

University of Groningen

Meeting the Challenge of Scientific Dissemination in the Era of COVID-19

Fuller, Clifton D.; van Dijk, Lianne V.; Thompson, Reid F.; Scott, Jacob G.; Ludmir, Ethan B.; Thomas, Charles R.

Published in:
International Journal of Radiation Oncology Biology Physics

DOI:
[10.1016/j.ijrobp.2020.06.066](https://doi.org/10.1016/j.ijrobp.2020.06.066)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2020

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Fuller, C. D., van Dijk, L. V., Thompson, R. F., Scott, J. G., Ludmir, E. B., & Thomas, C. R. (2020). Meeting the Challenge of Scientific Dissemination in the Era of COVID-19: Toward a Modular Approach to Knowledge-Sharing for Radiation Oncology. *International Journal of Radiation Oncology Biology Physics*, 108(2), 496-505. <https://doi.org/10.1016/j.ijrobp.2020.06.066>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

SCIENTIFIC COMMUNICATION

Meeting the Challenge of Scientific Dissemination in the Era of COVID-19: Toward a Modular Approach to Knowledge-Sharing for Radiation Oncology



Clifton D. Fuller, MD, PhD,^{*,†} Lisanne V. van Dijk, PhD,^{*,‡}
Reid F. Thompson, MD, PhD,[†] Jacob G. Scott, MD, DPhil,[§]
Ethan B. Ludmir, MD,^{*} and Charles R. Thomas, Jr, MD[†]

^{*}Department of Radiation Oncology, The University of Texas MD Anderson Cancer, Houston, Texas;

[†]Department of Radiation Medicine, Oregon Health & Science University, Oregon; [‡]Department of

Radiation Oncology, University Medical Center- Groningen, Groningen, Netherlands; and

[§]Department of Radiation Oncology, Taussig Cancer Institute, Cleveland Clinic, Cleveland, Ohio

Received Jun 27, 2020. Accepted for publication Jun 28, 2020.

On May 1 and May 22, 2020, a pair of high-profile articles were fast-track reviewed and published by the *New England Journal of Medicine (NEJM)* and *The Lancet*, venues widely regarded as among the most prestigious of medical journals.^{1,2} The *Lancet* article reported a multinational registry analysis of chloroquine

with or without macrolide antibiotics in patients who were infected with the novel severe acute respiratory syndrome corona virus-2 virus, and an *NEJM* manuscript from the same group investigated angiotensin-converting enzyme inhibitors and angiotensin-receptor blockers in patients who tested positive for coronavirus disease 2019

Corresponding author: Clifton D. Fuller, MD, PhD; E-mail: cdfuller@mdanderson.org

C.D.F. received/receives funding and salary support related to this project from the National Institutes of Health (NIH) National Institute of Biomedical Imaging and Bioengineering (NIBIB) Research Education Programs for Residents and Clinical Fellows Grant (R25EB025787-01).

Disclosures: C.D.F. has received direct industry grant support, honoraria, and travel funding from Elekta AB unrelated to this project. C.D.F. received/receives funding and salary support unrelated to this project from the National Institute for Dental and Craniofacial Research Establishing Outcome Measures Award (1R01DE025248/R56DE025248) and Academic Industrial Partnership Grant (R01DE028290); NCI Early Phase Clinical Trials in Imaging and Image Guided Interventions Program (1R01CA218148); an NIH/NCI Cancer Center Support Grant (CCSG) Pilot Research Program Award from the UT MD Anderson CCSG Radiation Oncology and Cancer Imaging Program (P30CA016672); an NIH/NCI Head and Neck Specialized Programs of Research Excellence (SPORE) Developmental Research Program Award (P50 CA097007); NIH Big Data to Knowledge (BD2K) Program of the National Cancer Institute (NCI) Early Stage Development of Technologies in Biomedical Computing, Informatics, and Big Data Science Award (1R01CA2148250);

National Science Foundation (NSF), Division of Mathematical Sciences, Joint NIH/NSF Initiative on Quantitative Approaches to Biomedical Big Data (QuBBD) Grant (NSF 1557679); NSF Division of Civil, Mechanical, and Manufacturing Innovation (CMMI) grant (NSF 1933369); and the Sabin Family Foundation. Direct infrastructure support is provided to Dr Fuller by the multidisciplinary Stiefel Oropharyngeal Research Fund of the University of Texas MD Anderson Cancer Center Charles and Daneen Stiefel Center for Head and Neck Cancer and the Cancer Center Support Grant (P30CA016672) and the MD Anderson Program in Image guided Cancer Therapy.

S.V.D. receives direct salary support unrelated to this manuscript from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Dutch Research Council) Rubicon Award (452182317). J.G.S. receives direct salary and resource support unrelated to this submission under an NIH/NCI Method to Extend Research in Time (MERIT) Award for Early Stage Investigators (R37CA244613). R.F.T. receives salary and project support unrelated to this manuscript through the Veterans Administration Clinical Science Research and Development (CSR) mechanism (IK2 CX002049).

Supplementary material for this article can be found at <https://doi.org/10.1016/j.ijrobp.2020.06.066>.

(COVID-19). These papers would have been a pinnacle achievement for the academic coauthors, in addition to the supporting company Surgisphere, who reportedly supplied the data. Led by the vascular surgeon Sapan Desai, this small company with “big data” aspirations redefined research priorities and patient study allocation with their remarkable results. Unfortunately, these august journals would soon be roiled by controversy when it became evident that the data may have been falsified for both papers.³ The subsequent debacle serves as a cautionary tale of the systematic failure modes of traditional avenues of sharing and verifying clinical science, particularly when applied to fast-tracked research.

