This is a "preproof" accepted article for *Journal of Clinical and Translational Science*.
This version may be subject to change during the production process.
10.1017/cts.2025.10212

Structured Data Management and Sharing Plan (DMSP) Templates Outperformed Non-Structured Ones in an Institutional Implementation of the NIH Data Management and Sharing (DMS) Policy

Muayad M Hamidi, MBChB[1]; Manju Bikkanuri, MBBS[1]; Camille Scott, MPH[1]; Monica Carrizal[1], Mari Martinez[1]; Andrea N. Schorr, MS[1]; Liu Qianqian, MS[1]; Jonathan Gelfond, MD, PhD[1]; Joseph Schmelz, PhD[1]; Jennifer Potter, PhD, MPH[1]; Meredith Zozus, PhD[1]

[1]University of Texas at San Antonio (UTSA)

Corresponding Author: Muayad M Hamidi, MBChB, Joe R. and Teresa Lozano Long School of Medicine, University of Texas at San Antonio, 7703 Floyd Curl Drive, San Antonio, TX 78229 hamidim@uthscsa.edu 210-567-0841

Abstract

Introduction: The National Institutes of Health Data Management and Sharing (DMS) policy (NOT-OD-21-013) mandates the submission of a Data Management and Sharing Plan (DMSP) for all NIH-funded research that generates scientific data. However, little information is available about how academic medical centers have implemented the policy.

Objectives: The study aimed to characterize our institution's implementation of the DMS policy and compare structured versus unstructured approaches to producing policy-conformant DMSPs.

Methods: We monitored all NIH grant submissions from our institution for 18 months, evaluating policy implementation through DMSP completeness and reviewer comments during the Just-in-Time period. A rubric was developed to assess whether each required DMSP element and sub-element was addressed. Eight DMSP templates (three NIH-provided, five institutionally developed) and two categories of investigator-created DMSPs were scored. Researchers' feedback was collected through surveys and interviews.

Results: 79.3% of submitted DMSPs addressed all NIH-required DMSP elements. Element-level compliance ranged from 98.9% (data type) to 82.7% (tools and software). Sub-element scores showed greater variability, with 98.9% completion for data description and 49.3% for data generation. Unstructured DMSPs consistently underperformed compared to structured DMSPs. Survey and interview feedback, along with reviewer comments, reinforced these findings.

Conclusion: A notable 20.7% of DMSPs omitted one or more required elements, indicating a need for improved DMS policy conformance. Structured DMSP templates demonstrated greater alignment with NIH policy. We recommend using structured templates to enhance the quality and consistency of data management and sharing plans.

https://doi.org/10.1017/cts.2025.10212 Published online by Cambridge University Press

Introduction

The National Institutes of Health[1] Data Management and Sharing (DMS) policy (NOT-OD-21-013, effective on January 25, 2023) mandates that a Data Management and Sharing Plan (DMSP) be submitted as a component of grant applications for studies that generate scientific data.[1] Two decades prior, and part of an international groundswell of support for open sharing of resources and results from publicly funded research, the NIH released its initial policy in 2003, requiring a data sharing plan for certain grants. The National Science Foundation (NSF) subsequently revised its policy on Dissemination and Sharing of Research Results, requiring submission of a two-page Data Management Plan starting January 18, 2011. The following years have seen a sharp increase in publications on data management and sharing, many from the library and information science discipline. Examples include reviews[2-4], inventory or evaluation of funder requirements for DMPs[5-7], inventory or assessment of tools for DMP creation[8,9], general guidance for DMP creation[5,10-14], discipline-specific guidance for DMP creation[6,15,16], case studies of DMP creation or implementation[17,18], characterizations of DMP implementations[19], descriptions of programs to help researchers with DMP creation[20-22], policy analysis or recommendations[18,23,24], evaluation of DMP content[3,23,25-33], development of rubrics to assess DMP content[29,34], evaluation of DMSP effectiveness resulting in well-shared data[23], advances toward machine-actionable DMPs[35-37], and a public registry for DMPs.[38]

As a new policy, little has been published regarding the institutional Implementation of the NIH DMS Policy and its requirements for DMSPs. This study aimed to characterize our institution's implementation of NIH DMS policy and evaluate structured vs. unstructured approaches to producing policy-conformant DMSPs. The NIH posted a general Microsoft Word DMSP template with headers and brief instructions for each header, along with policy information.[39] In March 2023, a Federal Demonstration Project (FDP)[40] was initiated to test two draft NIH versions of DMSP templates (*alpha* and *bravo*). At the same time, our institution initiated a support program to help our researchers comply with the new mandate. The support program produced five study-specific DMSP templates to support policy compliance. The templates averaged five pages in length and followed NIH grant font and margin requirements. They were structured with a section header for each of the six required elements of the NIH

https://doi.org/10.1017/cts.2025.10212 Published online by Cambridge University Press

DMSP. They provided additional resources, such as links to supplemental institutional and federal policy information. Example text for each DMSP required element was also provided to guide investigators in describing pertinent aspects of their research under the appropriate DMSP section. In this study, we obtained copies of DMSPs submitted with grant applications from our grant management system over an 18-month period. We tracked data related to Just-in-Time (JIT) comments. We created a rubric to assess DMSPs' compliance with the NIH DMS policy requirements. We surveyed and interviewed investigators who had requested support for preparing their DMSPs. We asked them about usability, usefulness, clarity, ease of understanding, fit to study design, and overall satisfaction with the template they used for their grant applications. The results revealed that unstructured DMSPs omitted significantly more information than the structured ones. Unstructured DMSPs were used as a natural comparison group to assess the relative performance of local and NIH structured templates. The researcher's feedback overwhelmingly supported the use of structured templates with detailed instructions. The data related JIT comments were significantly less in submissions using a structured DMSP template. We recommend using structured DMSP templates that guide researchers through each required element and provide example text to support completion. We also recommend the use of a standardized rubric to help researchers and NIH program officials consistently prepare and evaluate DMSPs. Together, these measures could significantly improve compliance, reduce ambiguity, and promote higher-quality data management planning across NIH-funded research.

