Received: 6 April 2024 | Revised: 8 August 2024 | Accepted: 24 August 2024

DOI: 10.1002/lrh2.10456

RESEARCH REPORT

Learning Health Systems

Medical researchers' perception of sharing of metadata from case report forms

Alexandra Meidt[1] | Carolin Walter[1] | Christoph U. Lehmann[2] | Martin Dugas[3]

[1]Institute of Medical Informatics, University of Münster, Münster, Germany

[2]University of Texas Southwestern Medical Center, Dallas, Texas, USA

[3]Institute of Medical Informatics, Heidelberg University Hospital, Heidelberg, Germany

Correspondence
Alexandra Meidt, Institute of Medical Informatics, University of Münster, Albert-Schweitzer-Campus 1, Building A11, 48149 Münster, Germany.
Email: alexandra.meidt@uni-muenster.de

Abstract

Introduction: Publishing medical metadata stored in case report forms (CRFs) is a prerequisite for the development of a learning health system (LHS) by fostering reuse of metadata and standardization in health research. The aim of our study was to investigate medical researchers' (MRs) willingness to share CRFs, to identify reasons for and against CRF sharing, and to determine if and under which conditions MRs might consider sharing CRF metadata via a public registry.

Methods: We examined CRF data sharing commitments for 1842 interventional trials registered on the German Clinical Trials Registry (DRKS) from January 1, 2020, to December 31, 2021. We invited 1360 individuals registered as contacts on DRKS to participate in a web-based survey between May 10, 2022, and June 30, 2022.

Results: Only 0.3% (5/1842) of data sharing commitments in DRKS included a plan to share blank CRFs. Survey results showed high support for CRF sharing. More than 70% of respondents (223/301) were willing to share their CRFs, and 83.7% (252/301) were interested in CRF reuse. The most frequently reported reason for CRF sharing was improvement of comparability and interpretability of patient data (244/301; 81.0%). The most frequently reported reason against CRF sharing was missing approval by the sponsor (160/301; 53.2%). Researchers conducting commercial trials were significantly less likely to share CRFs than those conducting noncommercial trials (63.3% vs. 76.2%, OR 0.54, 95% CI 0.32–0.92) and they were less likely to reuse CRFs (78.5% vs. 84.6%, OR 0.66, 95% CI 0.35–1.24). The most frequently mentioned prerequisite for publication of CRFs in a public registry was its trustworthiness (244/301, 81.1%).

Conclusion: Data sharing commitments in DRKS revealed a low awareness of CRF sharing. Survey results showed generally strong support for CRF sharing, including the willingness to publish CRFs in a public registry, although legal and practical barriers were identified.

KEYWORDS

case report forms, clinical research, cross-sectional survey, interoperability, metadata classification, metadata repository, metadata standards, metadata trends, sharing of metadata, translational research

1 | INTRODUCTION

Medical research is an important part of a learning health cycle. Understanding medical study results and drawing conclusions that then result in improvements in care depends on well-thought-out study designs, including error-free and complete data documentation of study results. Case report forms (CRFs) are essential for the documentation of clinical studies.[1] The layout and content of CRFs vary greatly depending on the research question. To guarantee study data accuracy, disambiguation, and completeness, CRFs must be designed carefully following standard guidelines.[2] CRFs congruency and standards on content and design exist usually only within an institution or a network (and sometimes only within one researcher's work), with negative effects on the reproducibility of studies. Despite existing recommendations to include study documentation forms in the study protocol and in the study's report,[3,4] most CRFs and their metadata are not publicly available with the exception of eligibility criteria published in study registries such as Clinicaltrials.gov or the International Clinical Trials Registry Platform (ICTRP).[5,6] This article refers to CRF metadata as a description of collected data elements in a case report form, including the item name (e.g. "age"), the description (e.g. "data from identity card"), the data type (e.g. "integer"), and the measurement units (e.g. "years").

Sharing of individual patient data (IPD) generated by clinical trials with other researchers is considered an ethical obligation.[7,8] However, other trial materials (e.g. study protocol, study reports, blank case report forms) are frequently unavailable and neglected in public discussions. Experts in clinical research have elaborated principles for IPD sharing, including the requirement to make the missing other data objects like CRFs available.[9,10]

The publication of CRF metadata is important for several reasons: They are important trial artifacts that allow interpretation and replication of trials.[11] Publishing CRF metadata supports the IPD request process since researchers could identify relevant trials more easily.[9] Open access to CRF metadata could further improve and speed up CRF development as metadata of similar studies could be recycled, which in turn would lead to standardization and harmonization of data in health research.[12] Comparing and analyzing metadata used in different studies supports the development of common data elements (CDEs) and solves interoperability and data sharing challenges in health care and research.[13]

Prior research explored aspects of IPD sharing.[14–16] Researchers most commonly expressed concerns about inappropriate data use and misinterpretation of data, participant privacy protection, costs and efforts associated with data sharing, and liability/legal concerns. While some of the concerns regarding IPD sharing may also apply to CRF sharing, others are likely to be less relevant. Since CRF sharing does not include sensitive patient data, we assume that generally the willingness to share CRFs may be higher than to share IPD.

2 | METHODS

Our study is comprised of a trial record analysis and an online cross-sectional survey. The study was informed by previously published studies on IPD data sharing[14–16] and was reported according to the **ST**rengthening the **R**eporting of **OB**servational studies in **E**pidemiology (STROBE) statement's principles.[17]

2.1 | Study design and setting

Our analysis of trial records included the interventional trials registered on the DRKS[18] from January 1, 2020, to December 31, 2021. The DRKS is the approved World Health Organization (WHO) Primary Trial Registry in Germany and member of the International Clinical Trials Registry Platform (ICTRP) network. We focused on interventional trials because public funders and journal policies require study registration as a prerequisite for funding or publication. Data sharing commitments for these trials were analyzed for CRF sharing. We invited individuals registered as study/trial contacts for scientific questions to take our online survey.

2.2 | Participants

We extracted the trial information of all interventional studies registered on DRKS from January 1, 2020, to December 31, 2021, from the registry.

When registering studies on DRKS, medical researchers are routinely asked to describe their IPD sharing plan (free-text field). The entries provide information about the researchers' willingness to share IPD and additional study material. For our trial record analysis, we examined manually the sharing plans of 1842 interventional trials with regard to the intention to share CRFs.

We invited the individuals registered as contacts for scientific questions for the trials by email to respond to our online survey. Since the DRKS is primarily used by German-speaking investigators, a questionnaire in German was designed to increase the response rate. Consequently, to avoid language barriers, we excluded 50 contact persons not coming from Germany, Austria, or Switzerland. If a contact person was registered for more than one interventional trial, we selected only one trial randomly and included that trial into the sample, excluding any others. After excluding duplicates, 1390 unique contact persons were invited to take the survey. If emails were rejected, alternative email addresses were identified via absence notices or through a Google search for the individual. We were unable to contact 30 individuals and excluded them, leaving a final sample size of 1360 persons. The flow chart of our survey sample is available as Appendix S1.

2.3 | Survey instrument and data collection

We developed a 16-item questionnaire using a nominal, Likert-type scale, single and multiple choice, and open-ended formats. The English translation of our questionnaire can be found in Appendix S2. As we addressed our survey to individuals, who may not have an informatics background, we referred to "CRF metadata" as "blank case report

forms" or simply "CRFs," and we replaced the term "metadata registry" with the term "public registry for case report forms."