Warning signs regarding the scientific integrity of these publications were posted not in a traditional journal, but via the *Zenodo* preprint server, as a near-immediate open letter to the *Lancet*.⁴ Statistician James Watson led signatories to critique the *Lancet* and *NEJM*'s fidelity to their own policies on data transparency, noting, among other issues: “the [Surgisphere] authors have not adhered to standard practices in the machine learning and statistics community. They have not released their code or data” nor external study preregistration with an ethics board. The letter demanded “Surgisphere provide[s] details on data provenance, [with] independent validation of the analysis [and] open access to all the data sharing agreements cited above...” to verify findings in the *Lancet* article.⁴ A retraction of the *Lancet* article followed, as the data could not be verified. In early June 2020, the results in *NEJM* were similarly repudiated, “after concerns were raised with respect to the veracity of the data and analyses conducted by Surgisphere Corporation.”⁴

The COVID-19 pandemic has exposed both long-standing and emerging issues with scientific review and dissemination. Although the pace and scope of scientific output in response to the pandemic are commendable and necessary, it has outstripped already fragile capacity and accountability mechanisms for ensuring scientific internal validity, rapid dissemination, credibility, and verifiability. Failure to achieve these critical components of scientific communication and credibility has tremendous potential for real-world harm (as in the Surgisphere debacle). Therefore, it is imperative that the scientific community optimally balance speed with rigor, and is held to account via transparent, modular, and verifiable standards to maximize reproducible research. Achieving these ambitious objectives for improving transmission of scientific knowledge requires using a diverse array of novel tools at our disposal.⁵

Preprints: Accelerated Research Transmission in a Pandemic

The COVID-19 crisis gripping the world has justifiably led to an increased need for efficient scientific dissemination, with a resultant rapidity observed in efforts at transmission across both traditional streams of scientific discourse

(eg, scientific manuscripts,^{6,7} society journal consensus guidelines⁸) as well as more novel mechanisms with various levels of peer review (eg, preprints,⁹ social media¹⁰). In 2019, a clinical medicine preprint repository, medRxiv, was made publicly available, allowing clinical research to be posted before peer review in a manner mimicking those widely used in physics (arXiv), psychology (psyarxiv), chemistry (chemRxiv), engineering (engRxiv), social sciences (SocArxiv), and basic biomedical science (bioRxiv). Coupled with the pressing thirst for usable information amid a fatal pandemic, the use of preprints has blossomed since late January 2020¹¹; similarly, reputable peer-reviewed journals from *NEJM* to the *International Journal of Radiation Oncology Biology Physics* have accelerated both their review processes and online posting of accepted peer-reviewed manuscripts with impressive legerity, leading to admirably paced communication in a time when timeliness is critical.

The value of preprints for near-immediate dissemination of research findings, their ease of use, their ability to circumvent traditional journal politics and dominant narratives, and their accompanying lack of restrictions before knowledge-sharing has been highlighted even further with COVID-19, with 32% of the National Institutes of Health (NIH) Office of Portfolio Analysis represented as preprints, whereas the PubMed/preprint ratio was at 3% as of last year. In the interval since COVID-19 became a global pandemic, the preprint phenomenon appears, anecdotally, to have grown substantially within radiation oncology. In a recent example, to combat initial shortfalls of personal protective equipment (PPE), Twitterati on social media forums began to discuss the practicability of radiation as a method for PPE sterilization and reuse.^{12,13} Within days, pilot protocols were developed and a preprint generated to use laboratory biosafety cabinets as a method to stretch dangerously short supplies of previously disposable PPE for health care workers.¹⁴ Near simultaneous efforts were made by other groups at ultraviolet-based sterilization and were disseminated directly via university website.¹⁵ At present, none of the relevant scientific content from these works has yet been credentialed as peer-reviewed.¹ Nonetheless, in the heart of the PPE shortage, decisions were actively being made how best to reuse life-saving PPE with these prereview data from radiation oncologists. Without preprints, we would not know these findings even existed in a timely enough fashion to consider for practice during an acute PPE shortage.

Although there has previously existed fierce debate about the potential and pitfalls of preprints, the wave of COVID-19 preprints has rendered even heated theoretical arguments about the acceptability of preprints practically moot, as a tsunami of research teams have raced to get results on the servers as fast as possible.¹⁶⁻¹⁹ Although the mass uptake of preprints for COVID-19 data are evident (medRxiv/biorXiv catalogs more than 5000 COVID-19–related preprints at present), cancer, which kills nearly 600,000 persons per year

in the United States, has not seen the same embrace of preprints to date by the radiation oncology community (ie, ~150 preprints met search criteria for the keywords “radiotherapy” OR “radiation oncology”).

A challenge of preprints, in many instances, is that the lay public, media, and some scientists have treated preprints as fully vetted scientific analyses, and amplified papers that would likely not pass muster in a more thorough journal review process; thus, indicators of internal validity that are evidence of some level of pre–peer-review rigor are a useful indicator of quality for rapid dissemination such as preprints. Critics of preprints also point to rampant dissemination of poor quality or methodologically thin research as drawbacks of the platform and suggest this demonstrates the necessity of peer-review (although extant preprint data suggest that the majority of the effects of peer-review may in fact be cosmetic); thus, efforts to mitigate publication biases in both preprint and ultimate peer-review contexts are warranted.

Preregistration: A Tool to Combat Internal Bias in, During, and After COVID?

The findings of the seminal work “Why Most Published Research Findings Are False” have become so recited as to almost be a mantra for critics of scientific discourse,²⁰ who note widespread biases across published scientific literature. Foremost among these are “p-hacking” (repetition of analyses until a “significant” result emerges), the “file drawer effect” (whereby positive trials are published and negative studies are either not reported or discarded to an editorial manuscript limbo), “salami slicing” or duplicate publication bias (publishing multiple pieces of redundant or nearly overlapping research), or any number of other identifiable biases plague clinical research.^{21,22} Among these, P-hacking and interpretation bias (eg, “borderline significance” for nonsignificant statistical endpoints, “statistically significant” but clinically inconsequential observed differences) have vexed statisticians to the extent of revising definitions of statistical significance.²³ This bias becomes especially critical in situations in which there is minimal established comparative prior knowledge, such as during the current pandemic involving a novel corona virus with high morbidity and mortality. Preregistration can serve as a powerful tool to mitigate these pitfalls.

By specifying one’s research plan on a registry in advance of performing the study (preregistration), or submitting the methodology and statistical design to a journal for review before performing the study (prereview), biases can be prevented, or at least identified more clearly. Preregistration also reduces the capacity for data molding, P-hacking, or convenient hypotheses shifting, and may thus increase analytical rigor, as early results show preregistration increases the publication of null findings.^{24,25}

Impressively, the *Red Journal* was among the first journals to pilot prereview²⁶; however, to date, most radiation oncology researchers have eschewed preregistration, anecdotally arguing that the process commits the authors to a single journal. Further, the current *Red Journal* website makes no notation of prereview, and many radiation oncologists are, anecdotally, unfamiliar with the concept. In addition to journal-based prereview checklists, groups like the Open Science Foundation provide an independent avenue of easily used templates to preregister hypotheses, planned experiments, sample sizes, and statistical analyses, and can be generated before or after data collection.^{27,28} The resulting time-stamped, digital object identifier (DOI)–labeled document, called a registered report, need not be constraining, but serves to demonstrate that alterations from planned research activities were fully transparent in intent and execution. Sadly, the process has not become normative in radiation oncology, with 1 glaring exception: clinical trials.

