



Registering Study Analysis Plans (SAPs) Before Dissecting Your Data—Updating and Standardizing Outcome Modeling

Maria Thor*, Jung Hun Oh, Aditya P. Apte and Joseph O. Deasy

Department of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, NY, United States

OPEN ACCESS

Edited by:

Tiziana Rancati,
Istituto Nazionale dei Tumori
(IRCCS), Italy

Reviewed by:

Alexander F. I. Osman,
Al-Neelain University, Sudan
Wilma Desiree Heemsbergen,
Erasmus University
Rotterdam, Netherlands

*Correspondence:

Maria Thor
thorm@mskcc.org

Specialty section:

This article was submitted to
Radiation Oncology,
a section of the journal
Frontiers in Oncology

Received: 20 January 2020

Accepted: 18 May 2020

Published: 24 June 2020

Citation:

Thor M, Oh JH, Apte AP and
Deasy JO (2020) Registering Study
Analysis Plans (SAPs) Before
Dissecting Your Data—Updating and
Standardizing Outcome Modeling.
Front. Oncol. 10:978.
doi: 10.3389/fonc.2020.00978

Public preregistration of study analysis plans (SAPs) is widely recognized for clinical trials, but adopted to a much lesser extent in observational studies. Registration of SAPs prior to analysis is encouraged to not only increase transparency and exactness but also to avoid positive finding bias and better standardize outcome modeling. Efforts to generally standardize outcome modeling, which can be based on clinical trial and/or observational data, have recently spurred. We suggest a three-step SAP concept in which investigators are encouraged to (1) Design the SAP and circulate it among the co-investigators, (2) Log the SAP with a public repository, which recognizes the SAP with a digital object identifier (DOI), and (3) Cite (using the DOI), briefly summarize and motivate any deviations from the SAP in the associated manuscript. More specifically, the SAP should include the *scope* (brief data and study description, co-investigators, hypotheses, primary outcome measure, study title), in addition to step-by-step details of the *analysis* (handling of missing data, resampling, defined significance level, statistical function, validation, and variables and parameterization).

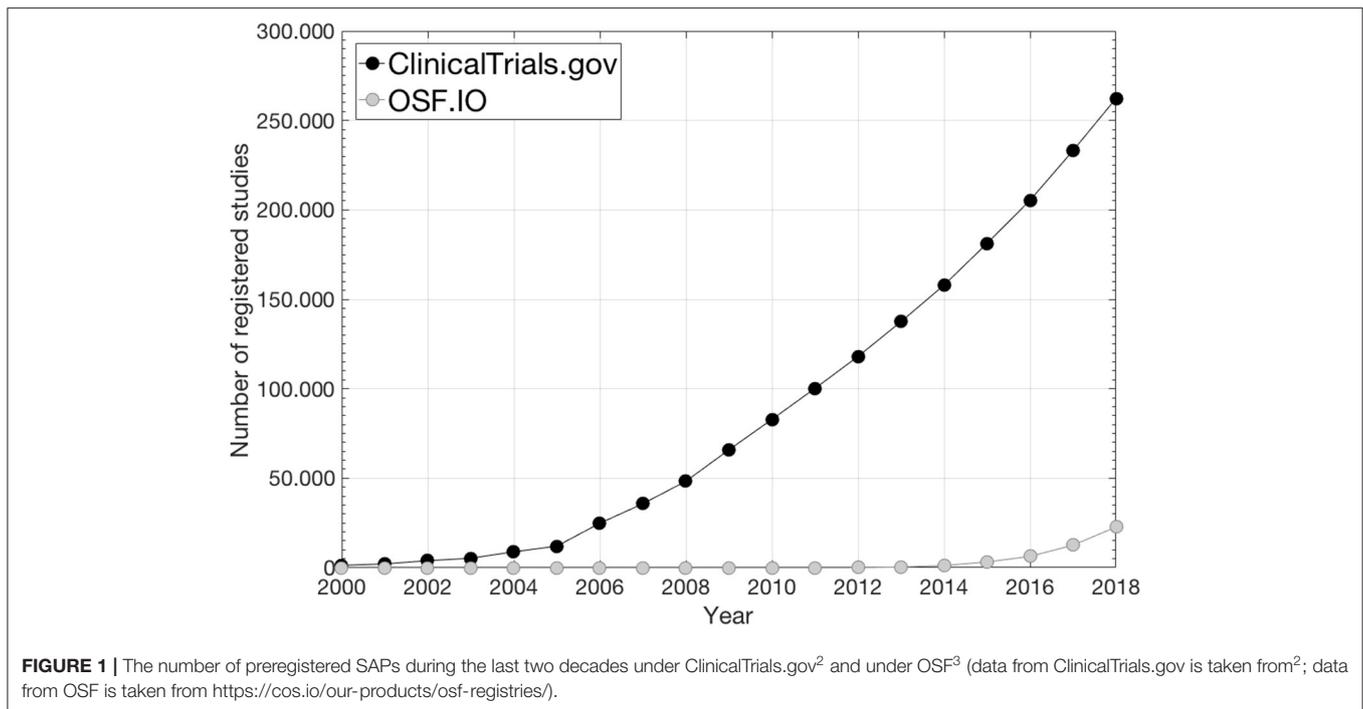
Keywords: cancer, clinical trial, observational study, outcome modeling, preregistration, public repository, radiotherapy, study plan