Beyond the ethical considerations for the public availability of publicly funded research, data sharing may increase the return on investment from such research, for example, by enabling subsequent use of the data to answer new research questions. Thus, the NIH has heavily invested in data repositories to ensure enduring access to these resources. Following the seminal articulation of FAIR principles by Wilkins *et al*. 2016[41], significant efforts have been undertaken to pursue methods and technology to make shared data and resources Findable, Accessible, Interoperable, and Reusable (FAIR)[41,] such as the Biomedical Data Fabric (BDF) initiative[42]. Although data sharing in the U.S. is far from achieving the FAIR ideal[43], the NIH DMS policy takes an essential step toward the reuse of research data in clinical translational science by addressing the upfront application of data standards and data management methods.

https://doi.org/10.1017/cts.2025.10212 Published online by Cambridge University Press

Methods

Local DMSP Templates: We designed the local DMSP templates to reduce cognitive load through their structure, specifically the external representation and the Proximity Compatibility Principle (PCP). Briefly, external representation refers to having the information needed for a task in the world, such as physical constraints, or, in the case of our DMSP templates, instructions and links to relevant information, rather than storing it in memory. The PCP in cognitive engineering states that information required by a task (related information) should be placed physically close to (proximally) or ideally at the location of task performance so that users can more easily integrate the information[44]. For example, adhering to the PCP for DMSPs obviates the need for investigators to conduct web searches for required policy information or requirements.

Data Collection: We obtained copies of Data Management and Sharing Plans (DMSPs) submitted with grant applications from our institutional grant management system for the period from January 1, 2023, to May 23, 2024. Similar to the DART Framework (Data Management Plans as a Research Tool), which is based on NSF data management plan requirements[29], and the DMP scorecard developed for the Belmont Forum grants[34], we created a standardized rubric to assess adherence to the DMSP with the NIH policy (Supplementary Material 1). The rubric assessed the six main DMSP elements (Tables 1 and 2 of the Rubric in Supplementary Material 1) using nominal binary (Yes/No) items. The rubric also assessed the DMSP sub-elements (Tables 3 of the rubric in Supplementary Material 1) using fourteen nominal binary items and seven nominal non-binary items to score each type of data listed in the DMSP. Because the DMS policy requires similar information for all kinds of scientific data generated by a grant, the assessment rubric was designed to be expandable, where a new section could be added for each type of data addressed in the DMSP (Supplementary Material 1). Two independent evaluators examined and scored each DMSP. The reviewers in regular study meetings adjudicated the discrepant assessments. Scores were entered in a scoring worksheet to calculate the adherence/conformance rate and to categorize the results.

Survey and Interview: We surveyed and interviewed investigators who had requested support through the DMSP Assistance Program. We expanded the original 14-item questionnaire

https://doi.org/10.1017/cts.2025.10212 Published online by Cambridge University Press

used in the NIH FDP pilot–comprising 10 quantitative and 4 qualitative items–for local use by adding 8 additional quantitative and 11 qualitative items (Supplementary Material 2). These supplementary questions were included to explore researchers' perspectives on the template they used, with a specific focus on usability, usefulness, clarity, ease of understanding, fit with the study design, and overall satisfaction. Responses were categorized based on the type of template used, dividing them into three groups: NIH templates, local templates, and investigator-created DMSPs. The survey data were collected and managed using the REDCap electronic data capture tool hosted at the University of Texas Health Science Center in San Antonio (UTHSA).[45,46] Survey data were collected two weeks after the investigator submitted their funding application. Non-respondents received up to three follow-up contact attempts via email. Most of the additional qualitative questions focused on identifying opportunities to improve the local DMSP assistance program. Responses to qualitative items were reviewed and thematically categorized for reporting.

Just-In-Time (JIT) Review: We also waited for the JIT comments from the grant review process. These were used as an additional source of quality control feedback to assess policy adherence.

Statistical analysis: The proportion of completed DMSP elements for each template type was calculated with exact 95% confidence intervals and compared using Fisher's exact test. Each required element in a DMSP could have multiple sub-elements corresponding to the different data types. These sub-elements were analyzed using Generalized Estimating Equations, with each data type clustered within a DMSP. For meaningful statistical analysis, we grouped the templates into three categories: NIH group templates, local group templates, or idiosyncratic DMSPs. The latter consisted of DMSPs created by investigators combining multiple templates, as well as those with no resemblance to any of the assessed templates. Within this model, the completeness rate of the sub-elements was predicted by a categorical variable representing the three DMSP template categories. No adjustments were made for multiple comparisons.

https://doi.org/10.1017/cts.2025.10212 Published online by Cambridge University Press

Results

During the evaluation period, from January 1, 2023, to May 23, 2024, 27% of the funding applications submitted by our institution did not require a DMSP. These grant applications were submitted to NIH for activities that do not generate scientific data, such as training, infrastructure development, or other non-research-focused efforts, or were submitted to funding agencies other than NIH. The remaining 73% of applications fell within the scope of the DMS Policy and therefore required a DMSP. Of the applications requiring a DMSP, 84% (n=358) included one. For 16.2% (n=69) of the submissions requiring a DMSP, no plan was retrievable from our grant management system (Figure 1).

We presumed that these DMSPs were either not submitted at the time of application or were submitted through alternative channels that bypassed our standard institutional tracking system, resulting in their omission from the internal record. We evaluated the content of the 358 DMSPs using the rubric developed for this study (Supplementary Material 1); 79.3% of them addressed all six required elements outlined in the NIH DMS Policy. The percentage of DMSP addressing required DMSP elements varied. Element 1 (Data description) was most frequently addressed (98.9%), while Element 2 (Related Tools, Software, or Code) was the least frequently addressed (82.7%). Similarly, the percent of sub-elements addressed varied from 98.9% for data type description to 49.3% for data generation process description (Table 1).