The *Red Journal*, like most within its scientific echelon, since 2015 has stipulated all clinical trials must be preregistered (eg, listed in clinicaltrials.gov), stating “Taken together, mandatory trial registration improves transparency, reduces the potential for bias, and should help to allay public concerns regarding possible manipulation of research findings for commercial or academic benefit.”²⁹ Clearly, when it comes to clinical evidence, preregistration, if not prereview, is considered a stable standard in radiation oncology (as in other fields) in scenarios where critical patient care decisions are concerned.

A recent episode illustrates the powerful combination of preregistration and preprints. This spring, interest in reviving historic methods of low-dose radiation therapy for pneumopathy from early in the last century resulted in a piquant social media discourse, a subsequent review, editorials, and a flurry of commentary in the *Green Journal*³⁰⁻³⁴ Rather than the standard peer-review timeline, or responses via editorial, a domestic group from Emory University had preregistered a trial, executed the pilot study, published results as preprint on medRxiv, and had an overview of the work featured online in *Forbes*^{35,36} before the next issue of the *Green Journal* had arrived in mailboxes. Given the fact that radiation oncology is rarely mentioned in general audience magazines like *Forbes*, and given the discussion generated regarding the article within the scientific community, this approach, from a pure dissemination assessment, rendered peer review ancillary, if not moot. Advocates of preregistration have accelerated review processes in response to the novel coronavirus; for example, after an initiative by the *Royal Society Open Science* aiming for review of registered reports in less than 1 week from submission, a bevy of other journals and more than 450 referees volunteered to assist in the effort.³⁷

As *Red Journal* Editor-in-Chief Anthony Zeitman presciently noted in a 2017 commentary: “Many investigators report their work in data repositories, on

archive sites, or on their own websites, all publicly available and easily searchable. This is the postjournal world toward which we appear to be heading.”³⁸ We appear to have entered an era in clinical science post-COVID; however, it remains to be seen whether these instances represent large-scale adoption or crisis-driven limited use-cases.

Open Access: Publication Accessibility and Equity After the Current Crisis

Overlaying this discussion of preprints is a similar discussion regarding full accessibility of findings, or open access. Starting in 2004, and mandated in 2008, the NIH stipulated that publicly funded research be made available to the public via PubMed Central “immediately upon acceptance for publication.”³⁹ This effort to democratize access to federally funded research has moved apace with the Berlin, Budapest, and Salvador declarations, which describe the need for global accessibility to research data as an essential matter of equity and which have driven a suite of international policies and activities designed to increase open access to manuscripts after peer review.^{40–42} The COVID-19 pandemic toppled barriers to publication access across platforms and stakeholders,^{43–45} as early in the pandemic the United States and international National Science and Technology Advisors challenged publishers to voluntarily make COVID-19–related publications (in addition to data) immediately accessible in public repositories such as PubMedCentral.⁴⁶

As we move (presumably) to a post-COVID-era, Plan S, a pan-European initiative, is slated to start in January 2020 (now pushed back to 2021). It would require that all EU-funded efforts immediately upon acceptance “must be published in compliant Open Access Journals or on compliant Open Access Platforms” and asserts “all researchers should be able to publish their work Open Access,”⁴⁷ ensconcing the current COVID-19–related open access imperatives across biomedical science (at least for the EU). It has been proposed that deposition of preprints would meet the mandated criteria for open-access, with a finalized version subsequently published by a traditional journal, leading to a confluence of consideration regarding the relationship between preprint, peer-review, and postpublication open access. Notably, Plan S also mandates that, to be compliant, a journal must provide within open access publications the ability to directly “link to raw data and code in external repositories.”

Data Sharing: FAIRness in Data Accessibility

The data provenance underlying the Surgisphere analyses was almost immediately suspect.⁴ Further, the editors of the *Lancet* and *NEJM* were, fascinatingly, not just critiqued for the acceptance of the manuscript, but also for their inability

to verify data provenance. This, at some level is, in our estimation, a fairly novel critique for journal editors; previously, the idea that the purview of a journal would entail, in any sense, provision of some certitude of experimental data quality or origin, would have seemed bizarre in the pre-electronic era, when laboratory notebooks were physical objects, rather than a JuPyTer notebook. Data and code accessibility (which, we must be clear, is an entirely different enterprise than traditional peer review has typically concerned itself with) has become not just an appendage, but a central normative component of peer review. Recent retractions of a set of articles from “*L’affaire Surgisphere*” are particularly informative and illustrate another issue: the limitations of both peer-reviewed manuscripts and preprints in terms of quality control with regard to data availability and data transparency.

The need for a structure to index and annotate these shared data has led to a segment of journals dedicated to publishing data (as opposed to analyses) using templated records explaining the structure and content of deposited information called data descriptors.⁴⁸ Some traditional journals (eg, The American Association of Physicists in Medicine’s flagship journals, *Medical Physics* and the *Journal of Applied Clinical Medical Physics*) facilitate publication of a citable data descriptor (ie, a formal description of a publicly accessible data set, describing the data and directing the interested reader to the relevant *data repository* as a PubMed-listed citation). Sadly, our corresponding flagship radiation oncology journals do not formally offer such a resource for data descriptors. Other virtual journals have stepped into the breach, offering avenues for publishing peer-reviewed data descriptors, provided the descriptor uses “field specific” standards of annotation and the data repositories meet (somewhat vague) community data norms. Adding another layer of complexity, some data repositories (ie, the storage location or data warehouse where shared files/information is permanently housed), such as the NIH Figshare, also serve as avenues for nonpeer-reviewed data publication, issuing a DOI but without external assessment. Thus, as with preprints, publication does not necessarily imply peer review, again confusing the uninitiated, who heretofore have been able to rely on the near one-to-one linkage between “published” and “peer-reviewed.”⁴⁹

COVID-19 has, if not for radiation oncology, for infectious disease research, again been transformative through disruption. In addition to the implementation of existing data repositories (such as the Global Initiative to Share All Influenza Data (GISAID) viral genome datasharing platform previously used for influenza and avian flu), experts have called for making full-scale bulk anonymized electronic medical record (EMR) data broadly accessible to researchers, stating, “In this interconnected world, we can imagine a unifying multinational COVID-19 electronic health record waiting for global researchers to apply their methodological and domain expertise.”⁵⁰