INTRODUCTION

Starting from 1997, the Food and Drug Administration Modernization Act (FDAMA) mandated the National Institute of Health (NIH) to design a platform in which information about FDA regulated clinical trials would become publicly available¹. As a result, NIH launched ClinicalTrials.gov shortly thereafter (1). Public pre-registration of clinical trials has since become a general publication requirement (2), and fast forwarded to two decades after FDAMA was introduced, ClinicalTrials.gov hosts 341 988 (as of June 11, 2020) registered studies conducted worldwide².

¹Food and Drug Administration Modernization Act of 1997: <https://www.govinfo.gov/content/pkg/PLAW-105publ115/pdf/PLAW-105publ115.pdf#page=16>

²<https://clinicaltrials.gov/ct2/resources/trends>



This site primarily focuses on interventional studies/clinical trials while study analysis plans (SAPs) and associated results from observational studies are scarce (1)² as also illustrated in **Figure 1** where the number of preregistrations from ClinicalTrials.gov and from the Open Science Foundation (OSF)³, which mainly holds SAPs from observational studies, is given over time. Consequently, for observational exploratory research it is often unclear as to the number of analyses undertaken, which further feeds into what is referred to as “*p-hacking*,” i.e., a positive finding publication bias since the vast majority of published studies that report *p*-values disclose positive/significant findings (3, 4). Further, SAP pre-registration is likely to facilitate researchers to better distinguish between confirmatory research (hypothesis-testing in which *p*-values retain diagnostic value) and exploratory research (hypothesis-generating in which *p*-values loose diagnostic value) in order to avoid overconfidence in *post-hoc* explanations in a finding that has not been proven, which could limit reproducibility (5).

The Transparent Reporting of a multivariate prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement has encouraged to better standardize outcome modeling (6). Outcome modeling can be based on data generated from clinical trials or observational studies. Here we propose to pre-register SAPs under public repositories for any outcome modeling study to further promote standardization, transparency and exactness and to mitigate the false positive inflation of published results.