The study designs represented in the assessed DMSPs were categorized into five groups: clinical interventional studies, clinical observational studies, basic science studies, secondary data or sample use, and other. Overall, the proportion of DMSP elements and sub-elements addressed was generally consistent across these categories (Figure 2). The most notable deviation occurred in the "Other" category, which included miscellaneous study designs not clearly assignable to any primary groups. Excluding this outlier, the percentage of elements addressed by each study category ranged from 76.3% to 92.9% and the percentage of sub-elements addressed per each study type ranged from 67.0% to 81.7% across the remaining study types (Figure 2). Data aggregation levels and descriptions of data processing were more frequently reported in clinical studies than in basic science studies. In contrast, the element

https://doi.org/10.1017/cts.2025.10212 Published online by Cambridge University Press

related to data generation was less frequently addressed in studies involving secondary data than in prospective clinical or basic science studies.

Types of templates used varied considerably; the most frequently used template was the NIH Generic template (n= 178), followed by the "Other" category (n= 71), the local animal studies template (n= 62), and the local prospective clinical study template (n= 30). In contrast, six of the ten assessed template types—the NIH *Alpha* and *Beta* templates, the human tissue only template, the non-animal/non-human studies template, the secondary data use template, and the mixed template—were used too infrequently to allow for meaningful statistical comparison, with usage frequencies of 1, 1, 3, 1, 4, and 7, respectively. To enable meaningful analysis, we grouped the templates into three broader categories: NIH, Local, and Idiosyncratic. The NIH group includes the *Alpha, Beta*, and Generic NIH templates (n= 180); the local group consists of the five institutionally developed templates (n= 100); and the idiosyncratic group comprises both the "Mixed" and "Other" template categories (n= 78), which reflect unstructured researcher-created formats. We calculated and compared the proportion of completed elements for each template type with exact 95% confidence intervals using Fisher's exact test. (Table 2).

Each required element in a DMSP could have multiple sub-elements for each type of data generated. The completion of these sub-elements was analyzed using Generalized Estimating Equations, with each data type clustered within a DMSP (Supplementary Material 3). Within this model, the rate of completeness of the sub-elements was predicted by a categorical variable representing the three DMSP template categories: NIH group template, local group template, or idiosyncratic DMSPs. The significance level (two-sided) was 0.05 for all tests. There were no adjustments for multiple comparisons (Supplementary Material 3). The results showed that the odds of addressing DMSP sub-elements increase significantly using structured templates. Compared to NIH DMSP templates, idiosyncratic (unstructured) templates underperformed, while local (structured) templates performed superiorly in addressing all sub-elements (Figure 3).

Researcher Survey and Interview results: There were 75 researchers that contacted the DMSP assistance program and accessed the DMSP templates via the provided web link. Of these, 43% (n=32) responded to the survey, and 27% (n=20) of those survey respondents also participated in follow-up interviews. In this small sample, we grouped the interview comments

https://doi.org/10.1017/cts.2025.10212 Published online by Cambridge University Press

from respondents completing a local template together and those completing an NIH template together. Locally developed templates were found to be favorable across all evaluated categories, including usability, clarity, alignment with the study design, and overall user satisfaction. Most respondents (87%, n=65), regardless of the template used, agreed that research trainees should receive formal data management and sharing training. All participants expressed that any template should be accompanied by instruction sheets tailored to common study types.

Regarding specific template features: 71% found the inclusion of example text within templates helpful, 57% indicated that links to NIH discipline-specific and general data repositories were beneficial, and 50% reported that including information and links related to genomic data sharing requirements was helpful. Researchers selected templates primarily based on how well they aligned with their study design and research aims. For instance, users of the NIH Generic Template cited reasons such as "followed NIH links" and its suitability for "basic science." The NIH Pilot *Alpha* Template was chosen because it "fits better with our research type and scopes," the NIH Pilot *Bravo* template user noted a desire to "try the NIH V2 template." Local templates were favored for their close alignment with specific project needs. Users described them as "most applicable," "discipline specific," "aligned with the type of data we are producing," and "very convenient for my study specifics." Others emphasized that these templates were "more project specific" and "most closely aligned with the proposed research plan aims". Other template selections seemed to be based on familiarity or adaptability. Comments included: "template used by our lab before," "simple, clear, and understandable," and "had to be customized to fit the study needs regarding data management and sharing.".

JIT comments related to data management and sharing: Out of 427 funding applications submitted, 19 (4.4%) received data-related comments or questions as part of the Just-In-Time (JIT) request process. The unstructured templates accounted for the highest number of JIT comments, 52.6% (n=10), followed by NIH templates, 42.1% (n=8). Notably, only one data-related JIT request was associated with a submission using a local template, suggesting potential advantages in clarity or completeness for those plans.

https://doi.org/10.1017/cts.2025.10212 Published online by Cambridge University Press

Discussion

Of the 358 Data Management and Sharing Plans (DMSPs) evaluated, 79% addressed all required sections of the NIH DMS policy, while 21% omitted one or more required sections. In our survey results, the majority of respondents indicated that data management training was desirable. Others continue to find lack of knowledge about research data management or DMP creation among researchers[8,29,33] similarly suggesting that institutional support programs should include training in DMSP creation if not in basic data management principles.

The proportion of DMSPs received data-related Just-In-Time (JIT) comments was significantly lower than the 21% (4.4%), suggesting that incomplete plans did not always trigger follow-up inquiries. There are several possible explanations. Omissions may have gone undetected during NIH program officials' reviews, may have been deemed inconsequential to the proposed study, or may be attributable to other factors that could not be determined from the available data. These findings suggest room for improvement in the completeness and review of DMSPs. Both investigators preparing DMSPs and NIH program officials reviewing them could benefit from a standardized assessment rubric to promote consistency, completeness, and adherence to policy requirements.