Six questions collected information on respondents' characteristics (role, institution, study type, number of planned and started trials within the last and next three years, proportion of electronically captured study documentation). The remaining items inquired about respondents' willingness to share CRFs and other trial data, their interest in reusing CRFs, reasons for and against CRF sharing, prerequisites for CRF sharing via a public registry, benefits of a public CRF registry for different user types, willingness to share different CRF types, time of study publication, and choice of platforms for CRF publication. The last, open-ended question allowed for general comments. To decrease the response burden, an adaptive, conditional questioning format was chosen. We used the web-based survey platform LimeSurvey (version 5.3.32)[19] for questionnaire creation, invitation of participants, and administration of participants' responses.

We pretested the questionnaire with six project team members (four physicians, two research associates) and three persons unaffiliated with the research team (one clinician scientist, one representative of a clinical research organization, and one member of the pharmaceutical industry). After incorporating their input, we sent the invitation to the survey via an email that included a short informational message and an individual access key to the survey to prevent duplicate entries from the same individual.

The anonymity of participants was preserved by storing responses and participant data separately. Participation was voluntary, and respondents could withdraw from participation in the survey at any time or they could skip any question. Furthermore, participants could pause their response at any time and continue the survey later by using the access key.

Participants provided informed consent and had to confirm that they understood the survey's terms before they could respond to the questionnaire. Results were reported in aggregate and did not contain personally identifiable information. The survey was available from May 10 to June 30, 2022. We sent two reminders in two-week intervals. Participants were incentivized to participate by a donation of three euros to a Pediatric Cancer Support organization for completing the survey.

2.4 | Statistical analysis

All statistical analyses were conducted with Excel, R (version 4.1.3), and the R package "stats".[20] We calculated odds ratios and corresponding confidence intervals (95% threshold) for all study subgroups, while we utilized a group-wise chi-squared test to test for significant differences in the trial characteristics for survey respondents versus nonrespondents and the factors for researchers' willingness or unwillingness to share and reuse CRFs, respectively. P-values were not adjusted for multiple testing. We calculated the survey's response rate as the ratio of persons completing the survey to eligible contact persons (American Association for Public Opinion Research response rate 4.0).[21]

3 | RESULTS

3.1 | Analysis of trials

Of the 1842 trials analyzed on May 30, 2022, only 5 sharing plans (0.3%) included commitments to CRF sharing either by direct ("CRF") or indirect mention (e.g., "tools used for data collection").

3.2 | Cross-sectional survey

Of 1360 contact persons invited to our survey, 301 completed the questionnaire (response rate = 22.1%). Incomplete questionnaires ($n = 78$) were excluded from the analysis.

3.3 | Characteristics of survey respondents

Most respondents (71.1%) were study coordinators, followed by principal investigators (41.5%) and sponsors (21.3%) (multiple answer possibilities). Most respondents (60.5%) were affiliated with an academic institution, 28.9% with health care, 6.6% with private industry, and 4.0% with nonprofit, governmental, and other institutions. Nearly all respondents (95.0%) ran noncommercial trials and 26.2% commercial trials (multiple answer possibilities). The largest group (40.5%) of respondents stated that they initiated 3–5 studies within the last 3 years and 42.2% planned to start 1–2 new studies within the next 3 years (Table 1).

A third of respondents (110–36.5%) indicated that more than 75% of their study documentation consisted of electronic case report forms. More detailed information can be found in Table 2.

Because clinical trial characteristics for survey respondents compared to nonrespondents were similar, we assumed that our respondents were representative of the overall survey sample. Summary statistics for all clinical trial variables are included in Appendix S3.

3.4 | Willingness to share blank CRFs and other trial data

The majority of respondents (223–74.1%) were willing to share the blank CRFs of their studies upon request. Even more (245–81.4%) would share the study protocol and 121 (40.2%) individual participant data. Table 3 provides additional data on willingness to share. The most commonly mentioned prerequisite for publication of blank CRFs in a public registry was the registry's trustworthiness ($n = 244$, 81.1%), followed by the requirement that the CRF publication should not generate any additional effort for the researcher ($n = 203$, 67.4%), and the option to delete previously published CRFs at any time ($n = 193$, 64.1%). Additional motivation to use a public CRF registry included the expressed interest to do so by a

journal ($n = 135$, 44.9%), the clinical trial funder ($n = 119$, 39.5%), or a sponsor ($n = 97$, 32.2%). Figure 1 provides an overview of all responses.

TABLE 1 Characteristics of survey respondents.

Characteristic	Survey respondents $N = 301$	*n* (%)
Role*		
Sponsor	64	21.3%
Principal Investigator	125	41.5%
Study coordinator	214	71.1%
Other	10	3.3%
Institution		
Academic institution	182	60.5%
Health care organization	87	28.9%
Private industry	20	6.6%
Nonprofit organization	3	1.0%
Government	7	2.3%
Other	2	0.7%
Study type[a]		
Commercial	79	26.2%
Noncommercial	286	95.0%
Number of studies in the last 3 years[b]		
0	3	1.0%
1–2	110	36.5%
3–5	122	40.5%
6–10	50	16.6%
>10	15	5.0%
Not indicated	1	0.3%
Number of studies anticipated in the next 3 years		
0	12	4.0%
1–2	127	42.2%
3–5	122	40.5%
6–10	25	8.3%
>10	15	5.0%

Abbreviations: %, percentage; n, number.
[a]Percentages do not sum to 100 as respondents could select more than one entry for these fields.
[b]Percentages do not sum exactly to 100 due to rounding.

Participants stated that clinical researchers would benefit most from a public CRF registry. The majority (257–85.4%) suggested a high or very high benefit for this user group (5-point Likert scale). Most respondents indicated high or very high benefits for all listed user groups (Figure 2).

Respondents, who reported to be willing to share their CRFs via a public registry, were asked which of their CRFs they would be willing to share and under which circumstances. More than half of the respondents would share all CRFs ($n = 161$, 53.5%). A third (96–31.9%) would share inclusion and exclusion criteria and 84 (27.9%) demographic forms (multiple responses possible – Table 4).

Almost half (136–45.2%) of respondents were willing to publish their CRFs jointly with the publication of study results (Table 5).

Only a minority (23–7.6%) of respondents was able to name potential venues to publish their CRFs. The most frequently mentioned option for CRF publication was "supplementary material of publications together with study results and/or the study protocol" (Table 6).

3.5 | Rationale for sharing blank CRFs

Participants were asked to rate a set of provided reasons for CRF sharing as very important, less important, or of no importance.

TABLE 3 Willingness to share data and/or documents by predefined categories.

What data or documents of your study/studies would you make available to other researchers on request?	$N = 301$ *n* (%)[a]	*n* (%)[a]
None	16	5.3%
Anonymized, individual patient data (IPD)	121	40.2%
Study protocol	245	81.4%
Statistical analysis plan (SAP)	144	47.8%
Informed consent form (ICF)	183	60.8%
Clinical study report (CSR)	118	39.2%
Blank case report forms (CRFs)	223	74.1%
Other	2	0.7%

[a]Percentages do not sum to 100 as respondents could select more than one entry for these fields.

TABLE 2 Proportion of electronic case report forms (eCRFs) in study documentation by institution.

Share in %	Academic institution	Health care[a]	Private industry	Nonprofit organization	Government	Other	Total = 301
<25	43 (23.6%)	29 (33.3%)	3 (15.0%)	1 (33.3%)	2 (28.6%)	0 (0.0%)	78 (25.9%)
25–49	32 (17.6%)	19 (21.8%)	0 (0.0%)	0 (0.0%)	1 (14.3%)	0 (0.0%)	52 (17.3%)
50–75	41 (22.5%)	10 (11.5%)	8 (40.0%)	0 (0.0%)	0 (0.0%)	2 (100.0%)	61 (20.3%)
>75	66 (36.3%)	29 (33.3%)	9 (45.0%)	2 (66.7%)	4 (57.1%)	0 (0.0%)	110 (36.5%)

[a]Percentages do not sum exactly to 100 due to rounding.