Table 1 The FAIR guiding principles⁵²

Principles	Concepts
To be Findable:	F1. (Meta)data are assigned a globally unique and persistent identifier F2. Data are described with rich metadata (defined by R1 below) F3. Metadata clearly and explicitly include the identifier of the data it describes F4. (Meta)data are registered or indexed in a searchable resource
To be Accessible:	A1. (Meta)data are retrievable by their identifier using a standardized communications protocol A1.1 The protocol is open, free, and universally implementable A1.2 The protocol allows for an authentication and authorization procedure, where necessary A2. Metadata are accessible, even when the data are no longer available
To be Interoperable:	I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation I2. (Meta)data use vocabularies that follow FAIR principles I3. (Meta)data include qualified references to other (meta)data
To be Reusable:	R1. Meta(data) are richly described with a plurality of accurate and relevant attributes R1.1. (Meta)data are released with a clear and accessible data usage license R1.2. (Meta)data are associated with detailed provenance R1.3. (Meta)data meet domain-relevant community standards

Driven by the fiery manifesto of Future Of Scholarly Communication and e-Scholarship (FORCE11) to “Rethink the unit and form of the scholarly publication,”⁵¹ NIH has recently embraced data-sharing via adoption of findable, accessible, interoperable, and reusable (FAIR) guiding principles. Data must be findable, accessible, interoperable, and reusable, not only at the human level, but also for machines (eg, indexing tools or software) (Table 1).^{52,53} The FAIR principles describe the “how” of data infrastructure as an outgrowth of an ethos of transparency and equity that has been growing in medicine for years, and are supported by the Institute of Medicine, encouraging stakeholders (eg, cooperative groups, journal editors, specialty societies) that “data sharing [is]...the expected norm.”⁵⁴ Throughout the scientific enterprise, data dissemination, rather than merely article acceptance, is now a direct end-goal of the scientific process. This is a radical break from past eras, when data were hoarded as a precious commodity and not treated as intellectual commons, and norms are evolving in many fields.⁵⁵ The need for data sharing and repositories has also become more apparent via technical innovations such as machine and deep learning, as these require large pooled data sets to generate data-driven clinical decision models.⁵⁶⁻⁵⁸ This need particularly has accelerated the FAIR data principles’ integration in radiation oncology for structured machine readability (ie, index-ability/search-ability) and annotated data curation, as opposed to data *qua* unrefined data, are imperative. These principles are structurally reflected and supported by recent investment in shared data infrastructure, such as the NIH Strategic Plan for Data Science and Data Commons Framework.^{59,60} COVID-19 has raised the imperative for data deposition, sprouting new data sharing avenues such as the European COVID-19 Data Platform, Australia’s Academic and Research Network (AARNet) COVID-19 Data Resource Repository, and other open-access data venues.⁶¹⁻⁶³

In our minds, although preregistration, preprints, peer-reviewed, postpublication open access, and FAIR-principled data publications have synergistic value, they are designed to serve fundamentally different purposes, offered through different venues, and serve different communities of stakeholders. Simply put, these scientific “modules” serve to provide gains in distinct domains of scientific dissemination:

- Preregistration increases internal validity of scientific knowledge dissemination.
- Preprints increase speed/transferability of scientific knowledge dissemination.
- Peer review increases credibility/interpretability of scientific knowledge dissemination.
- Open access increases the availability/equity of scientific knowledge dissemination.
- Data availability increases reproducibility/reusability of scientific knowledge dissemination.

A Proposal for Transparent Modular Scientific Dissemination

The fact that these structural modules serve distinct, related, but nonoverlapping goals has led to the current ecosystem where a multitude of iterative, unlinked steps results in a decentralized and poorly standardized corpus, all indexed in different manners and scarcely identifiable as a single consolidated scientific enterprise (Fig. 1). For example, in a scenario where a study is corrected, retracted, or edited or there is a variance in published data found, there may be no direct method to ensure, outside of the original author’s good judgment, that corrections, errata, or retractions are perpetuated “upstream” (ie, back to the original preregistration or preprint) or “downstream” (ie, to a prior journal, open-access, or data repository), although there are avenues in PubMed for linking the retraction notice to the prior publication.