Seven sub-elements were less than 75% complete, meaning that at least 25% of DMSPs failed to address these components (Table 1). While omitting a sub-element does not necessarily indicate that the corresponding data-related tasks will not be performed by the investigator or study team, it does prevent NIH officials from evaluating the adequacy of the plan for completing them. If missing information is not identified and requested upon review, opportunities to intervene in potentially inadequate plans may be lost. Timely intervention is particularly concerning for data generation, collection, and handling aspects that cannot be easily corrected after the fact.[47] To mitigate this risk, we recommend that DMSP authors use a standardized template and rubric to ensure completeness and a standardized way to explicitly indicate that a required element or sub-element does not apply to a study. NIH program officials could similarly use a rubric during review to consistently evaluate conformance. Completeness—addressing each required element and sub-element—is critical to compliance with the NIH DMS policy. A complete DMSP enhances transparency and fosters early and meaningful discussions

https://doi.org/10.1017/cts.2025.10212 Published online by Cambridge University Press

about data management and sharing. These early and meaningful discussions can improve planning, implementation, and ultimately, the overall integrity and utility of the research. The completeness of DMSPs, and by extension, compliance with the NIH DMS policy, varied significantly across the most used templates. The local prospective clinical study template (n=30) had the highest completeness rate at 92.5%, followed by local animal study templates (n=62) at 84.5%, the NIH Generic template (n=178) at 75.3%, and "Other" templates (n=71) at 42.6%. These variations underscore the significance of template structure in facilitating compliance with NIH requirements. While the "Other" template category stood out as distinctly different in having little to no structure, the structured templates provided more specific guidance with respect to section and sub-section content (the six elements and sub-elements), leading to higher completeness rates (Table 2). The survey and interview data supported these findings with majority of respondents indicating that structural elements of the templates were desirable. Others also advocate domain-specific DMP templates or tools[8,33]

Interestingly, unstructured DMSPs commonly included data and resource sharing elements consistent with prior NIH resource sharing policies, but often omitted those newly required by the NIH DMS policy. This pattern suggests that researchers were generally aware of and responsive to prior sharing expectations, rather than aware of the current policy. Therefore, while researcher-level factors such as awareness and institutional support undoubtedly influence template selection and completeness, our findings also suggest that unstructured templates–lacking prompt for policy-required information–may have also contributed to missing information in DMSPs. The NIH Generic template offered semi-structure, with headers for the six required DMSP sections but no additional framework. This template performed well at the element level (98.9% completeness) but less at the sub-element level (75.4% completeness), reinforcing that greater structure within a template improves overall completeness.

Limitations and Generalizability: A randomized controlled trial (RCT) was neither feasible nor appropriate due to the nature of the policy implementation and the real-world constraints. The observational nature of this study limits the strength of evidence. Additionally, the implementation environment itself was evolving throughout the study. Additional challenges included barriers and the time required to gain access to the institutional grant management

https://doi.org/10.1017/cts.2025.10212 Published online by Cambridge University Press

system and Just-In-Time (JIT) information. These results, from a case study conducted at a single institution, are not directly generalizable to other institutions or settings. However, we believe our context may share significant similarities with other institutions, and thus, some of our findings may be applicable or informative to others in similar settings.

Conclusion

In our case study, the current adherence rate with the NIH Data Management and Sharing (DMS) policy elements was just over 79%, meaning that approximately 21% of submitted DMSPs missed one or more required elements. This gap highlights the opportunity for targeted interventions to enhance policy adherence, specifically in terms of data management plan completeness. Our findings indicate that structured DMSPs outperform non-structured ones regarding completeness and alignment with policy expectations (Table 2). To address these challenges, we recommend a two-pronged approach: use of structured DMSP templates that guide researchers through each required element and sub-element–with example text to clarify expectations– and implementation of a standardized rubric to assist researchers, institutional officials, and NIH program officials in consistently preparing or evaluating DMSPs. Based on this case study, policy adherence can be significantly improved by adopting these measures, reducing ambiguity, and promoting higher-quality data management planning across NIH-funded research.

Financial Support

Funding source: This work is funded by the National Institute of Health (NIH), National Center for Advancing Translational Sciences (NCAT). The Clinical and Translational Science Award (CTSA) UL1 TR 002645-02 - Element E.

Role of sponsor or funder: The NIH-NCAT awarded The Clinical and Translational Science Award (CTSA) to the University of Texas San Antonio (UTSA) to pursue integrating clinical and translational research and career development across UTSA schools and diverse public and private partners in South Texas. The CTSA to UTSA provides a broad range of support for research infrastructure to reduce barriers to research and stimulate the transformation of knowledge into improved health care. The funder will have no input on the interpretation or publication of the study results.

Competing Interests

All authors report no conflicts of interest.

References

1. NIH. National Institutes of Health. Final NIH Policy for Data Management and Sharing. Bethesda, MD: US Department of Health and Human Services; 2023. Accessed October 31, 2025. https://grants.nih.gov/policy-and-compliance/policy-topics/sharing-policies/dms/policy-overview

2. Vitale C, Moulaison-Sandy H. Data Management Plans A Review. *DESIDOC Journal of Library & Information Technology*. 12/16 2019;39:322-328. doi:10.14429/djlit.39.06.15086

3. Smale N, Unsworth K, Denyer G, Barr D. The history, advocacy and efficacy of data management plans. *Int J Digit Curation.* 2020;15(1):30. doi:10.2218/ijdc.v15i1.525

4. Stanciu A. Data Management Plan for Healthcare: Following FAIR Principles and Addressing Cybersecurity Aspects. A Systematic Review using InstructGPT. *Romanian Cyber Security Journal*. 2023;5(1):23-43. doi:10.54851/v5i1y202303

5. Jennifer LT. Examination of federal data management plan guidelines. *J eSci Librariansh.* 2015;4(1). doi:10.7191/jeslib.2015.1072

6. Williams M, Bagwell J, Nahm Zozus M. Data management plans: the missing perspective. *J Biomed Inform.* 2017;71:130–142. doi:10.1016/j.jbi.2017.05.004

7. Henning P, Silva L, Ferreira Pires L, van Sinderen M, Moreira J. The FAIRness of data management plans: an assessment of some European DMPs. *Rev Eletron Comum Inf Inov Saúde.* 2021;15. doi:10.29397/reciis.v15i3.2270

8. Gajbe SB, Tiwari A, Gopalji, Singh RK. Evaluation and analysis of data management plan tools: a parametric approach. *Inf Process Manag.* 2021;58(3):102480. doi:10.1016/j.ipm.2020.102480

9. Becker C, Hundt C, Engelhardt C, Sperling J, Kurzweil M, Müller-Pfefferkorn R. *Data management plan tools: overview and evaluation.* Vol 1. 2023.