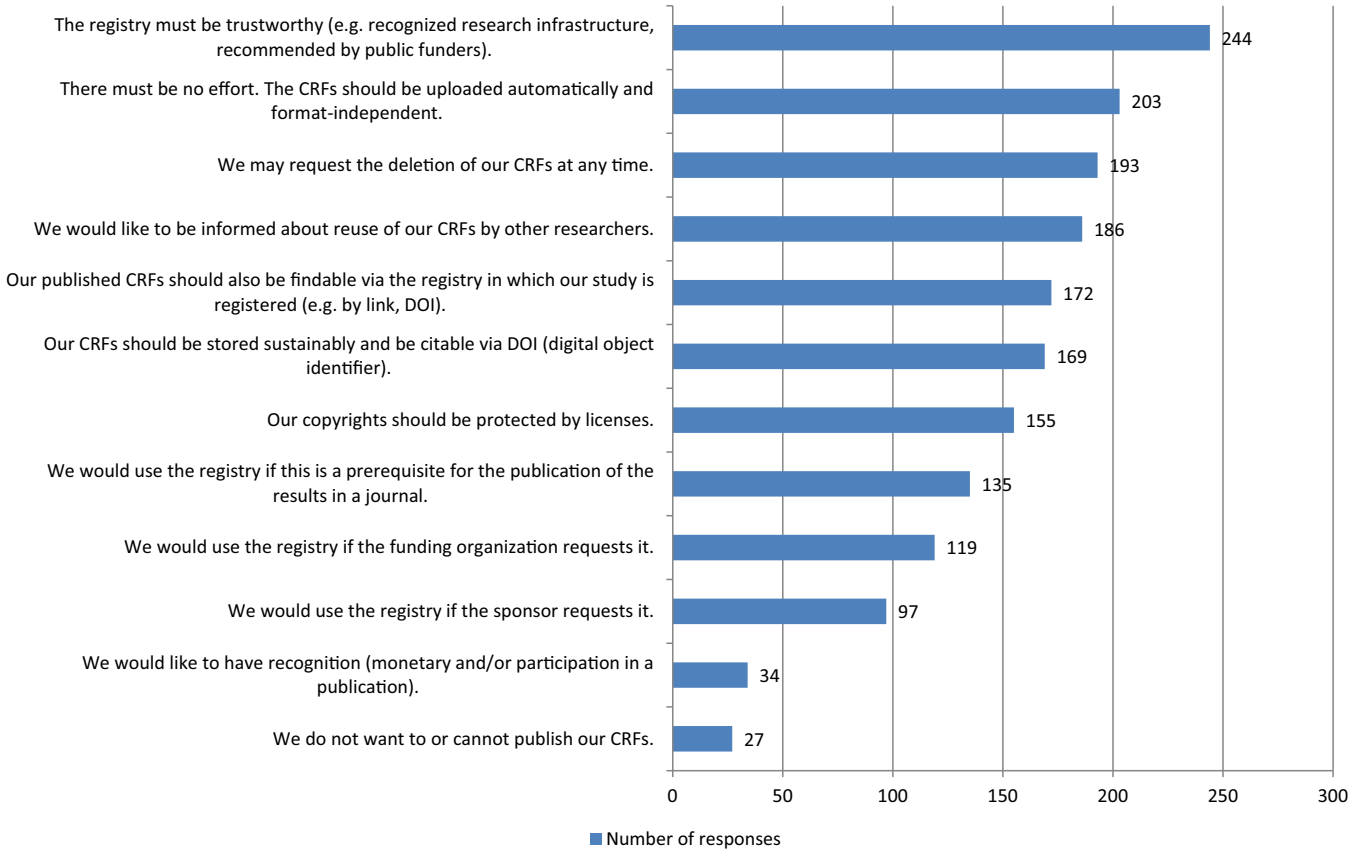

FIGURE 1 Prerequisites for sharing the blank CRFs via a public registry (more than one response possible).

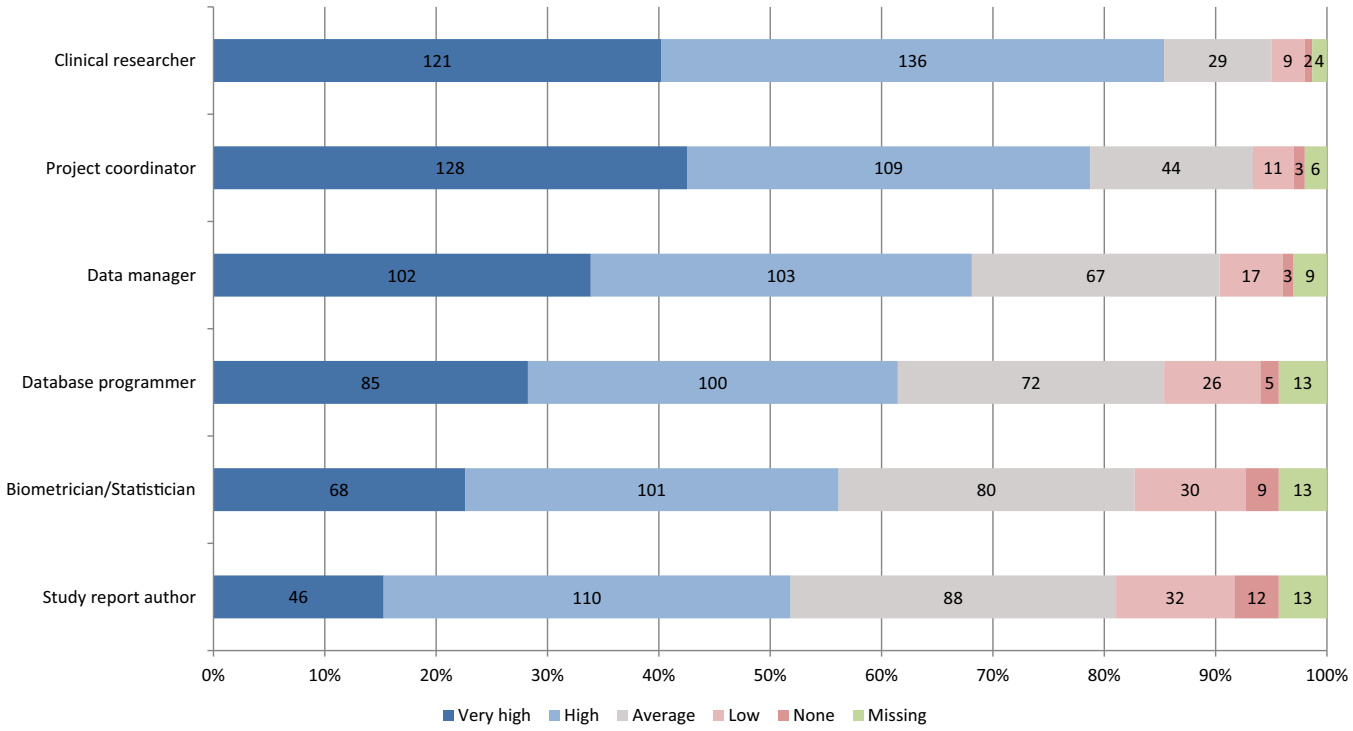

FIGURE 2 Perceived Benefits of a public CRF registry for different user groups (5 point Likert Scale).

TABLE 4 Willingness to share CRFs for predefined categories.