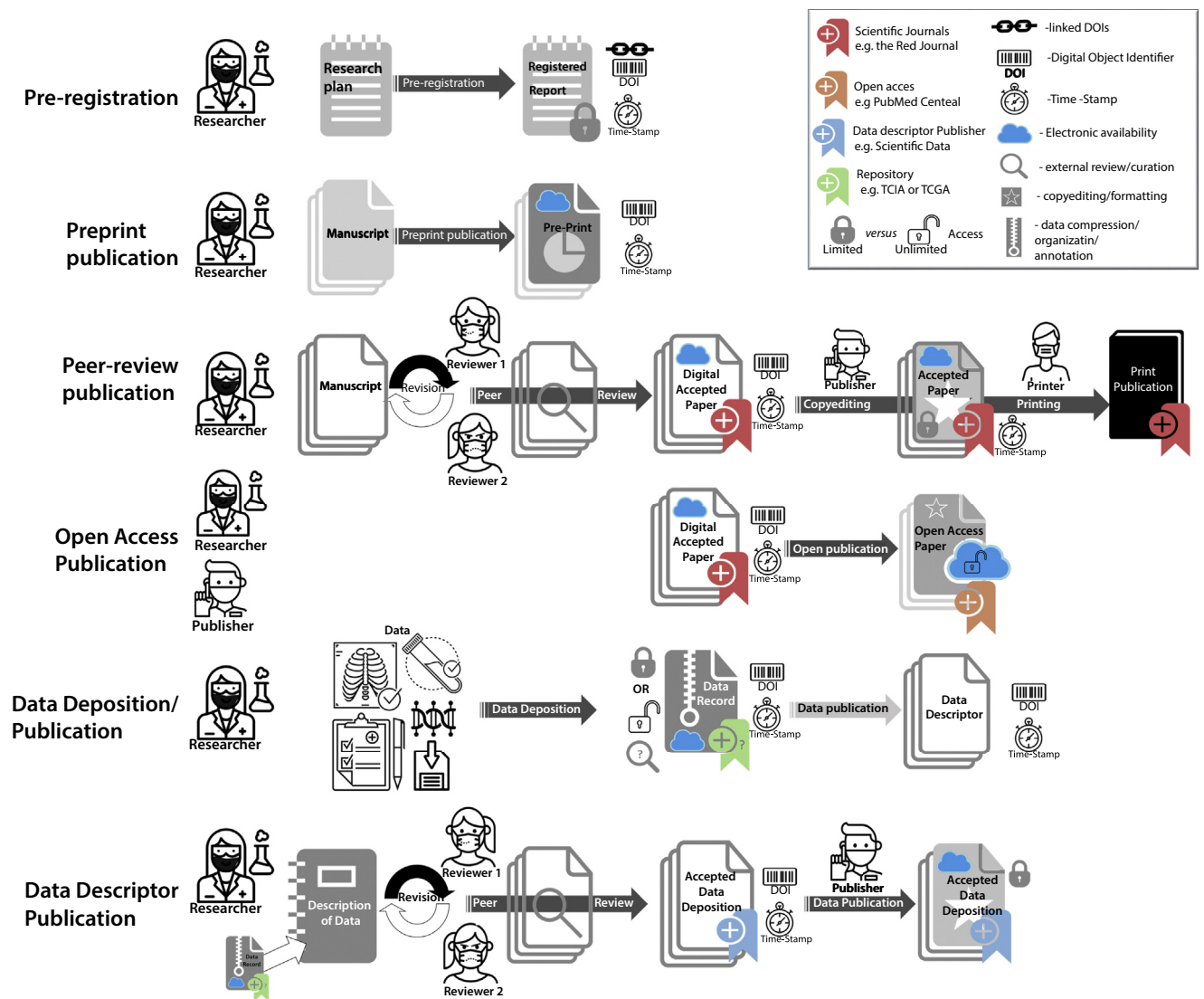


Fig. 1. Graphical representation of current disparate and independent scientific dissemination processes, showing actors, activities, and component documents/data/software.

In our estimation, a major conceptual limitation that precludes the scientific enterprise in COVID19 is the historic reliance on the peer-reviewed manuscript as the definitive “quanta” of scientific information. Historically, manuscripts were considered as complete, self-enclosed, and self-contextualizing, using the standard introduction, methods, results, and discussion format in such a way that, presumably, the peer referee and ultimate reader could reconstruct the scientific enterprise of the author with some veracity. In an era of simple experiments and high-trust nondiverse stakeholders, reputational assessment served to preclude inadvertent lack of scientific rigor or value (as peer review presumes good faith on all actors and is not calibrated particularly well for adversarial fraud detection). However, with modern experimental design (eg, Bayesian statistics, Markov models, machine learning approaches) and massive big data analyses using increasingly complex statistical techniques on increasingly arcane data elements (viz EMR data, epidemiologic records, genome-wide

association studies, or radiomics variables), how could any reasonable person expect the interpretable whole of a scientific undertaking to be performed in a manner derived when the requisite referee skillset was a deep knowledge of the (admitted much smaller) corpus of extant literature and some basic knowledge of statistics?

Instead, we advocate the ideal of a unified thematic project, with direct linkages between component processes, in what might be termed as “scientific modules” as puzzle pieces of a larger holistic process (Fig. 2). In the most simple instance, this would involve a “check-list” notification that records the linked metadata or simply the DOIs of all prior “modules,” with manual posthoc updating of indexed archival documents at “project” completion.⁶⁴ For example, at preprint submission, the existence of a registered report or an extant data deposition would be formally affirmed or denied; if affirmed, the relevant DOI(s) would be provided and added as a link on the preprint server. Similarly, at peer review, the referees would

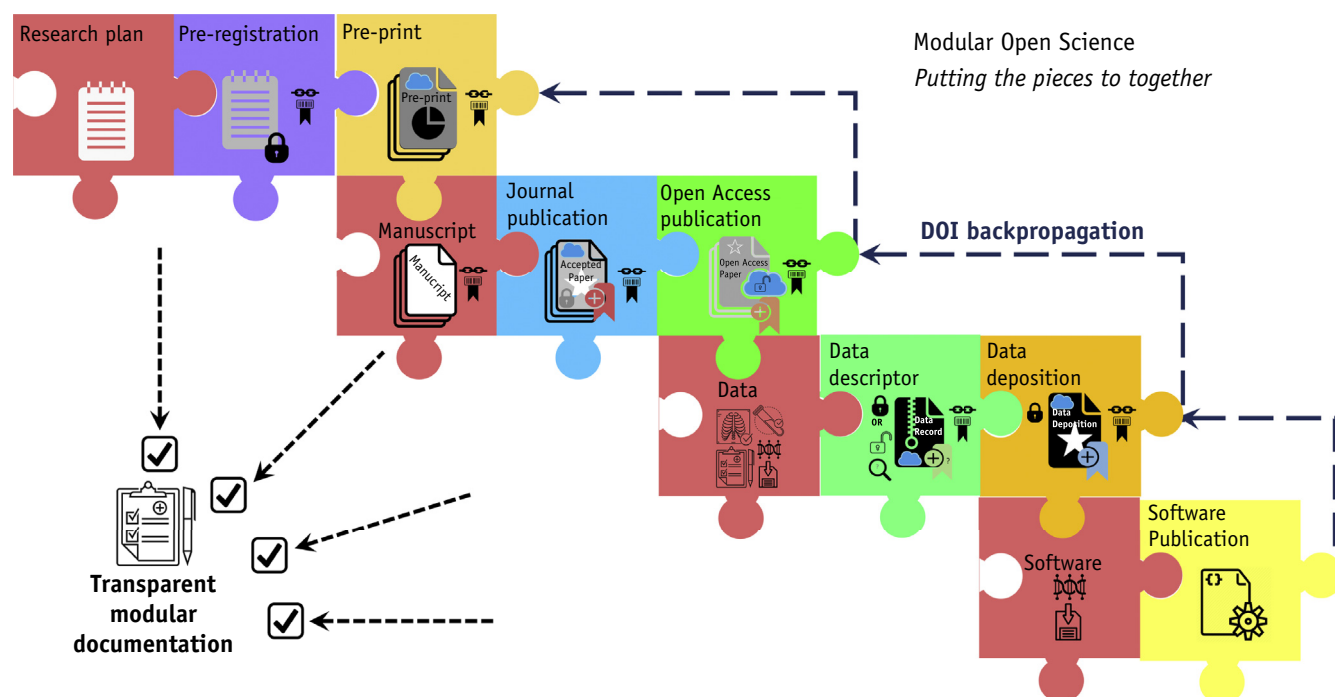


Fig. 2. Proposed transparent modular scientific dissemination process, using either provenance metadata or digital object identifier (DOI) linkages to link iterative or sequential processes for a scientific process. Presumably, these linkages could be further refined via technical interoperability between current discrete processes in [Figure 1](#).

have confirmation of the existence or absence of an available registered report, preprint, and data deposition, and these could then be used as ancillary justification for acceptance or revision. Finally, at “final” article or data publication, the prior work would be listed as serial DOIs, allowing ready access to all modular components as a single package or linked “provenance metadata.”^{65,66} A preliminary checklist ([Modular Science Checklist](#)) has been drafted by the authors; conceivably it or an analogous document could be submitted with each “scientific module” (preprint, data descriptor, peer-review submission, etc) for clarity, with a final version completed after deposition/publication of all modules, as an analog “content tracker form” to assure transparency across a series of currently disparate steps,⁶⁷ until end-user usable standardized provenance metadata solutions (such as those proposed by Mahmood et al⁶⁷) are realized in radiation oncology specifically or medical science generally.