³<https://osf.io/>

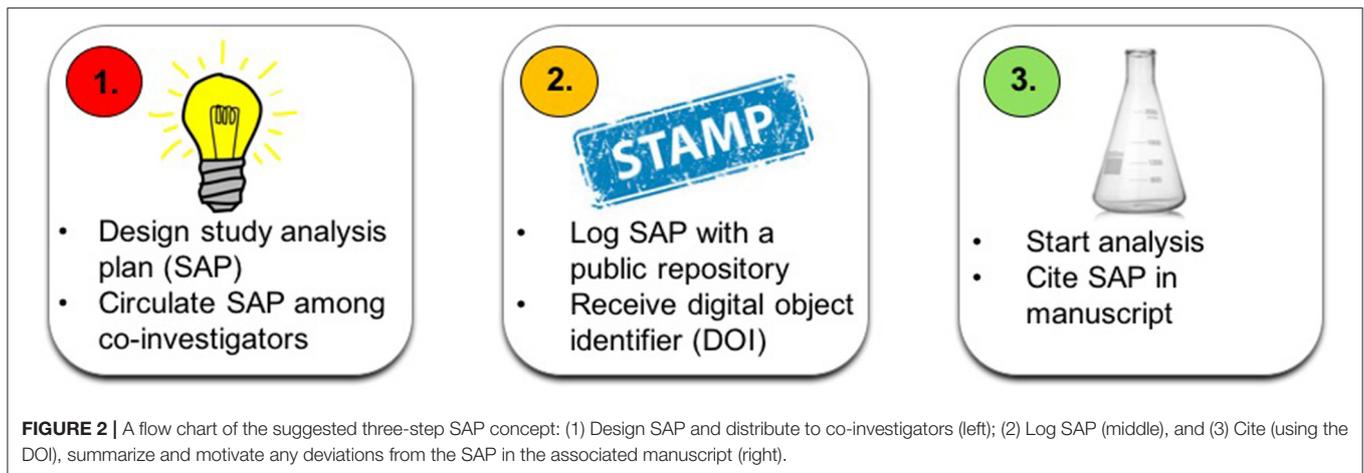
METHODS AND MATERIALS

Public pre-registration of SAPs could be thought of as committing to an analytical path but without advancing knowledge of the research outcome (4). To date, the two most commonly used public SAP repositories, which both provide SAP unique digital object identifiers (DOIs), are located under ClinicalTrials.gov (1) and under the OSF³. As previously pointed out, ClinicalTrials.gov has primarily been used to register clinical trials, while under OSF a larger extent of SAPs from observational studies can be found.

THE SAP CONCEPT

The suggested SAP concept consists of three steps: (1) Designing the SAP and circulating it among the co-investigators; (2) Logging the SAP with a public repository, which recognizes the SAP with a DOI, and (3) Citing (using the DOI), briefly summarizing and motivating any deviations from the SAP in the associated manuscript (*Note: any new major post-SAP analysis should only be considered hypothesis-generating/exploratory*). The three-step SAP concept is summarized in **Figure 2**.

The outcome modeling pipeline in the SAP should adhere to the modeling procedures defined in the TRIPOD landmark paper on how to model outcomes (6). This refers to description of data, outcomes and input variables and parameterization in addition to detailed step-by-step lay-out of the analysis. Below we list more specifically what the SAP should include (at a minimum) inspired by OSF’s preregistration template, which is available as a GoogleDoc here: <https://docs.google.com/document/d/1DaNmJEtBy04bq1l5OxS4JAscDZEKUGATURWwnBKLYxk/>



edit?pli=1, but more directed toward outcome modeling assuming an observational study design. An associated example SAP template for the purpose of outcome modeling is provided in the **Supplementary Material**.

Scope (Description of Data and Study)

The study scope should include title, co-investigators, and a brief study description, and the underlying study hypothesis/hypotheses. The brief study description should be accompanied by a description of data/patient population (inclusion criteria, number of patients, primary tumor site, treatment era, etc.) and primary outcome measure with range and minimum follow-up time and censoring defined if applicable. The study type should also be clearly stated (e.g., validation, exploration, and/or prediction).

Analysis (Description of the Analysis)

All variables considered for analysis should be described in detail along with their parameterization (binary, categorical and continuous; specify increments if applicable). Handling of missing data (if excluding data then describe how this will be accounted for) should be disclosed, and if applicable data transformation (or normalization) as well as definition of variable interaction terms should be given. The exact definition of the studied outcome, e.g., timing and scoring of radiation-induced toxicity and how pre-treatment status was taken into account, should be given. Although this SAP concept work focuses on outcome modeling in general, the expected minimum level of detail on reported variables is exemplified for RT dose, which is central for outcome modeling following RT: Specify if dose was parameterized as 2D dose-volume histograms (denote metrics, interval investigated and sampling), and/or summary measures such as the mean dose or the generalized equivalent uniform dose, and/or if being represented spatially (denote metrics and describe method) and if dose originated from the planned dose distribution, if being accumulated (plus type of dose accumulation), and/or if during treatment dose was used (applies possibly only for acute toxicity). Denote if and how fractionation effects were handled, and give the exact anatomical

definition of the investigated organ(s) along with the associated segmentation approach. Please refer to the Results section for a practical example of the level of detail in describing dose.

The statistical functions/methods of analysis (e.g., regression (and type), time to event, competing risk, etc.) should be explained in detail along with a description of risk groups and defined errors/confidence intervals (if valid), any considered resampling (e.g., iterated cross-validation hold-out or bootstrapping; number of iterations, etc.), validation (external/internal), and if and how univariate and/or multivariate analysis will be performed. Any considered level of significance/model quality should be specified and the associated performance metric described. If investigating more than one variable authors should denote how multiple testing will be corrected for.

Lastly, the SAP should include the statistical software tools (and version) that are being considered.

LITERATURE RESOURCES FOR OUTCOME MODELING IN RADIOTHERAPY

Aside from advocating the use of TRIPOD (6) as a guideline for outcome modeling in general, we below provide a short introduction to relevant literature for outcome modeling in radiotherapy (RT) with a particular emphasis on standardization.