Which CRFs of your study/studies would you share with other researchers?	$N = 301$ n (%)[a]	n (%)[a]
All	161	53.5%
Anamnesis	56	18.6%
Demography	84	27.9%
Inclusion and exclusion criteria	96	31.9%
Follow-up examination	58	19.3%
Physical examination	46	15.3%
Laboratory/ECG	32	10.6%
Target effect size	61	20.3%
Pathology/Histology	14	4.7%
Score	49	16.3%
Study completion	45	15.0%
Adverse event	56	18.6%
Other	0	0.0%

[a]Percentages do not sum to 100 as respondents could select more than one response option.

TABLE 5 Time of publication as indicated by respondents.

When would you be willing to publish the CRFs of your study?	$N = 301$	n (%)[a]
Before start of studies	38	12.6%
During the ongoing study	35	11.6%
Immediately after completion of the study	39	13.0%
Together with the publication of the study results	136	45.2%
>6 months after the end of the study, regardless of publication of the results	21	7.0%
Other	3	1.0%
Not indicated	29	9.6%

[a]Single choice, percentages sum to 100.

Participants considered "Improvement of comparability and interpretability of individual patient data" as the most important reason for CRF sharing (244–81.1%), followed by "improvement of new CRF quality as proven CRFs can be reused" (234–77.7%) and "standardization of data management in clinical trials" (223–74.1%). Figure 3 presents a summary of participants' assessments for all options.

3.6 | Reasons for not sharing blank CRFs

The most frequently cited reason for not sharing CRFs was "missing approval by the sponsor" (160–53.2%), followed by "missing approval of the company commissioned with CRF development" (137–45.5%). "Lack of time and of financial resources" were considered by 125 respondents (41.5%) as a very important reason. A third (111–36.9%) stated that their "use of licensed CRFs" was a very important obstacle to CRF sharing. Additional reasons against CRF sharing assessed by participants can be found in Figure 4.

TABLE 6 Potential platforms for CRF publication as indicated by respondents.

What possibilities for the publication of CRFs do you know?	Frequency of mention[a]
Supporting information of publications (publication of study results and/or study protocol)	17
Study registries (e.g. ClinicalTrials.gov, DRKS)	3
Open Science Framework (OSF)	3
Portal of Medical Data Models (MDM Portal)	2
Own website, cloud or project website	2
Own validation studies	2
ResearchGate	1
Guidelines	1
Books	1
Metadata repositories	1
RedCap	1

[a]Free-text field, more than one response possible.

3.7 | Factors associated with medical researchers' willingness to CRF sharing

Table 7 displays the six factors analyzed in connection with the medical researchers' willingness to share their blank CRFs: "role of respondent," "institution," "study type," "number of studies started within the last three years," "studies planned within the next three years," and the "share of electronic case report forms (eCRFs) used in study documentation." In bivariate analysis, "institution" and "study type" were significantly associated with the willingness to share CRFs ($p = 0.004$ and $p = 0.031$). Investigators affiliated with an academic institution were over five times more likely willing to share CRFs compared to those who were affiliated with private industry (OR 0.18, 95% CI 0.07–0.48). Investigators who conducted noncommercial trials were nearly two times more willing to share CRFs compared to those who conducted commercial trials (OR 0.54, 95% CI 0.32–0.92). In addition, investigators who were planning less than 2 studies within the next 3 years were significantly less likely to share CRFs compared to those who planned 3–5 further studies (50.0% vs 81.1%, OR 0.23, 95% CI 0.07–0.79 and 70.1% vs 81.1%, OR 0.54, 95% CI 0.30–0.98). None of the other factors were associated significantly with willingness to CRF publication.

3.8 | Medical researchers' willingness to CRF reuse

Most respondents (252–83.7%) were willing to reuse CRFs from studies with similar research questions. We analyzed six factors regarding

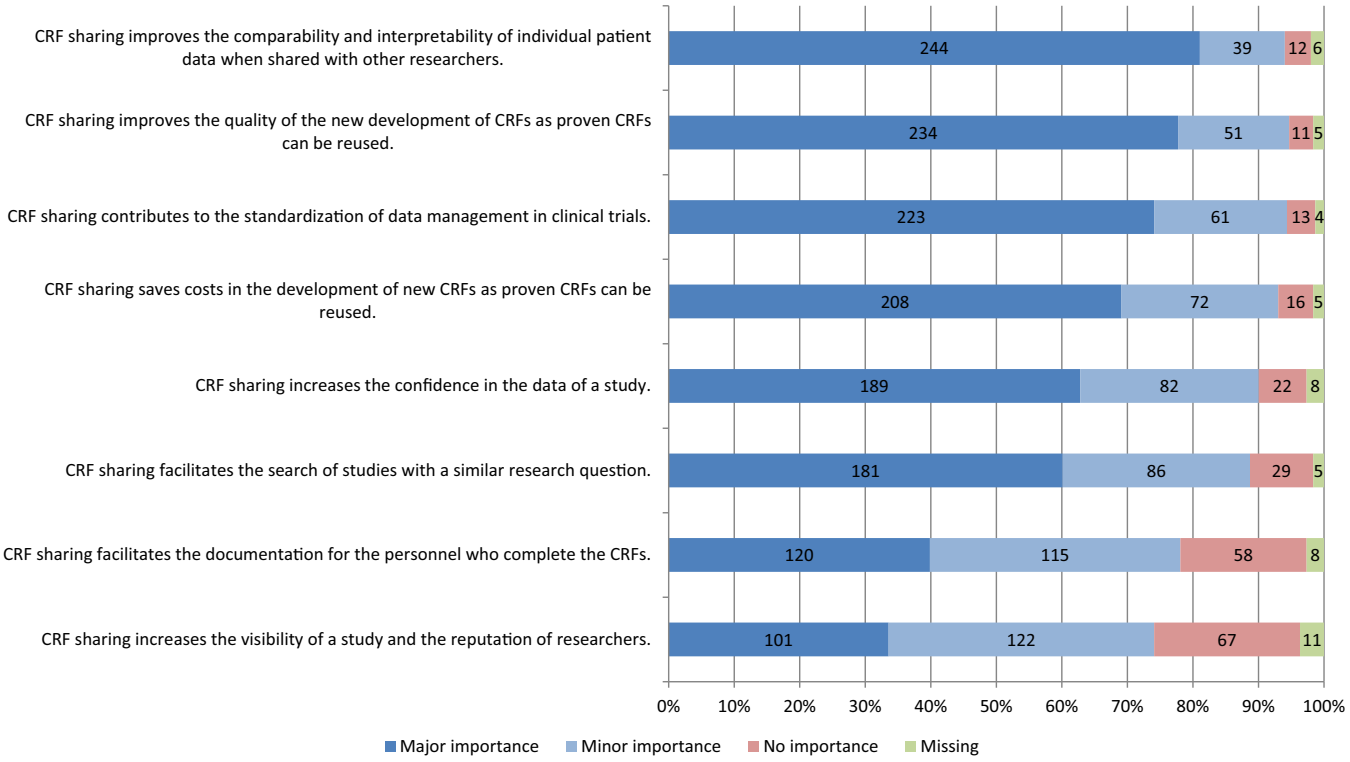

FIGURE 3 Reasons for sharing blank case report forms.