In the future, however, we can envision an integrated (or at least, interoperable) modular science dissemination process ([Fig. 2](#)), where a software infrastructure is capable of linking prior, current, and future modules dynamically, such that a researcher need not bear the onus of modular science unassisted. For example, linked DOIs (if not metadata) of all modules across the project could be forward- or back-propagated dynamically. This would prove especially valuable in cases of retraction, errata, correction, or data updates, as the relevant information would be “embedded” not only in later references but in previously submitted modules; this is truly transparent

science, but would require that the current puzzle of systems be formalized at some level, which requires a singular committed vision of the scientific process above and beyond each modular aim.⁶⁸

Reasonably, we feel that this will happen, if at all, through the leadership of concerted scientific societies and publication, journal, or repository agents serving as mediators of a modular science infrastructure, as individual researchers are unlikely to see added value, even in the instance of increased scientific transparency and rigor, in the context of substantively increased *ad hoc* clerical burden. However, as COVID-19 has shown us, in a crisis, new avenues of dissemination (ie, preprints) may surge in popularity, even in the absence of direct integration with traditional publication venues; by consolidating the process, scientific communities can enhance the quality of the entire modular scientific publication “chain” rather than myopically concerning themselves only with the traditional safe harbor of peer review.

Epilogue: Back to the Future

This is not the first time a novel zoonotic pathogen has served to spur speed in scientific missives. During the 17th century, when most print media was heavily government controlled, the waves of plague afflicting England necessitated rapid public tracking of up-to-date regional mortality. In that era, as in ours, the information dissemination rapidity resulted in a lack of central control.

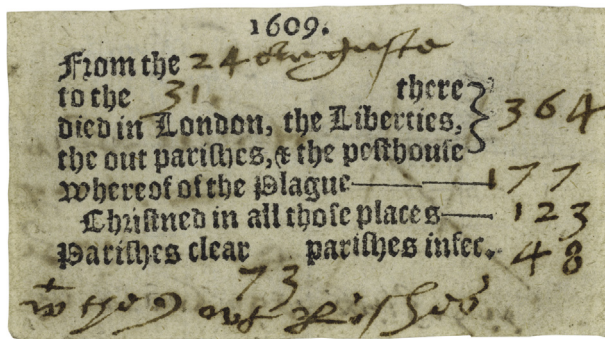


Fig. 3. Simplified plague report cutting from a London handbill of mortality for August 24 to August 31, 1609 (courtesy of the Folger Shakespeare Library, used under CC BY-SA license).⁷¹

What came to be known as the “Pamphlet Wars” was a feud between the upstart Chemical Physicians Society and the exclusive College of Physicians. With a desire for increased influence, these educational societies, which were limited groups of ~50 elite members, each published their own recommendations and guidance and disseminated highly divergent (and often erroneous) information. The varying and contradictory (mis)information of these pamphlets served to confuse the public; fraud and quackery were rampant. However, soon, a heretofore unconventional printing of handbills began (Fig. 3), with lay demographers jumping in to track rates of disease, publishing in English rather than Latin (the professional language of physicians). This data democratization meant that, rather than only being the purview of the learned and initiated through the filter of medical jargon, *everyone* had access to raw data, meaning that “ordinary Londoners, in addition to their governors, [could] anticipate a rise or fall in mortality, and...turn to medicine or prayers as circumstances or inclinations dictated.”⁶⁹

This historical move toward open data publication not only reified the idea that research was not just for the elite, but also served as the accepted beginning of statistical epidemiology, allowing a rapid dissemination of data at a time generally thought to be starved of information.⁷⁰ Today, we must decide—as a specialty—whether we will turn the 2019-nCoV coronavirus crisis into an impetus for improved faster, better, FAIR-er science, or whether we ignore the pressing need for transparent transmission of timely knowledge and consequently produce research that may benefit the endemic cancer crisis worldwide. In our minds, the concept of transparent modular science is a “disruptive integration” that brings the strengths of various scientific formats into a cohesive whole and represents an avenue for our specialty to lead into the post-COVID era with the full employ of each approach; otherwise, like the physicians of previous eras, we may find our pontifications ignored by those who can more ably share and disseminate data in democratic and interpretable formats in a timely and accountable manner.