10. Michener WK. Ten simple rules for creating a good data management plan. *PLoS Comput Biol.* 2015;11(10):e1004525. doi:10.1371/journal.pcbi.1004525

https://doi.org/10.1017/cts.2025.10212 Published online by Cambridge University Press

11. Cauchick-Miguel PA, Moro SR, Rivera R, Amorim M. Data management plan in research: characteristics and development. In: *Springer International Publishing*; 2020:3–14.

12. Kung S. Developing a Data Management Plan. In: Carlson J, Choudhury S, eds. Demystifying the Data Management Plan: A Guide for the Research Community. Cambridge, MA: MIT Press; 2022:101-116. doi:10.7551/mitpress/12200.001.0001

13. Chhagan M, Kauchali S, van den Broeck J. The Data Management Plan. In: van den Broeck J, Brestoff JR, eds. Epidemiology: Principles and Practical Guidelines. Springer Netherlands; 2013:261-280

14. Zozus MN. The Data Book: Collection and Management of Research Data. Boca Raton, FL: CRC Press/Taylor & Francis Group; 2017:1-336

15. Brand S, Bartlett D, Farley M, et al. A model data management plan standard operating procedure: results from the DIA Clinical Data Management Community, Committee on Clinical Data Management Plan. Ther Innov Regul Sci. 2015;49(5):720-729. doi:10.1177/2168479015579520

16. Lebedys E, Famatiga-Fay C, Bhatkar P, Johnson D, Viswanathan G, Zozus MN. Data Management Plan. J Soc Clin Data Manag. 2021;1(4). doi:10.47912/jscdm.116

17. Burnette M, Williams S, Imker H. From plan to action: successful data management plan implementation in a multidisciplinary project. J eSci Librariansh. 2016;5(1):e1101. doi:10.7191/jeslib.2016.1101

18. Bishop B, Neish P, Kim J, et al. Data management plan implementation, assessments, and evaluations: Implications and recommendations. Data Sci J. 2023;22:1-27. doi:10.5334/dsj-2023-027

19. Spichtinger D. Data management plans in Horizon 2020: what beneficiaries think and what we can learn from their experience. Open Res Eur. 2021;1:42. doi:10.12688/openreseurope.13342.2

20. Mischo WH, Wiley CA, Schlembach MC, Imker HJ. An Integrated Data Management Plan Instructional Program. Paper presented at: 2017 ASEE Annual Conference & Exposition; June 2017; Columbus, OH. doi:10.18260/1-2--27572.

21. S B. The data management plan – the common thread through the project. DMP consulting at the University of Vienna. Proc Sci. 2021:014

22. Karimova Y. Research data description in multiple domains: supporting researchers with data management plans [doctoral thesis]. University of Porto; 2023

23. Van Tuyl S, Whitmire AL. Water, water, everywhere: defining and assessing data sharing in academia. PLoS One. 2016;11(2):e0147942. doi:10.1371/journal.pone.0147942

24. Pasek JE. Historical development and key issues of data management plan requirements for National Science Foundation grants: a review. Issues Sci Technol Librariansh. 2017;(87). doi:10.29173/istl1709

25. Parham SW, Doty C. NSF DMP content analysis: What are researchers saying? Bull Am Soc Inf Sci Technol. 2012;39(1):37-38. doi:10.1002/bult.2012.1720390113

26. Mischo WH, Schlembach MC, O'Donnell MN. An analysis of data management plans in University of Illinois National Science Foundation grant proposals. J eSci Librariansh. 2014;3(1):3. doi:10.7191/jeslib.2014.1060

27. Bishoff C, Johnston L. Approaches to data sharing: an analysis of NSF data management plans from a large research university. J Librariansh Scholarly Commun. 2015;3(2):eP1231. doi:10.7710/2162-3309.1231

28. Lalwani L, Carlson J, Samuel S, Grochowski P. Analyzing data management plans: where librarians can make a difference. Paper presented at: 122nd ASEE Annual Conference & Exposition; 2015

29. Parham SW, Carlson J, Hswe P, Westra B, Whitmire A. Using data management plans to explore variability in research data management practices across domains. Int J Digit Curation. 2016;11(1):53-67. doi:10.2218/ijdc.v11i1.423

30. Berman EA. An exploratory sequential mixed methods approach to understanding researchers' data management practices at UVM: findings from the quantitative phase. J eSci Librariansh. 2017;6(1):e1098. doi:10.7191/jeslib.2017.1098

31. Van Loon JE, Akers KG, Hudson C, Sarkozy A. Quality evaluation of data management plans at a research university. IFLA J. 2017;43(1):98-104. doi:10.1177/0340035216682041.

32. Mannheimer S. Toward a better data management plan: the impact of DMPs on grant funded research practices. J eSci Librariansh. 2018;7(3):e1155. doi:10.7191/jeslib.2018.1155

33. Veiga V, Ferreira Pires L, Henning P, Moreira J, Pereira I. Perception and practice of data management plans in health: an exploratory study. Data Intelligence. 2024;7(2):381-396. doi:10.3724/2096-7004.di.2024.0030

34. Bishop BW, Ungvari J, Gunderman H, Moulaison-Sandy H. Data management plan scorecard. Proc Assoc Inf Sci Technol. 2020;57(1):e325. doi:10.1002/pra2.325

35. Bakos A, Miksa T, Rauber A. Research data preservation using process engines and machine-actionable data management plans. In: 22nd International Conference on Theory and Practice of Digital Libraries (TPDL 2018), Porto, Portugal, September 10–13, 2018, Proceedings. 2018:69-80.