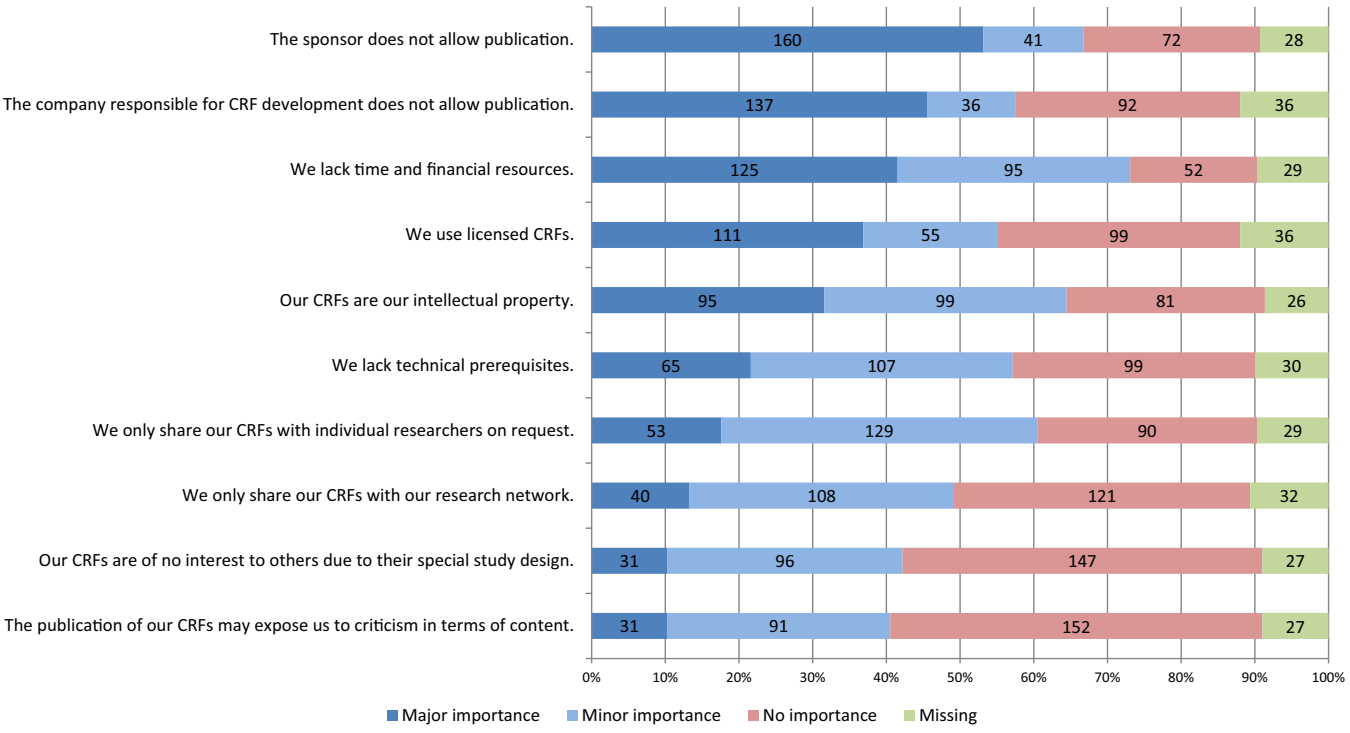

FIGURE 4 Reasons for not sharing blank case report forms.

medical researchers' willingness to reuse CRFs (Table 8). Only the variable "Number of planned studies within next 3 years" was significantly associated ($p = 0.043$). Investigators who did not plan any further studies were significantly less likely to be interested in CRF reuse than investigators who planned 3–5 studies within the next three years (OR 0.17, 95% CI 0.05–0.60).

TABLE 7 Factors associated with medical researchers' willingness to share CRFs.

Variables	Willing to share *n* (%)	Not willing to share *n* (%)	ODDS ratio	95% confidence interval
Role	$p = 0.338$			
Sponsor ($n = 64$)	42 (65.6%)	22 (34.4%)	0.58	0.32–1.07
Principal Investigator ($n = 125$)	91 (72.8%)	34 (27.2%)	0.82	0.49–1.35
Study coordinator ($n = 214$)	164 (76.6%)	50 (23.4%)	1.00	Reference
Other ($n = 10$)	8 (80.0%)	2 (20.0%)	1.22	0.25–5.93
Institution	$p = 0.004$			
Academic institution ($n = 182$)	143 (78.6%)	39 (21.4%)	1.00	Reference
Health care ($n = 87$)	64 (73.6%)	23 (26.4%)	0.76	0.42–1.37
Private industry ($n = 20$)	8 (40.0%)	12 (60.0%)	0.18	0.07–0.48
Nonprofit organization ($n = 3$)	1 (33.3%)	2 (66.7%)	0.14	0.01–1.54
Government ($n = 7$)	5 (71.4%)	2 (28.6%)	0.68	0.13–3.65
Other ($n = 2$)	2 (100.0%)	0 (0.0%)		
Study type	$p = 0.031$			
Commercial ($n = 79$)	50 (63.3%)	29 (36.7%)	0.54	0.32–0.92
Noncommercial ($n = 286$)	218 (76.2%)	68 (23.8%)	1.00	Reference
Number of studies last 3 years	$p = 0.257$			
0 (n = 3)	3 (100.0%)	0 (0.0%)		
1–2 ($n = 110$)	76 (69.1%)	34 (30.9%)	0.70	0.39–1.25
3–5 ($n = 122$)	93 (76.2%)	29 (23.8%)	1.00	Reference
6–10 ($n = 50$)	40 (80.0%)	10 (20.0%)	1.25	0.56–2.80
>10 ($n = 15$)	11 (73.3%)	4 (26.7%)	0.86	0.25–2.90
Not indicated ($n = 1$)	0 (0.0%)	1 (100.0%)		
Number of studies next 3 years	$p = 0.095$			
0 ($n = 12$)	6 (50.0%)	6 (50.0%)	0.23	0.07–0.79
1–2 ($n = 127$)	89 (70.1%)	38 (29.9%)	0.54	0.30–0.98
3–5 ($n = 122$)	99 (81.1%)	23 (18.9%)	1.00	Reference
6–10 ($n = 25$)	18 (72.0%)	7 (28.0%)	0.60	0.22–1.60
>10 ($n = 15$)	11 (73.3%)	4 (26.7%)	0.64	0.19–2.19
Share of eCRFs used in study documentation	$p = 0.364$			
<25% ($n = 78$)	61 (78.2%)	17 (21.8%)	1.67	0.86–3.27
25%–49% ($n = 52$)	40 (76.9%)	12 (23.1%)	1.56	0.73–3.33
50%–75% ($n = 61$)	47 (77.0%)	14 (23.0%)	1.57	0.76–3.22
>75% ($n = 110$)	75 (68.2%)	35 (31.8%)	1.00	Reference

Note: Differences between groups were considered significant at the 5% level.

Analysis of free-text comments revealed general support for CRF sharing, but also showed concerns regarding legal barriers such as licenses and missing requirements by funders or by law in addition to concerns about unrewarded efforts. All comments were summarized in Appendix S4

4 | DISCUSSION

4.1 | Principal findings

We aimed to gain insight into medical researchers' attitude toward sharing metadata stored in case report forms. While the survey showed strong general support of CRF sharing (74.1%), the analysis of data sharing plans in trial registration records revealed very limited actual sharing of these documents (0.3%). As expected, the willingness of researchers to share original CRFs was higher than the willingness to share original IPD. Previous studies showed that protecting participant confidentiality constituted an important reason for withholding IPD.[14–16] However, actual CRF sharing was also uncommon, even though it does not involve the sharing of sensitive personal data.

Our findings indicate that willingness to share CRFs was significantly associated with funding and institution type. Researchers receiving industry funding were less willing to share CRFs than researchers of noncommercial trials. Researchers working for the

TABLE 8 Factors associated with medical researchers' willingness to reuse CRFs.