References

- Mehra MR, Desai SS, Ruschitzka F, Patel AN. Retracted: Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. *Lancet* 2020. [https://doi.org/10.1016/S0140-6736\(20\)31180-6](https://doi.org/10.1016/S0140-6736(20)31180-6).
- Mehra MR, Desai SS, Kuy S, Henry TD, Patel AN. Cardiovascular disease, drug therapy, and mortality in COVID-19. *N Engl J Med* 2020;102:1-7.
- Mehra MR, Desai SS, Kuy S, Henry TD, Patel AN. Retraction: Cardiovascular disease, drug therapy, and mortality in COVID-19. *N Engl J Med*. <https://doi.org/10.1056/nejmc2021225>. Accessed June 27, 2020.
- Watson J. An open letter to Mehra et al and The Lancet. *Zenodo*. <https://doi.org/10.5281/zenodo.3862789>. Accessed June 27, 2020.
- Brainard J. Scientists are drowning in COVID-19 papers. Can new tools keep them afloat? *Sci Community/Coronavirus* 2020. <https://doi.org/10.1126/science.abc7839>.
- Rivera A, Ohri N, Thomas E, Miller R, Knoll MA. The impact of COVID-19 on radiation oncology clinics and patients with cancer in the United States [epub ahead of print]. *Adv Radiat Oncol*. <https://doi.org/10.1016/j.adro.2020.03.006>. Accessed June 27, 2020.
- Rao YJ, Provenzano D, Gay HA, Read PW, Ojong M, Goyal S. A radiation oncology departmental policy for the 2019 novel coronavirus (COVID-19) pandemic. *Preprints*. <https://doi.org/10.20944/preprints202003.0350.v1>. Accessed June 27, 2020.
- Thomson DJ, Palma D, Guckenberger M, et al. Practice recommendations for risk-adapted head and neck cancer radiation therapy during the COVID-19 pandemic: An ASTRO-ESTRO consensus statement. *Int J Radiat Oncol Biol Phys* 2020;107:618-627.
- Kwon D. How swamped preprint servers are blocking bad coronavirus research. *Nature* 2020;581:130-131.
- Simcock R, Thomas TV, Estes C, et al. COVID-19: Global radiation oncology’s targeted response for pandemic preparedness. *Clin Transl Radiat Oncol* 2020;22:55-68.
- Kupferschmidt K. Preprints bring ‘firehose’ of outbreak data. *Science* 2020;367:963-964.
- @joe_deasy. “Sterilizing N95 masks for reuse. In the current crisis we at MSK are considering giving bags of N95 masks say 5kilo Gy of xrt to get an estimated 3 log10 virus deactivation level. Anyone else considering similar? Your thoughts?” March 21, 2020. https://twitter.com/joe_deasy/status/1241451201830555648?s=20. Accessed June 27, 2020.
- @CancerConnector. “Our team has finished updates to our preprint to use idle biosafety cabinets to sterilize #PPE for reuse during the #COVID19 pandemic. We just submitted to @medrxivpreprint but it will be a few days before it’s online. Please RT.” March 25, 2020. <https://twitter.com/CancerConnector/status/1242849194609725442>. Accessed June 27, 2020.
- C. C. L. R. I. Theory Division et al. UV sterilization of personal protective equipment with idle laboratory biosafety cabinets during the COVID-19 pandemic. *medRxiv*. <https://doi.org/10.1101/2020.03.25.20043489>. Accessed June 27, 2020.
- Lowe JJ, Paladino KD, Farke JD, et al. N95 filtering facemask respirator ultraviolet germicidal irradiation (uvgi) process for decontamination and reuse. Available at: <https://www.nebraskamed.com/sites/default/files/documents/covid-19/n-95-decon-process.pdf>. Accessed June 27, 2020.
- Bagdasarian N, Cross GB, Fisher D. Rapid publications risk the integrity of science in the era of COVID-19. *BMC Med* 2020;18:192.
- Fraser N, Brierley L, Dey G, Polka JK, Pálffy M, Coates JA. Preprinting a pandemic: the role of preprints in the COVID-19 pandemic. *bioRxiv*. <https://doi.org/10.1101/2020.05.22.111294>. Accessed June 27, 2020.
- Trott KR, Zschaek S, Beck M. Radiation therapy for COVID-19 pneumopathy. *Radiother Oncol* 2020;147:210-211.

