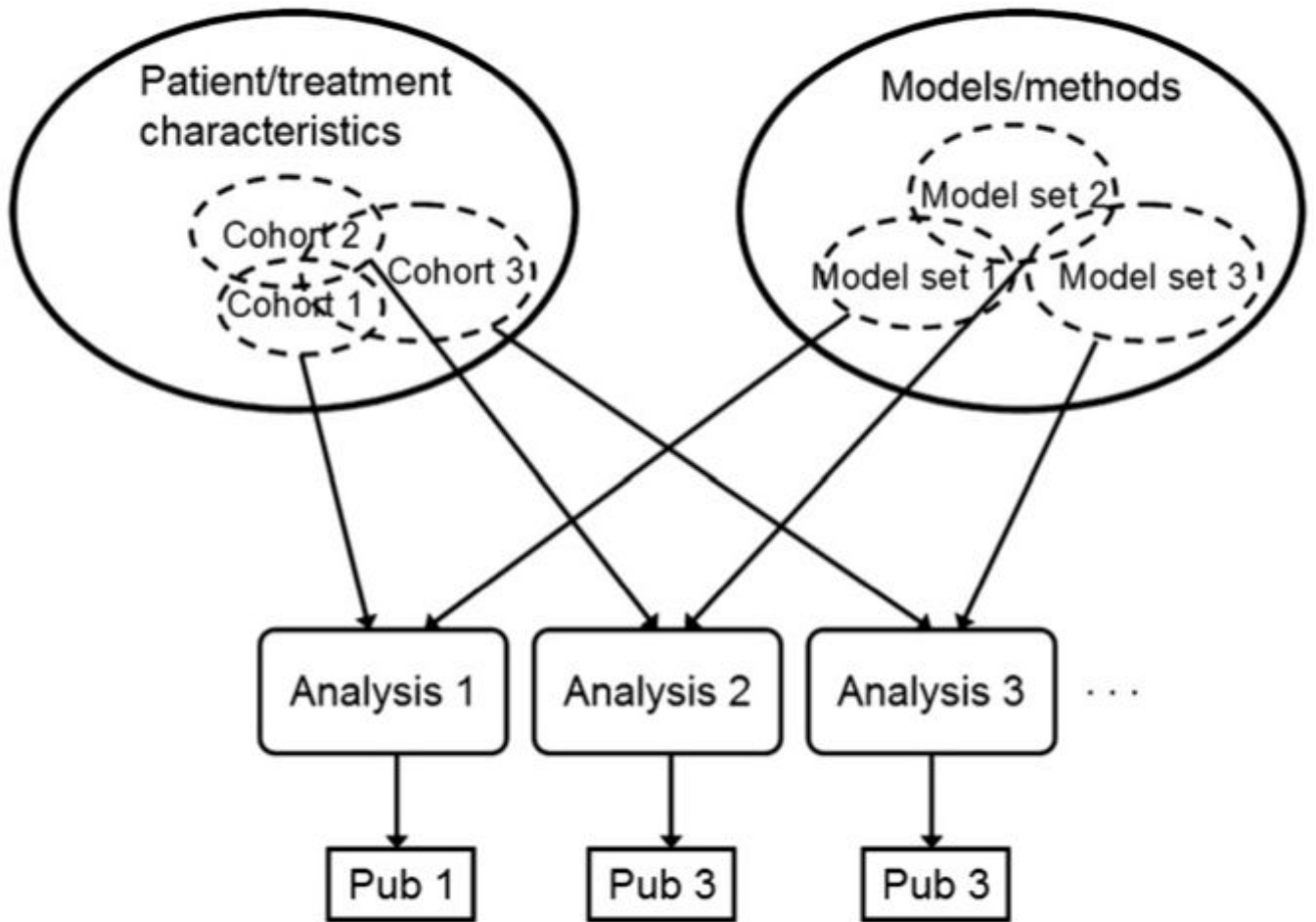


Fig. 1.

Why does normal tissue complication probability (NTCP) modeling frequently lead to incompatible results? The current paradigm consists of applying a range of evolving methods (models tested, structures included, etc. to datasets that at least partially differ in patient, disease, and treatment characteristics). This inevitably leads to inconsistent results and impedes the validation of NTCP models for broad clinical use. It will be necessary to pool data to escape this trap.

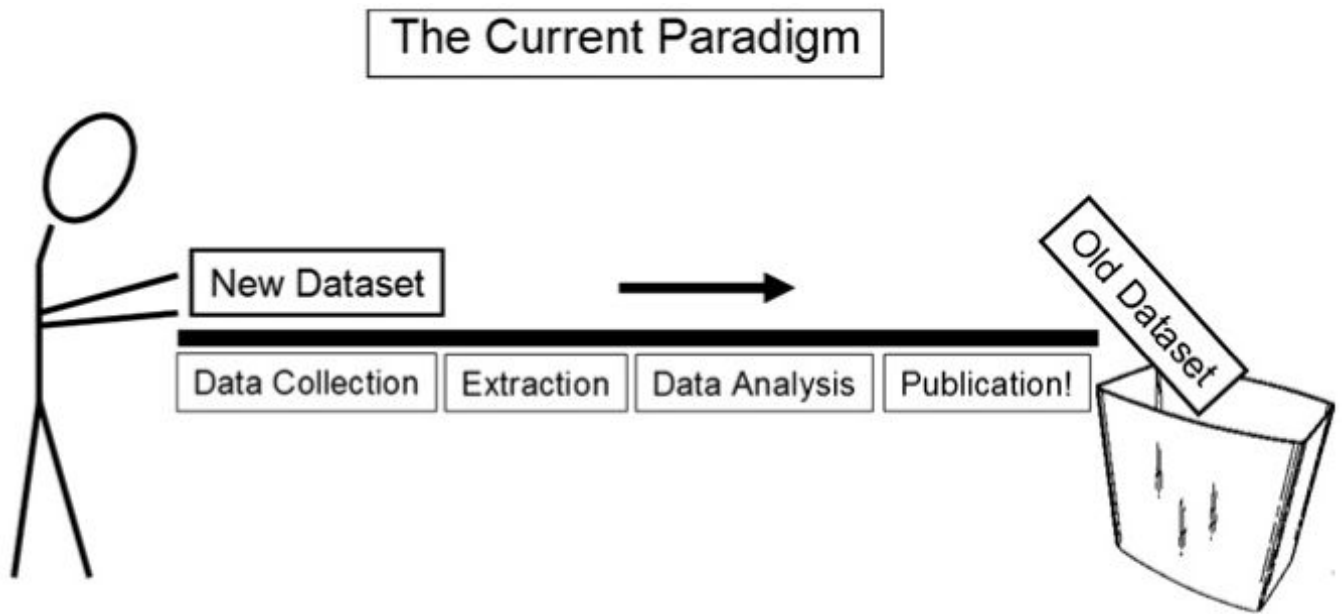


Fig. 2. “The current (data-loss) paradigm.” Data are effectively lost to the wider scientific community after publication. Capturing key datasets in query-able data repositories would accelerate the discovery of causative factors and increase the accuracy of parameter estimates.