Variables	Willing to reuse n (%)	Not willing to reuse n%	ODDS ratio	95% confidence interval
Role	$p = 0.303$			
Sponsor ($n = 64$)	49 (76.6%)	15 (23.4%)	0.57	0.29–1.14
Principal Investigator ($n = 125$)	98 (78.4%)	27 (21.6%)	0.64	0.36–1.13
Study coordinator ($n = 214$)	182 (85.0%)	32 (15.0%)	1.00	Reference
Other ($n = 10$)	8 (80.0%)	2 (20.0%)	0.70	0.14–3.46
Institution	$p = 0.630$			
Academic institution ($n = 182$)	152 (83.5%)	30 (16.5%)	1.00	Reference
Health care ($n = 87$)	73 (83.9%)	14 (16.1%)	1.03	0.51–2.06
Private industry ($n = 20$)	15 (75.0%)	5 (25.0%)	0.59	0.20–1.75
Nonprofit organization ($n = 3$)	3 (100.0%)	0 (0.0%)		
Government ($n = 7$)	7 (100.0%)	0 (0.0%)		
Other ($n = 2$)	2 (100.0%)	0 (0.0%)		
Study type	$p = 0.261$			
Commercial ($n = 79$)	62 (78.5%)	17 (21.5%)	0.66	0.35–1.24
Noncommercial ($n = 286$)	242 (84.6%)	44 (15.4%)	1.00	Reference
Number of studies last 3 years	$p = 0.385$			
0 ($n = 3$)	3 (100.0%)	0 (0.0%)		
1–2 ($n = 110$)	88 (80.0%)	22 (20.0%)	0.52	0.25–1.07
3–5 ($n = 122$)	108 (88.5%)	14 (11.5%)	1.00	Reference
6–10 ($n = 50$)	39 (78.0%)	11 (22.0%)	0.46	0.19–1.10
>10 ($n = 15$)	13 (86.7%)	2 (13.3%)	0.84	0.17–4.13
Not indicated ($n = 1$)	1 (100.0%)	0 (0.0%)		
Number of studies next 3 years	$p = 0.043$			
0 ($n = 12$)	7 (58.3%)	5 (41.7%)	0.17	0.05–0.60
1–2 ($n = 127$)	105 (82.7%)	22 (17.3%)	0.57	0.27–1.19
3–5 ($n = 122$)	109 (89.3%)	13 (10.7%)	1.00	Reference
6–10 ($n = 25$)	19 (76.0%)	6 (24.0%)	0.38	0.13–1.12
>10 ($n = 15$)	12 (80.0%)	3 (20.0%)	0.48	0.12–1.91
Share of eCRFs used in study documentation	$p = 0.547$			
<25% ($n = 78$)	66 (84.6%)	12 (15.4%)	1.38	0.64–2.98
25–49% ($n = 52$)	46 (88.5%)	6 (11.5%)	1.92	0.73–5.06
50–75% ($n = 61$)	52 (85.2%)	9 (14.8%)	1.45	0.62–3.37
>75% ($n = 110$)	88 (80.0%)	22 (20.0%)	1.00	Reference

Note: Differences between groups were considered significant at the 5% level ($p \leq 0.05$).

pharmaceutical industry were less willing to share CRFs than academic researchers. These findings may be attributed to more severe restrictions on researchers within commercial trials, limiting researchers' control of data.

The most frequently reported reason for CRF sharing was improvement of comparability and interpretability of IPD (81.1%). This finding is congruent with results of surveys investigating rationales against IPD sharing among medical researchers. The most commonly indicated reason for withholding IPD was prevention of inappropriate data use and misleading interpretations.[14–16]

As a study's success depends on the development of high-quality data collection forms, which is a time- and cost-consuming endeavor, improvement of CRF quality (77.7%) and cost savings (69.1%) were important reasons for CRF sharing. Standardization and harmonization of data management in clinical trials were seen by 74.1% as an important reason for CRF sharing, highlighting the demand of stakeholders to publish CRF metadata as a means to foster data standardization in health research.[12]

The most frequently reported barrier to CRF sharing was lack of approval by the study's sponsor (53.2%). Reasons may include the obligation to protect commercially sensitive information and intellectual property, but since only 26.2% of the respondents indicated to be engaged in industry funded trials, additional motives must be considered. Lack of time and financial resources for CRF sharing were

indicated by 41.5% of the respondents, which was consistent with drivers against IPD sharing.[14–16] In addition, respondents cited legal issues such as subcontracts with CRF developing companies (45.5%) and the use of licensed CRFs (36.9%). Moreover, medical researchers regarded CRF development as individual academic achievement, declaring it as their own intellectual property, resulting in reluctance to share (31.6%).

Participants' responses showed strong interest in reusing proven CRFs of similar studies (83.7%), which confirmed results of our prior survey dealing with secondary use of CRFs via an open access metadata repository.[22]

4.2 | Implications for policy and practice

An effective learning cycle needs a "multistakeholder learning community"—researchers, clinicians, implementation scientists, and innovators.[23] While we only surveyed a subset of this community—the researchers—our data showed that most researchers surveyed would support CRF sharing and reuse, although their awareness for CRF sharing and its implementation as reflected in trial records was almost nonexistent. It is our belief that we would find similar support if we surveyed clinicians, implementation scientists, and innovators. The reason for the broad support for CRF sharing is rooted in the need of researchers, innovators, and clinicians alike for interoperability of data and data sharing among different entities to reduce the burden and effort associated with an LHS.

For industry-funded trials, protection of commercial interests restricted access to study documents, including blank CRFs, creating a conflict with perceived public benefit from the availability of clinical trial data.[24] Only a few pharmaceutical companies (e.g., GlaxoSmithKline) routinely publish protocols, study reports, and CRFs on their websites.[25] In 2013, the European Federation of Pharmaceutical Industries and Associations released principles for clinical trial data sharing, including the commitment to voluntary share IPD, trial protocols, and results.[26] The online portal Clinical Study Data Request (CSDR) and the European Medicines Agency (EMA) provide access to clinical study reports, including sample CRFs, on request.[27,28] These developments demonstrate that efforts to increase transparency for industry-funded trials are being made. Unfortunately, blank CRFs are still not yet freely available on these portals. One potential solution could be to incentivize the publication of CRFs by extending patent time for new pharmaceuticals or by legislative mandates linked to the new drug approval processes by the Food and Drug Administration and the European Medicines Agency.

Investigator-initiated trials financed by public funds routinely place a stronger obligation to share data on researchers. Stakeholders of health research developed data sharing goals for nonprofit funders of clinical trials.[29] They recommended that these trials should implement data sharing policies in their funding guidelines, define standards such as CDISC,[30] and fund and promote sustainable infrastructures for data sharing and the development of common data elements (CDEs). Usually, CDEs are developed disease-specific,[31–33] but potentially an exchange of CDEs across diseases could "increase the power of data collection".[29] While CDE development would be facilitated through the use of electronic case report forms (eCRFs), only a minority of survey respondents (36.5%) stated that more than 75% of their study documentation was captured electronically. Compared to a prior study in 2004, which reported that 75% of studies used paper-based CRFs,[34] the situation seems to have improved in favor of electronic data capture. Accelerating this development would foster the automated processes of eCRF transformation into machine-readable data, allowing standardized data collection and exchange, which are highly relevant for the study of rare diseases.[35] Further, there is value in eCRFs permitting export into open, standardized formats, such as CDISC ODM,[36] rendering them machine-readable, regardless of the study database used.

Learning systems require "simultaneously operating learning cycles" supported by an "infrastructure" "that can be viewed as an integrated set of services" among others to make "knowledge computable und sharable."[23] Open access platforms could serve an important role as infrastructure to publish and merge data objects generated by medical studies, such as the prototype platform OpenTrials or the database of Genotypes and Phenotypes (dbGaP).[37,38] As of December 14, 2022, dbGaP offered access to data from more than 2300 studies, including more than 2600 CRFs, without prior registration or request.