36. Stodden V, Ferrini V, Gabanyi M, Lehnert K, Morton J, Berman H. Open access to research artifacts: Implementing the next generation data management plan. Proc Assoc Inf Sci Technol. 2019;56:481-485. doi:10.1002/pra2.51

37. Romanos N, Kalogerini M, Koumoulos EP, Morozinis AK, Sebastiani M, Charitidis C. Innovative data management in advanced characterization: implications for materials design. Materials Today Communications. 2019;20:100541. doi:10.1016/j.mtcomm.2019.100541

38. Silva FCC, Monteiro E, Mundim Rodrigues M, Arellano M, Oliveira A. National registry of scientific data management plans by IBICT. 2022:29-38

39. NIH. Data Management and Sharing Plan Format Page.Accessed 8/27/2025, https://grants.nih.gov/grants-process/write-application/forms-directory/data-management-and-sharing-plan-format-page

40. FDP. NIH Data Management and Sharing Pilot. accessed 05/01/2025, https://thefdp.org/demonstrations-resources/nih-data-management-sharing-pilot/?highlight=template+DMSP

41. Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR guiding principles for scientific data management and stewardship. Sci Data. 2016;3:160018. doi:10.1038/sdata.2016.18

42. ARPA-H. ARPA-H announces effort to develop an integrated data toolbox improving access to biomedical research to bolster innovation. ARPA-H. May 1, 2025. Accessed October 31, 2025. https://arpa-h.gov/news-and-events/arpa-h-announces-effort-develop-single-data-system-biomedical-research

43. Duke J. U.S. Department of the Interior: Sharing FAIR Data Fairly. University of Arkansas at Little Rock; 2022. Accessed October 31, 2025. https://ualr.edu/eitcolloquium/spring-2023-colloquium-schedule/

44. Wickens CD, Carswell CM. The proximity compatibility principle: its psychological foundation and relevance to display design. Hum Factors. 1995;37(3):473-494. doi:10.1518/001872095779049408.

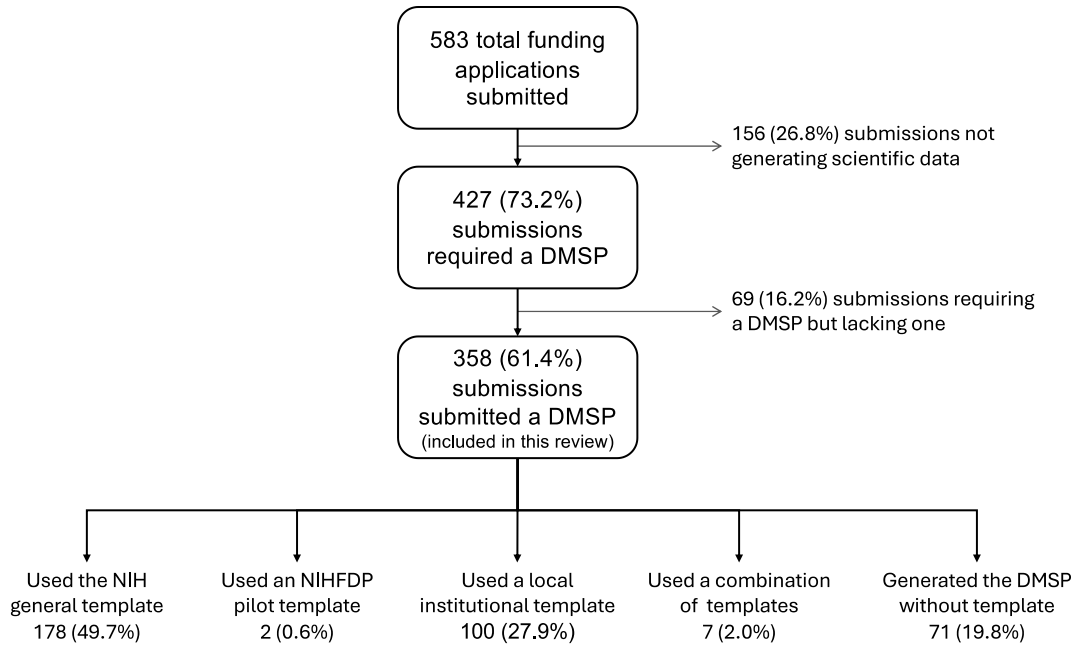
45. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. Apr 2009;42(2):377-81. doi:10.1016/j.jbi.2008.08.010

46. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform*. Jul 2019;95:103208. doi:10.1016/j.jbi.2019.103208