19. Fidahic M, Nujic D, Runjic R, et al. Research methodology and characteristics of journal articles with original data, preprint articles and registered clinical trial protocols about COVID-19. *BMC Med Res Methodol* 2020;20:161.
20. Ioannidis JPA. Why most published research findings are false. *PLOS Med* 2005;2:e124.
21. Sterling TD, Rosenbaum WL, Weinkam JJ. Publication decisions revisited: The effect of the outcome of statistical tests on the decision to publish and vice versa. *Am Stat* 1995;49:108-112.
22. Nagarajan PJ, Garla BK, Taranath M, Nagarajan I. The file drawer effect: a call for meticulous methodology and tolerance for non-significant results. *Indian J Anaesth* 2017;61:516-517.
23. Wasserstein RL, Schirm AL, Lazar NA. Moving to a world beyond 'p < 0.05'. *Am Stat* 2019;73:1-19.
24. Kupferschmidt K. More and more scientists are preregistering their studies. Should you? *Science*. <https://doi.org/10.1126/science.aav4786>. Accessed June 27, 2020.
25. Allen C, Mehler DMA. Erratum: Open science challenges, benefits and tips in early career and beyond. *PLoS Biol* 2019;17:1-14.
26. Mell LK, Zietman AL. Introducing prospective manuscript review to address publication bias. *Int J Radiat Oncol Biol Phys* 2014;90:729-732.
27. Foster ED, Deardorff A. Open science framework (OSF). *J Med Libr Assoc* 2017;105:203-206.
28. Dehaven A. Open Science Foundation Wiki: Templates of OSF registration forms. Available at: <https://osf.io/zab38/wiki/home/>. Accessed June 27, 2020.
29. Palma DA, Zietman A. Clinical trial registration: A mandatory requirement for publication in the red journal. *Int J Radiat Oncol Biol Phys* 2015;91:685-686.
30. @subatomicdoc. "Quick outline of a trial to tear apart here @ProfAJChalmers @ShankarSiva @cd_fuller @CancerConnector @roentgen66 @CarmeloT2681 @CormacSmall." April 4, 2020. <https://twitter.com/subatomicdoc/status/1246416459452080128>. Accessed June 27, 2020.
31. Kirsch DG, Diehn M, Cucinotta FA, Weichselbaum R. Lack of supporting data make the risks of a clinical trial of radiation therapy as a treatment for COVID-19 pneumonia unacceptable. *Radiother Oncol* 2020;147:217-220.
32. Kirsch DG, Diehn M, Cucinotta FA, Weichselbaum R. Response letter: Radiation therapy for COVID-19 pneumopathy. *Radiother Oncol* 2020;149:238-239.
33. Kirkby C, Mackenzie M. Is low dose radiation therapy a potential treatment for COVID-19 pneumonia? *Radiother Oncol* 2020;147:221.
34. Kefayat A, Ghahremani F. Low dose radiation therapy for COVID-19 pneumonia: A double-edged sword. *Radiother Oncol* 2020;147:224-225.
35. Conca J. Preliminary data suggests low-dose radiation may be successful treatment for severe COVID-19. *Forbes* 2020. Available at: <https://www.forbes.com/sites/jamesconca/2020/06/12/1st-human-trial-successfully-treated-covid-19-using-low-doses-of-radiation/>. Accessed June 27, 2020.
36. Conca J. How low-dose radiation could be the trick for treating COVID-19. *Forbes* 2020. Available at: <https://www.forbes.com/sites/jamesconca/2020/05/13/researchers-explore-low-doses-of-radiation-to-treat-severe-coronavirus-cases/>. Accessed June 27, 2020.
37. Cardiff U, Chambers. NeuroChambers Blog: CALLING ALL SCIENTISTS: Rapid evaluation of COVID19-related registered reports at Royal Society Open Science. Available at: <http://neurochambers.blogspot.com/2020/03/calling-all-scientists-rapid-evaluation.html>. Accessed June 27, 2020.
38. Zietman AL. The ethics of scientific publishing: Black, white, and 'fifty shades of gray. *Int J Radiat Oncol Biol Phys* 2017;99:275-279.
39. Zerhouni EA. NIH public access policy. *Science* 2004;306:1895.
40. Berlin declaration on open access to knowledge in the sciences and humanities. Available at: <https://openaccess.mpg.de/Berlin-Declaration>. Accessed June 27, 2020.
41. Forgo PN, Nnetwich M, et al. Vienna declaration: 10 theses on freedom of information. Available at: http://web.archive.org/web/20060304153832/http://www.chaoscontrol.at:80/2005/we_english.htm. Accessed June 27, 2020.
42. O. S. I. Open Society Institute. The Budapest open access initiative. Available at: <https://www.budapestopenaccessinitiative.org/read>. Accessed June 27, 2020.
43. United Nations Educational Scientific and Cultural Organization. Open access to facilitate research and information on COVID-19. Available at: <https://en.unesco.org/covid19/communicationinformationresponse/open-solutions>. Accessed June 27, 2020.
44. Grove J. Open-access publishing and the coronavirus. *Insid High Ed*. Available at: <https://www.insidehighered.com/news/2020/05/15/coronavirus-may-be-encouraging-publishers-pursue-open-access>. Accessed June 27, 2020.
45. Carr D. Publishers make coronavirus covid-19 content freely available and reusable. Available at: <https://wellcome.ac.uk/press-release/publishers-make-coronavirus-covid-19-content-freely-available-and-reusable>. Accessed June 27, 2020.
46. COVID-19 open-access letter from chief scientific advisors to members of the scholarly publishing community. Available at: <https://www.whitehouse.gov/wp-content/uploads/2020/03/COVID19-Open-Access-Letter-from-CSAs.Equivalents-Final.pdf>. Accessed June 27, 2020.
47. cOALition S/European Science. Plan S. Available at: https://www.coalition-s.org/wp-content/uploads/PlanS_Principles_and_Implementation_310519.pdf. Accessed June 27, 2020.
48. Ward CH. Introducing the data descriptor article. *Integr Mater Manuf Innov* 2015;4:190-191.
49. Poline JB. From data sharing to data publishing. *MNI Open Res* 2019;2:1.
50. Cosgriff CV, Ebner DK, Celi LA. Data sharing in the era of COVID-19. *Lancet Digit. Heal* 2020;2:e24.
51. Bourne PE, Clark TW, Dale R, et al. Force11 White paper: Improving the future of research communication and e-scholarship. <https://doi.org/10.4230/DagMan.1.1.41>, Accessed June 27, 2020.
52. Wilkinson MD, Dumontier M, Aalbersberg IJJ, et al. The FAIR guiding principles for scientific data management and stewardship. *Sci Data* 2016;3:160018.
53. Wilkinson MD, Sansone SA, Schultes E, Doorn P, Da Silva Santos LOB, Dumontier M. A design framework and exemplar metrics for FAIRness. *Sci Data* 2018;5:1-4.
54. Committee on Strategies for Responsible Sharing of Clinical Trial Data; Board on Health Sciences Policy; Institute of Medicine, *Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk*. Washington, DC: National Academies Press (US); 2015.
55. Tenopir C, Dalton ED, Allard S, et al. Changes in data sharing and data reuse practices and perceptions among scientists worldwide. *PLoS One* 2015;10:e0134826.
56. Lustberg T, van Soest J, Jochems A, et al. Big data in radiation therapy: Challenges and opportunities. *Br J Radiol* 2017;90:1069.
57. Goble C, Cohen-Boulakia S, Soiland-Reyes S, Garijo D, Gil Y, Crusoe MR. FAIR computational workflows. *Data Intell* 2019;2:108-121.
58. Draxl C, Scheffler M. Big data-driven materials science and its FAIR data infrastructure. *Handb Mater Model* 2020;49:73.
59. U.S. Dept of Health and Human Services. National Institutes of Health (NIH) strategic plan for data science. *Natl Heal Inst Strateg Plan Data Sci*. <https://doi.org/10.1371/journal.pbio.1002195>. Accessed June 27, 2020.
60. Grossman RL. Data commons framework. Available at: <https://datascience.cancer.gov/news-events/blog/introducing-data-commons-framework>. Accessed June 27, 2020.
61. National Institutes of Health. Open-access data and computational resources to address COVID-19. Available at: <https://datascience.nih.gov/covid-19-open-access-resources>. Accessed June 27, 2020.
62. Covid-19 data portal. Available at: <https://www.covid19dataportal.org/>. Accessed June 27, 2020.

63. Rosewall G. How COVID-19 data is being shared with researchers globally. Available at: <https://news.aarnet.edu.au/how-covid-19-data-is-being-shared-with-researchers-globally/>. Accessed June 27, 2020.
64. Metadata 2020: Best practices. Available at: <http://www.metadata2020.org/blog/2020-03-17-metadata-practices/>. Accessed June 27, 2020.
65. Hartig O, Zhao J. Publishing and consuming provenance metadata on the web of linked data. *Lect Notes Comput Sci* 2010;6378:78-90.
66. Hartig O. Provenance information in the Web of data. *CEUR Workshop Proc* 2009;538.
67. Mahmood T, Jami SI, Shaikh ZA, Mughal MH. Toward the modeling of data provenance in scientific publications. *Comput Stand Interfaces* 2013;35:6-29.
68. Wood I, Larson JW, Gardner H H. A Vision and Agenda for Theory Provenance in Scientific Publishing. Chen L, Liu C, Liu Q, Deng K, editors. Database Systems for Advanced Applications DASFAA 2009, Lecture Notes in Computer Science. vol 5667. Berlin, Heidelberg: Springer; 2009. ISBN 978-3-642-04205-8.
69. Slack P. Metropolitan government in crisis: The response to plague [in] London 1500-1700: The making of the metropolis. *Longman* 1986;60:80.
70. Greenberg S. Plague, the printing press, and public health in seventeenth-century London. *Huntington Library Quarterly* 2004;67: 508-527.
71. *London - BILLS OF MORTALITY - Blank Bills (Briefs) 1609*. Folger Shakespeare Library; 1609.