Metadata repositories created for secondary use of CRFs or single CRF items expand on this service.[39,40] They can contribute to the development of common data elements[41,42] and to data standardization as well as to element exchange in medical research. Furthermore, the creation of a public repository for CRF metadata has the potential to speed up the development of trials through metadata reuse, reducing the burden of developing new CRFs. For an effective LHS, a rapid cycle from study results to novel health care delivery to its subsequent evaluation is essential. Reuse of CRF metadata has the potential to not only standardize care but to speed up this important cycle by speeding up the "performance to data" arm of the cycle.[43] A public repository would drastically reduce the effort and time required to develop novel trials.

A prior study investigated seven publicly accessible metadata repositories in the health care domain.[44] Only the Portal of Medical Data Models (MDM Portal) offered reuse of complete annotated CRFs.[45] As of December 14, 2022, MDM contained more than 24 800 medical forms, with more than 18 600 forms related to clinical trials.

Although German funders already incorporate or plan to incorporate the publication of CRFs in their funding guidelines,[46,47] this study and our prior survey[22] showed that only a minority of respondents, who were mostly conducting publicly funded trials, could name publication platforms for CRFs. Consequently, nonprofit funders should specify recommended publication platforms in their funding guidelines. Moreover, they should collaborate in the development of the infrastructure for CRF sharing to avoid storing them in diverse systems, rendering them unmanageable. Study registries could serve as interim solutions. Medical researchers could be required to upload sample case report forms as PDF files while registering their studies. This solution would require little effort, which increases the success of CRF sharing.

Further, organizations should store CRF metadata in a single place in a standardized, machine-readable format to foster metadata analysis. An infrastructure should be developed according to FAIR principles[48] to guarantee that published CRF metadata is findable, accessible, interoperable, and re-usable. As this requirement is more complex and time-consuming than just uploading CRFs as PDF files, funders should include extra funding for this sharing methodology.

Our survey showed that funders and journals have a major influence on sharing CRF metadata via a public registry. Nonprofit funders and journal editors should use their influence and adopt stronger policies on data sharing.[29,49]

The status of data sharing remains unsatisfactory despite years of sharing initiatives.[50] Industry and academic researchers must collaborate to improve the situation.[51] Public institutions such as universities or university hospitals should also oblige their researchers to share study data and metadata, including CRFs. Research data from medical studies, which were generated with support from public universities' budgets, are public property that should be available to all researchers and not be commercialized by paid licenses. To give credit to CRF authors, CC-BY licenses[52] can be used, which mandate citation of CRF sources.

4.3 | Limitations

Our study was limited to interventional studies registered on DRKS, which may have introduced a bias. Although a representative sample for the survey participants could be obtained from the DRKS, we cannot generalize the results for all international researchers. Furthermore, the risk of typical sampling bias should be considered. It is possible that investigators who were more open-minded toward Open science data themes were more likely to participate in our survey.

5 | CONCLUSION

CRF sharing was strongly supported by most surveyed medical researchers. However, it is implemented rarely. Sharing of CRF metadata could contribute significantly to reasonable IPD sharing, reuse of metadata for similar studies, and data exchange in health research. Funders, journals, and public institutions should collaborate in the development of a common infrastructure for CRF and metadata sharing and implement sharing policies. Further studies with a wider range of users should be carried out to explore the requirements for such an infrastructure, which can become an integral part of the learning health system once it is used by a large number of medical researchers.

AUTHOR CONTRIBUTIONS

Alexandra Meidt conceived the idea for the study, developed the study design, performed the manual and Excel analyses, wrote the manuscript, and prepared all graphs and tables. Carolin Walter performed the analyses in R. Christoph U. Lehmann reviewed and commented on the manuscript. Martin Dugas supervised the workflow. All authors read and edited the manuscript and approved the final version.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the advice of the following experts: Max Blumenstock, Martin Krauss, Philipp Neuhaus, Christian Niklas, Sarah Riepenhausen, Sonja Ständer, Johann Proeve, and Julian Varghese, who pre-tested the questionnaire. Philipp Neuhaus and Leon Schöpfner, who provided technical support in survey administration. The authors would also like to thank all participants for their contributions to this study. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflicts of interest.

DATA AVAILABILITY STATEMENT

The datasets used and/or analyzed during the current study are available from the Corresponding Author on reasonable request.

ETHICS STATEMENT

Permission to conduct the study was granted by the ethics committee of the University Münster on 6 May 2022 (2022-291-f-N). Respondents provided informed consent to take part in the study by reading the information provided on the first page of the study and indicating consent.

ORCID

Alexandra Meidt https://orcid.org/0000-0001-9644-8771

REFERENCES

1. Dugas M, Blumenstock M, Dittrich T, et al. Next-generation study databases require FAIR, EHR-integrated, and scalable electronic data capture for medical documentation and decision support. *NPJ Digit Med*. 2024;7(1):10. doi:10.1038/s41746-023-00994-6 PMID: 38216645; PMCID: PMC10786912.
2. Bellary S, Krishnankutty B, Latha MS. Basics of case report form designing in clinical research. *Perspect Clin Res*. 2014;5:159-166. doi: 10.4103/2229-3485.140555
3. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med*. 2013;158:200-207. doi:10.7326/0003-4819-158-3-201302050-00583
4. International Council for Harmonisation (ICH). Topic E 3 Structure and Content of Clinical Study Reports. https://database.ich.org/sites/default/files/E3_Guideline.pdf. Accessed 9 Aug 2022
5. U.S. National Library of Medicine. ClinicalTrials.gov. https://www.clinicaltrials.gov. Accessed 9 Aug 2022
6. World Health Organization (WHO). International Clinical Trials Registry Platform (ICTRP). https://www.who.int/clinical-trials-registry-platform. Accessed 9 Aug 2022
7. The World Medical Association (1964). The declaration of Helsinki. Ethical principles for medical research involving human subjects (Last amended 2013). https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects. Accessed 9 Aug 2022