47. Zozus M. *The Data Book : Collection and Management of Research Data*. First edition. ed. CRC Press; 2017.

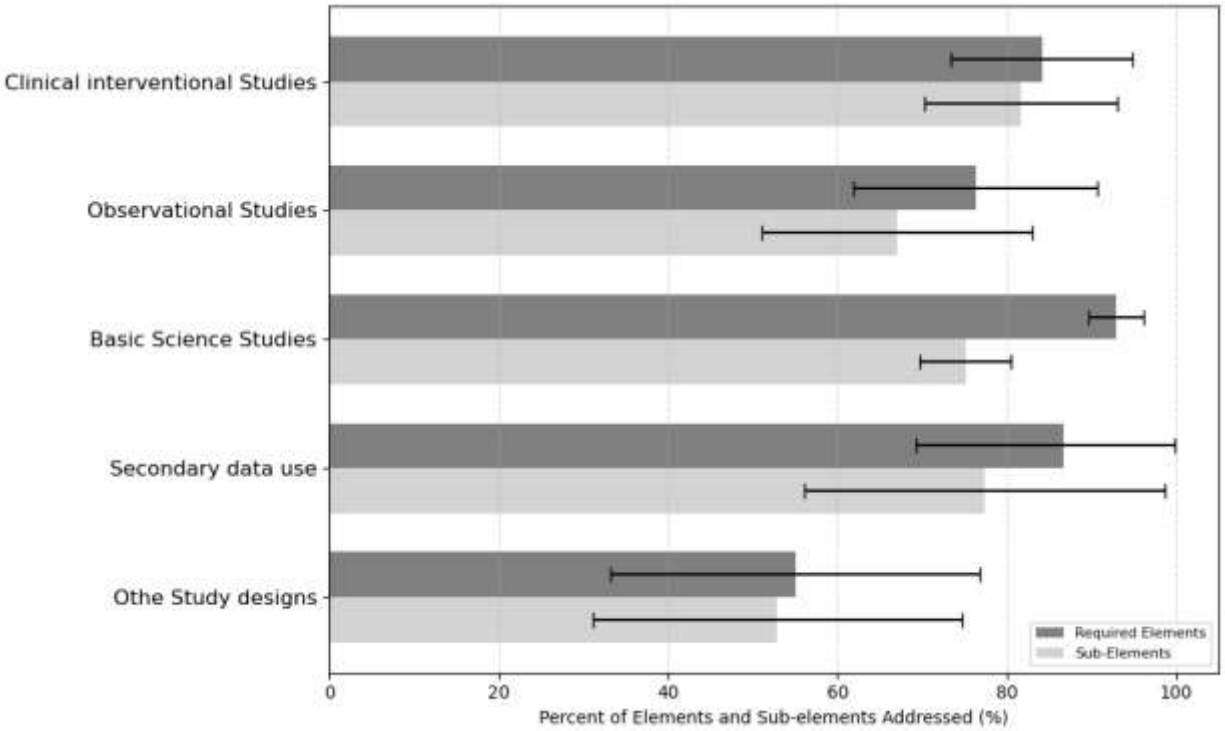
https://doi.org/10.1017/cts.2025.10212 Published online by Cambridge University Press

Figure 1 DMSPs Included in the Analysis



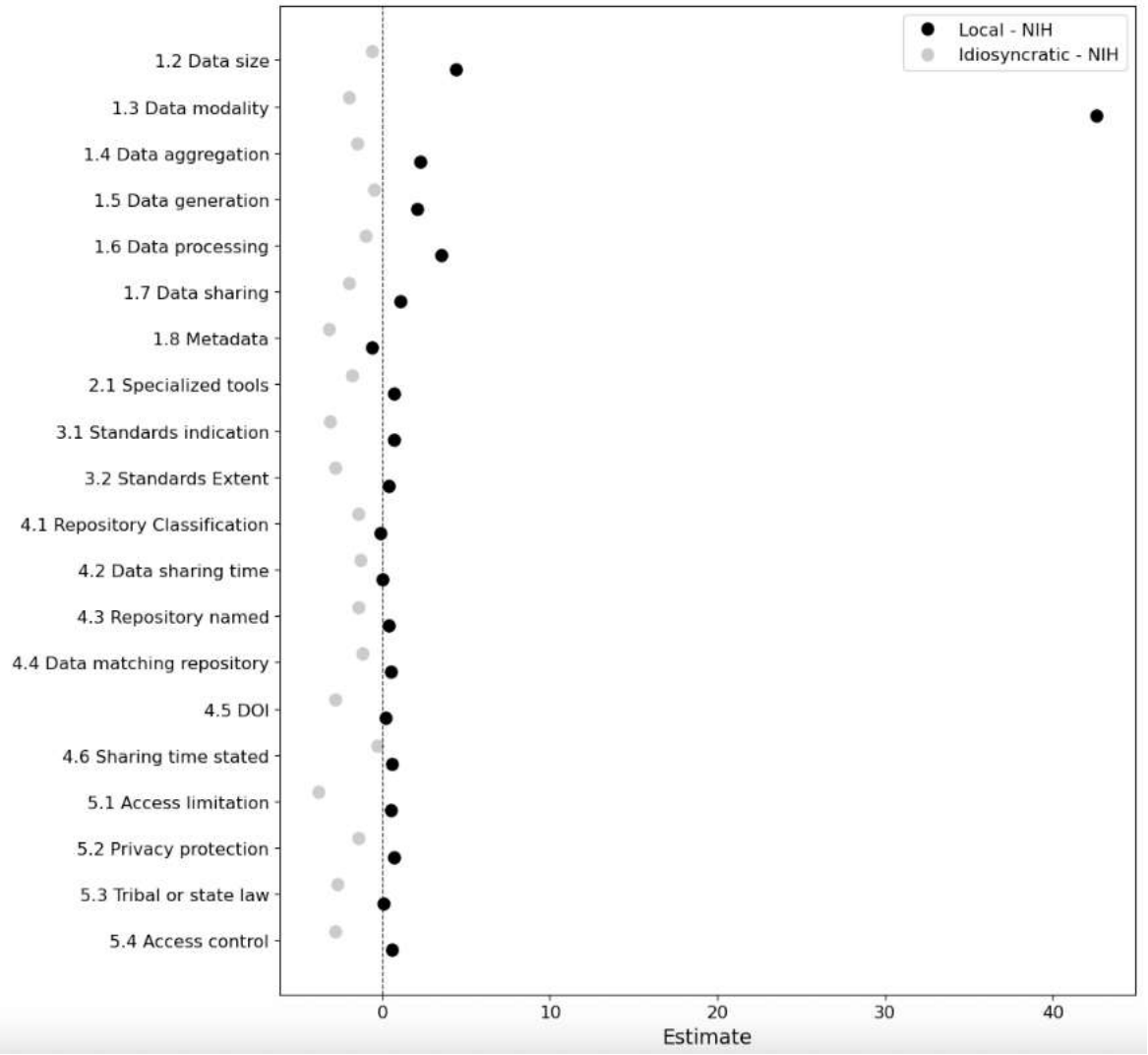
https://doi.org/10.1017/cts.2025.10212 Published online by Cambridge University Press

Figure 2 Percentage of DMSP Elements and Sub-Elements Addressed with 95% Confidence Intervals per Study Types



https://doi.org/10.1017/cts.2025.10212 Published online by Cambridge University Press

Figure 3 Estimating performance of local and idiosyncratic compared to NIH templates in addressing the Sub-Elements DMSP requirements using the Generalized Estimating Equations.