To obtain reliable information about toxicities that influence patient's quality of life, normal tissue toxicities are likely best represented by patient-reported outcomes (PROs) (7). Using clinical decision-support tools (8, 9) and keeping the number of items/questions as few as possible (10) are necessary for actionability to patient-reported complaints. Dose-volume histogram (DVH) metrics of interest depend on a large variety of factors as pointed out within the 21 papers by the QUANTEC effort (11). Gathering published DVH metrics to better understand the reliability and generalizability of such metrics was first initiated by QUANTEC and is, as illustrated in their offspring efforts [pediatric RT (12) and hypo fractionated RT (13)] and work by other groups (14), a continuous process. These and related efforts (ideally multi-institutional) in which

models have been validated (6, 15, 16) probably hold the most reliable DVH findings. Also, incorporation of additional sources of data is likely to shed much further light on the complex mechanisms of both tumor response and normal tissue toxicity following RT. Examples are shown in studies focusing on genome-wide assays (17) and immune status (18) as well as medical imaging within associated standardization efforts (19).

RESULTS

A SAP Pre-registration Example

The authors recent experience in depositing an outcome modeling SAP with a public repository (15) will be used as an example of the SAP pipeline and content for outcome modeling.

After circulating the SAP among co-investigators, the SAP was logged with the OSF on July 23rd 2018 (15), and the analysis was, thereafter, initiated [the associated full-length manuscript was recently accepted for publication (16)].

As stated under the study *scope* (15), data were generated from a clinical trial (20) but the trial was not part of the outcome modeling itself. The *primary outcome measure* was overall survival defined from the start date of randomization and right-censoring was applied if alive at the last follow-up. For *Input data* related to disease, patient and treatment characteristics (the latter included 2D DVH parameterizations of the atria, lung, pericardium, and ventricles [please see (16) for exact anatomical definitions and parameterization of the remaining input data] structures: the minimum dose to the hottest 5–95% volume in steps of 5%, mean dose, minimum dose, max dose and the mean of the hottest 5–100% volume in steps of 5%, all metrics were corrected for fractionation effects assuming $\alpha/\beta = 3$ Gy), significance was denoted at a 5% Bonferroni-corrected level. *Validation* was considered using a holdout subset on which performance would be assessed after settling the final model. The validation procedures were directly adopted from a previously published study (21).

REFERENCES

1. McCray AT. Better access to information about clinical trials. *Ann Intern Med.* (2000) 133:609–14. doi: 10.7326/0003-4819-133-8-200010170-00013
2. De Angelis C, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, et al. Clinical trials registration: a statement from the International committee of medical journal editors. *N Engl J Med.* (2004) 351:1250–1. doi: 10.1056/NEJMe048225
3. Ioannidis JPA. The proposal to lower p value thresholds to 0.05. *JAMA.* (2018) 319:1429–30. doi: 10.1001/jama.2018.1536
4. Kaplan RM, Irvin VL. Likelihood of null effects of large NHLBI clinical trials has increased over time. *PLoS ONE.* (2015) 10:e0132382. doi: 10.1371/journal.pone.0132382
5. Nosek BA, Ebersole CR, DeHaven AC, Mellor DT. The preregistration revolution. *PNAS.* (2017) 115:2600–06. doi: 10.1073/pnas.1708274114
6. Moons KG, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al. Transparent reporting of a multivariate prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med.* (2015) 162:1–73. doi: 10.7326/M14-0698
7. Yeung AR, Pugh SL, Klopp AH, Gil KM, Wenzel L, Westin SN, et al. Improvement in patient-reported outcomes

The main *statistical function* was Cox Proportional Hazard regression. Both univariate and multivariate analyses, with a clear advancement criterion ($p < 0.05$ of the log-likelihood statistics), were undertaken, and re-sampling was considered using Bootstrapping with 1,000 iterations. Lastly, arriving at the final model, two alternative approaches were explored—the $\geq 10\%$ most frequently selected multivariate models or the ensemble thereof.

CONCLUSION

We have suggested a SAP pre-registration pipeline to be used for outcome modeling studies, which typically use observational data. An example of an already submitted SAP and cited for outcome modeling is given along with an outcome modeling directed SAP template. The ambition of the authors is that pre-registration of SAPs, using the suggested layout and pipeline, is becoming standard, like it has for clinical trials, also in outcome modeling.