8. Taichman DB, Sahni P, Pinborg A, et al. Data sharing statements for clinical trials: a requirement of the International Committee of Medical Journal Editors. *Ethiop J Health Sci*. 2017;27:315-318.
9. Tudur Smith C, Hopkins C, Sydes MR, et al. How should individual participant data (IPD) from publicly funded clinical trials be shared? *BMC Med*. 2015;13:298. doi:10.1186/s12916-015-0532-z
10. Ohmann C, Banzi R, Canham S, et al. Sharing and reuse of individual participant data from clinical trials: principles and recommendations. *BMJ Open*. 2017;7(12):e018647. doi:10.1136/bmjopen-2017-018647
11. Hoffmann T, Glasziou P, Beller E, Goldacre B, Chalmers I. Focus on sharing individual patient data distracts from other ways of improving trial transparency. *BMJ*. 2017;357:j2782. doi:10.1136/bmj.j2782
12. Dugas M, Jöckel KH, Friede T, et al. Memorandum "open metadata". Open access to documentation forms and item catalogs in healthcare. *Methods Inf Med*. 2015;54:376-378. doi:10.3414/ME15-05-0007j
13. Kush RD, Warzel D, Kush MA, et al. FAIR data sharing: the roles of common data elements and harmonization. *J Biomed Inform*. 2020; 107:103421.
14. Rathi V, Dzara K, Gross CP, et al. Sharing of clinical trial data among trialists: a cross sectional survey. *BMJ*. 2012;345:e7570. doi:10.1136/bmj.e7570
15. Rathi VK, Strait KM, Gross CP, et al. Predictors of clinical trial data sharing: exploratory analysis of a cross-sectional survey. *Trials*. 2014; 15:384. doi:10.1186/1745-6215-15-384
16. Tan AC, Askie LM, Hunter KE, Barba A, Simes RJ, Seidler AL. Data sharing-trialists' plans at registration, attitudes, barriers and facilitators: a cohort study and cross-sectional survey. *Res Synth Methods*. 2021;12:641-657. doi:10.1002/jrsm.1500
17. von Elm E, Altman DG, Egger M, et al. The Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med*. 2007; 147:573-577. doi:10.1016/j.jclinepi.2007.11.008
18. Deutsches Register Klinischer Studien. DRKS. https://www.drks.de. Accessed 9 Aug 2022.
19. LimeSurvey. https://www.limesurvey.org/de. Accessed 9 Aug 2022.
20. R Core Team (2022). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing*. The R Foundation for Statistical Computing. https://www.R-project.org. Accessed 9 Aug 2022.
21. American Association for Public Opinion Research. AAPOR. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 2016. https://www.aapor.org/aapor_main/media/publications/standard-definitions20169theditionfinal.pdf. Accessed 9 Aug 2022
22. Meidt A, Riepenhausen S, Neuhaus P, et al. Portal of medical data models: stakeholder feedback and requirements. *Stud Health Technol Inform*. 2021;281:488-489. doi:10.3233/SHTI210208
23. Friedman CP. What is unique about learning health systems? *Learn Health Syst*. 2022;6(3):e10328. doi:10.1002/lrh2.10328
24. Rodwin MA, Abramson JD. Clinical trial data as a public good. *Jama*. 2012;308:871-872. doi:10.1001/jama.2012.9661
25. GlaxoSmithKline. Clinical Study Register. https://www.gsk-studyregister.com. Accessed 14 Dec 2022.
26. European Federation of Pharmaceutical Industries and Associations (EFPIA) and The Pharmaceutical Research and Manufacturers of America (PhRMA). Principles for Responsible Clinical Trial Data Sharing. 2013 Available: http://phrma-docs.phrma.org/sites/default/files/pdf/PhRMAPrinciplesForResponsibleClinicalTrialDataSharing.pdf. Accessed 14 Dec 2022
27. Clinical Study Data Request. 2022. https://clinicalstudydatarequest.com/Default.aspx Accessed 14 Dec 2022.
28. European Medicines Agency. Clinical Data. https://clinicaldata.ema.europa.eu. Accessed 14 Dec 2022
29. Coetzee T, Ball MP, Boutin M, et al. Data sharing goals for nonprofit funders of clinical trials. *J Particip Med*. 2021;13:e23011. doi:10.2196/23011
30. Clinical Data Interchange Standards Consortium. CDISC 2022. https://www.cdisc.org/about. Accessed 14 Dec 2022.
31. Jones CE, Munoz FM, Kochhar S, et al. Guidance for the collection of case report form variables to assess safety in clinical trials of vaccines in pregnancy. *Vaccine*. 2016;34:6007-6014. doi:10.1016/j.vaccine.2016.07.007
32. Sass J, Bartschke A, Lehne M, et al. The German Corona consensus dataset (GECCO): a standardized dataset for COVID-19 research in university medicine and beyond. *BMC Med Inform Decis Mak*. 2020; 20:341. doi:10.1186/s12911-020-01374-w
33. World Health Organization. WHO. Global COVID-19 Clinical Platform Case Report Form (CRF) for Post COVID condition. https://www.who.int/publications/i/item/global-covid-19-clinical-platform-case-report-form-(crf)-for-post-covid-conditions-(post-covid-19-crf-) 2022 Accessed 14 Dec 2022.
34. Alschuler L, Bain L, Kush RD. Improving data collection for patient care and clinical trials. *Sci Career Mag*. 2004. https://www.science.org/content/article/improving-data-collection-patient-care-and-clinical-trials.
35. Kersloot MG, Jacobsen A, Groenen KHJ, et al. De-novo FAIRification via an electronic data capture system by automated transformation of filled electronic case report forms into machine-readable data. *J Biomed Inform*. 2021;122:103897. doi:10.1016/j.jbi.2021.103897
36. Clinical Data Interchange Standards Consortium. CDISC. 2022. https://www.cdisc.org/standards/data-exchange/odm. Accessed 9 Aug 2022.
37. Goldacre B, Gray J. OpenTrials: towards a collaborative open database of all available information on all clinical trials. *Trials*. 2016;17: 164. doi:10.1186/s13063-016-1290-8
38. National Library of Medicine. Database of Genotypes and Phenotypes (dbGaP). 2022. https://www.ncbi.nlm.nih.gov/gap. Accessed 14 Dec 2022
39. Kadioglu D, Breil B, Knell C, et al. Samply.MDR – a metadata repository and its application in various research networks. *Stud Health Technol Inform*. 2018;253:50-54.
40. Dugas M. Design of case report forms based on a public metadata registry: re-use of data elements to improve compatibility of data. *Trials*. 2016;17:566. doi:10.1186/s13063-016-1691-8
41. Varghese J, Holz C, Neuhaus P, et al. Key data elements in myeloid leukemia. *Stud Health Technol Inform*. 2016;228:282-286.
42. Kentgen M, Varghese J, Samol A, Waltenberger J, Dugas M. Common data elements for acute coronary syndrome: analysis based on the unified medical language system. *JMIR Med Inform*. 2019;7:e14107. doi:10.2196/14107
43. Friedman CP, Rubin JC, Sullivan KJ. Toward an information infrastructure for Global Health improvement. *Yearb Med Inform*. 2017;26(1): 16-23. doi:10.15265/IY-2017-004 Epub 2017 Sep 11.
44. Hegselmann S, Storck M, Gessner S, et al. Pragmatic MDR: a metadata repository with bottom-up standardization of medical metadata through reuse. *BMC Med Inform Decis Mak*. 2021;21:160. doi:10.1186/s12911-021-01524-8
45. Dugas M. Portal of medical data models. https://medical-data-models.org Accessed 14 Dec 2022.
46. Bundesministerium für Bildung und Forschung. Richtlinie zur Förderung von klinischen Studien mit hoher Relevanz für die Patientenversorgung, Bundesanzeiger vom. 2019. https://www.gesundheitsforschung-bmbf.de/de/14731.php. Accessed 14 Dec 2022
47. Deutsche Forschungsgemeinschaft, Stellungnahme der Arbeitsgruppe "Klinische Studien" der DFG-Senatskommission für Grundsatzfragen in der Klinischen Forschung, p. 27. 2018 https://www.dfg.de/download/pdf/dfg_im_profil/geschaeftsstelle/publikationen/stellungnahmen_papiere/2018/181025_stellungnahme_ag_klinische_studien.pdf Accessed 14 Dec 2022.

48. The Future of Research Communications and e-Scholarship. Force11. https://www.force11.org/fairprinciples. Accessed 14 Dec 2022
49. Naudet F, Siebert M, Pellen C, et al. Medical journal requirements for clinical trial data sharing: ripe for improvement. *PLoS Med*. 2021;18: e1003844. doi:10.1371/journal.pmed.1003844
50. Ohmann C, Moher D, Siebert M, Motschall E, Naudet F. Status, use and impact of sharing individual participant data from clinical trials: a scoping review. *BMJ Open*. 2021;11:e049228. doi:10.1136/bmjopen-2021-049228
51. European Federation of Pharmaceutical Industries and Associations. EFPIA 2013. https://www.efpia.eu/news-events/the-efpia-view/statements-press-releases/130724-efpia-and-phrma-release-joint-principles-for-responsible-clinical-trial-data-sharing-to-benefit-patients. Accessed 14 Dec 2022.
52. Creative Commons. 2022. https://creativecommons.org/about/cclicenses. Accessed 14 Dec 2022.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Meidt A, Walter C, Lehmann CU, Dugas M. Medical researchers' perception of sharing of metadata from case report forms. *Learn Health Sys*. 2025;9(1): e10456. doi:10.1002/lrh2.10456