*Sub-element 1.1 is a count of the data types collected for a study. As a continuous variable (and as the denominator for the remainder of the sub-elements), it was excluded from the Generalized Estimating Equations analysis in this table.

https://doi.org/10.1017/cts.2025.10212 Published online by Cambridge University Press

Table 1 Addressed DMSP Required Elements and Sub-Elements

Required Elements	Addressed% (lower, upper 95%CI)	Sub-Elements	Addressed% (lower, upper 95%CI)
I. Data	98.9 (97.8,99.9)	1.1 Data type: What was the data type of each scientific data or group of data elements? Create a column for each.	98.9% (97.8,99.9)
		1.2 Data size: Did the DMSP provide an estimate of the amount or size of the scientific data to be generated and/or used? For example, the number of participants, data elements, time points, images, files, or the anticipated data volume (e.g., in KB, MB, GB, or TB). Yes, No	54.9% (49.7,60.0)
		1.3 Data modality: Was the data modality, such as imaging, genomic, mobile, survey, or data sources, described? Yes, No	90.0 (86.9,93.1)
		1.4 Data aggregation: Was the level of data aggregation (e.g., individual, aggregated, summarized) stated? Yes, No	65.6 (60.7,70.5)
		1.5 Data generation: Was each data group's data generation process stated? For example, if patient activity is measured, is it from minutes of exercise recorded in a daily diary or steps measured with a wearable accelerometer? Yes, No	49.3 (44.1,54.5)
		1.6 Data processing: Was the level of data processing –i.e., how raw or processed the data will be– Stated? Yes, No	61.2 (56.1,66.2)
		1.7 Data sharing: Did the DMSP include a statement indicating whether each type or group of scientific data will be preserved and shared? If specific data is not shared, did the plan provide a rationale for the decision? Yes, No	87.0 (83.6,90.5)
		1.8 Metadata Description: Did the DMSP list the metadata, other relevant data, and any associated documentation (e.g., study protocols, data dictionaries, data collection instruments) that will be made accessible to facilitate interpretation of the scientific data? Yes, No	87.0 (83.6,90.5)
II. Tools	82.7 (78.8,86.6)	2.1 Specialized tools: If any specialized tools, software, or code are required to access, manipulate, or analyze the shared scientific data for replication or reuse, did the plan describe how they can be accessed and the likelihood that they will remain available over time? Yes, No	62.5 (57.5,67.5)
III. Standards	83.5 (79.6,87.3)	3.1 Standards indication: Were the applied data standards and associated metadata indicated? Yes, No	78.1 (73.8,82.4)
		3.2 Standards Extent: Was the extent of standardization stated, i.e., fully, partially, not needed, or not stated? Answer options: STD, Not STD, Part STD, not stated.	75.6 (71.1,80.0)
IV. Preservation	89.7 (86.5,92.8)	4.1 Repository Classification: Was the data repository classified? Answer options: NIH Discipline-specific vs. Other Generalist vs. NIH Generalist vs. Other Generalist vs. shared as supplemental data with publication vs. project website vs. call PI vs. none stated	85.4 (81.7,89.0)
		4.2 Data sharing time: Within what timeframe will data be shared? Answer options: (Before pub, with pub, after pub, not stated)	87.5 (84.0,90.9)
		4.3 Repository named: The DMSP named repository(s) for each source/modality of data to be shared. Yes, No	83.0

https://doi.org/10.1017/cts.2025.10212 Published online by Cambridge University Press

			(79.1,86.9)
		4.4 Data matching repository: Did the data source/modality to be shared match the DMSP-stated repository? Yes, No	83.1 (79.2,87.0)
		4.5 DOI: Was a dataset-level persistent unique identifier (used to facilitate discovery) stated? Yes, No	73.4 (68.8,78.0)
		4.6 Sharing time stated: Was the timeframe when the scientific data will be made available stated? Yes, No	62.5 (57.4,67.5)
V. Access	86.6 (83.0,90.1)	5.1 Access limitation: Did the DMSP address any restrictions on data sharing in the planned informed consent? (e.g., disease-specific limitations, particular communities' concerns). Yes/No	85.2 (81.6,88.9)
		5.2 Privacy protection: Did the DMSP state or confirm no regional or local privacy and confidentiality protections, such as de-identification, Certificates of Confidentiality, and other protective measures? Yes, No	62.2 (57.2,67.3)
		5.3 Tribal or state law: If applicable, did the DMSP state or confirm no tribal or state law, regulation, policy, or existing or anticipated agreement limitations on data sharing? Answer options: Yes, No, N/A	85.5 (81.9,89.2)
		5.4 Access control: If applicable, did the DMSP state whether access to shared scientific data derived from humans will be controlled? Answer: Yes, No, N/A	89.5 (86.3,92.6)
VI. Oversight	85.8 (82.1,89.4)	No Sub-Elements	

https://doi.org/10.1017/cts.2025.10212 Published online by Cambridge University Press

Table 2 Elements Addressed by Each DMSP Group

Required Elements	Idiosyncratic (N=78)	Local (N=100)	NIH (N=180)	Total (N=358)	P value
I. Data	74 (94.9%)	100 (100.0%)	180 (100.0%)	354 (98.9%)	0.002
II. Tools	21 (26.9%)	97 (97.0%)	178 (98.9%)	296 (82.7%)	< 0.001
III. Standards	22 (28.2%)	99 (99.0%)	178 (98.9%)	299 (83.5%)	< 0.001
IV. Preservation	44 (56.4%)	99 (99.0%)	178 (98.9%)	321 (89.7%)	< 0.001
V. Access	33 (42.3%)	99 (99.0%)	178 (98.9%)	310 (86.6%)	< 0.001
VI. Oversight	30 (38.5%)	100 (100.0%)	177 (98.3%)	307 (85.8%)	< 0.001

https://doi.org/10.1017/cts.2025.10212 Published online by Cambridge University Press