AUTHOR CONTRIBUTIONS

MT and JD contributed to the conception and design of the study. MT drafted the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This research was funded in part through the NIH/NCI Cancer Center Support Grant P30 CA008748.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2020.00978/full#supplementary-material>

- with intensity-modulated radiotherapy (RT) compared with standard RT: a report from the NRG Oncology RTOG 1203 study. *J Clin Oncol.* (2020) 38:1685–92. doi: 10.1200/JCO.19.02381
8. Salz T, Schnall RB, McCabe MS, Oeffinger KC, Corcoran S, Vickers AJ, et al. Incorporating multiple perspectives into the development of an electronic survivorship platform for head and neck cancer. *JCO Clin Cancer Inform.* (2018) 2: CCI.17.00105 doi: 10.1200/CCI.17.00105
9. Kildea J, Battista J, Cabral B, Hendren L, Herrera D, Hijal T, et al. Design and development of a person-centered patient portal using participatory stakeholder co-design. *J Med Internet Res.* (2019) 21:e11371 doi: 10.2196/11371
10. Young Afat DA, Gibbons C, Klassen AF, Vickers AJ, Cano SJ, Pusic AL. Introducing BREAST_Q computerized adaptive testing - short and individualized patient-reported outcome assessment following reconstructive breast surgery. *Last Reconstr Surg.* (2019) 143:679–84. doi: 10.1097/PRS.00000000000005314
11. Bentzen SM, Constine LS, Deasy JO, Eisbruch A, Jackson A, Marks LB, et al. Quantitative analyses of normal tissue effects in the clinic (QUANTEC): an

- introduction to the scientific issues. *Int J Radiat Oncol Biol Phys.* (2010) 76 (3 Suppl.):3–9. doi: 10.1016/j.ijrobp.2009.09.040
12. Constone LS, Ronckers CM, Hua CH, Olch A, Kremer LCM, Jackson A, et al. Pediatric normal tissue effects in the clinic (PENTEC): an international collaboration to analyze normal tissue radiation dose-dose-volume response relationships for paediatric cancer patients. *Clin Oncol.* (2019) 31:199–207. doi: 10.1016/j.clon.2019.01.002
 13. Ohri N, Tomé WA, Méndez Romero A, Miften M, Haken RKT, Dawson LA, et al. Local control after stereotactic body radiation therapy for liver tumors. *Int J Radiat Oncol Biol Phys.* (in press). doi: 10.1016/j.ijrobp.2017.12.288
 14. Olsson CE, Jackson A, Deasy JO, Thor M. A systematic post-quantec review of tolerance doses for late toxicity after prostate cancer radiation therapy. *Int J Radiat Oncol Biol Phys.* (2018) 102:1514–32. doi: 10.1016/j.ijrobp.2018.08.015
 15. Lee S, Kerns S, Ostrer H, Rosenstein B, Deasy JO, Oh JH. Machine learning on a genome-wide association study to predict late genitourinary toxicity after prostate radiation therapy. *Int J Radiat Oncol Biol Phys.* (2018) 101:128–35. doi: 10.1016/j.ijrobp.2018.01.054
 16. Thor M, Montovano M, Hotca A, Luo L, Jackson A, Wu AJ, et al. Are unsatisfactory outcomes after concurrent chemoradiotherapy for locally advanced non-small cell lung cancer due to treatment-related immunosuppression? *Radiother Oncol.* (2020) 143:51–7. doi: 10.1016/j.radonc.2019.07.016
 17. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: standardized quantitative radionics for high-throughput image-based phenotyping. *Radiology.* (2020) 295:328–38. doi: 10.1148/radiol.2020191145
 18. Thor M, Deasy JO. The role of heart-related dose-volume metrics on overall survival in the RTOG 0617 clinical trial. *OSF.* (2018). doi: 10.17605/OSF.IO/HZSVA
 19. Thor M, Deasy JO, Hu C, Gore E, Bar-Ad V, Robinson C, et al. Modeling the impact of cardio-pulmonary irradiation on overall survival in NRG Oncology trial RTOG 0617. *Clin Can Res.* doi: 10.1158/1078-0432.CCR-19-2627
 20. Bradley JD, Paulus R, Komaki R, Masters G, Blumenschein G, Schildet S, et al. Standard-dose versus high-dose conformal radiotherapy with concurrent and consolidation carboplatin plus paclitaxel with or without cetuximab for patients with stage IIIA or IIIB non-small cell lung cancer (RTOG 0617): a randomized, two-by-two factorial phase 2 study. *Lancet Oncol.* (2015) 16:187–99. doi: 10.1016/S1470-2045(14)71207-0
 21. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. *BMC Med Res Methodol.* (2013) 33:1–15. doi: 10.1186/1471-2288-13-33

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Thor, Oh, Apte and Deasy. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.