

BMJ Open Data-sharing recommendations in biomedical journals and randomised controlled trials: an audit of journals following the ICMJE recommendations

Maximilian Siebert ^{1,2}, Jeanne Fabiola Gaba,^{1,2} Laura Caquelin,¹ Henri Gouraud,¹ Alain Dupuy,² David Moher ³, Florian Naudet¹

To cite: Siebert M, Gaba JF, Caquelin L, *et al*. Data-sharing recommendations in biomedical journals and randomised controlled trials: an audit of journals following the ICMJE recommendations. *BMJ Open* 2020;**10**:e038887. doi:10.1136/bmjopen-2020-038887

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-038887>).

Received 27 March 2020
Revised 23 April 2020
Accepted 30 April 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Univ Rennes, CHU Rennes, Inserm, CIC 1414 (Centre d'Investigation Clinique de Rennes), F-35000, Rennes, France

²Univ Rennes, CHU Rennes, REPERES (Recherche en Pharmaco-épidémiologie et Recours aux Soins), EA 7449, 35000, Rennes, France

³Center for Journalology, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada

Correspondence to

Maximilian Siebert;
maximilian.siebert@univ-rennes1.fr

ABSTRACT

Objective To explore the implementation of the International Committee of Medical Journal Editors (ICMJE) data-sharing policy which came into force on 1 July 2018 by ICMJE-member journals and by ICMJE-affiliated journals declaring they follow the ICMJE recommendations.

Design A cross-sectional survey of data-sharing policies in 2018 on journal websites and in data-sharing statements in randomised controlled trials (RCTs).

Setting ICMJE website; PubMed/Medline.

Eligibility criteria ICMJE-member journals and 489 ICMJE-affiliated journals that published an RCT in 2018, had an accessible online website and were not considered as predatory journals according to Beall's list. One hundred RCTs for member journals and 100 RCTs for affiliated journals with a data-sharing policy, submitted after 1 July 2018.

Main outcome measures The primary outcome for the policies was the existence of a data-sharing policy (explicit data-sharing policy, no data-sharing policy, policy merely referring to ICMJE recommendations) as reported on the journal website, especially in the instructions for authors. For RCTs, our primary outcome was the intention to share individual participant data set out in the data-sharing statement.

Results Eight (out of 14; 57%) member journals had an explicit data-sharing policy on their website (three were more stringent than the ICMJE requirements, one was less demanding and four were compliant), five (35%) additional journals stated that they followed the ICMJE requirements, and one (8%) had no policy online. In RCTs published in these journals, there were data-sharing statements in 98 out of 100, with expressed intention to share individual patient data reaching 77 out of 100 (77%; 95% CI 67% to 85%). One hundred and forty-five (out of 489) ICMJE-affiliated journals (30%; 26% to 34%) had an explicit data-sharing policy on their website (11 were more stringent than the ICMJE requirements, 85 were less demanding and 49 were compliant) and 276 (56%; 52% to 61%) merely referred to the ICMJE requirements. In RCTs published in affiliated journals with an explicit data-sharing policy, data-sharing statements were rare (25%), and expressed intentions to share data were found in 22% (15% to 32%).

Strengths and limitations of this study

- This original research is a comprehensive assessment of data-sharing policies based on the International Committee of Medical Journal Editors (ICMJE) recommendations.
- In our approach we focused on both ICMJE-member and ICMJE-affiliated journals.
- Data-sharing policies are set in a fast-changing environment and results obtained today might be outdated in the future.
- One limitation is that we only relied on online information.

Conclusion The implementation of ICMJE data-sharing requirements in online journal policies was suboptimal for ICMJE-member journals and poor for ICMJE-affiliated journals. The implementation of the policy was good in member journals and of concern for affiliated journals. We suggest the conduct of continuous audits of medical journal data-sharing policies in the future.

Registration The protocol was registered before the start of the research on the Open Science Framework (<https://osf.io/n6whd/>).

INTRODUCTION

In June 2017, the International Committee of Medical Journal Editors (ICMJE) published a statement supporting data-sharing practices for randomised controlled trials (RCTs). For the ICMJE, 'there is an ethical obligation to responsibly share data generated by interventional clinical trials because trial participants have put themselves at risk' with the aim to 'maximize the knowledge gained' from these outstanding studies. The ICMJE policy requires a specific data-sharing statement to be included in each newly submitted paper (and prespecified in study registration) containing clinical trial data, starting 1 July 2018.¹

Examples of medical journals having a data-sharing policy before this requirement were few. In 2007, the *Annals of Internal Medicine* was the first journal to adopt a policy encouraging data-sharing practices.² *The BMJ* adopted a similar policy encouraging data-sharing in 2009,³ and went further by making it mandatory in 2013 for drugs and devices⁴ and for all RCTs in 2015.⁵ *PLOS* journals also adopted a strict policy enforcing RCT data-sharing in 2014.⁶ No other leading general medical journal has had a specific policy for data-sharing in RCTs.

The ICMJE policy could therefore have an impact on biomedical literature as a whole. At the time of the present research, the ICMJE included two organisations (the US National Library of Medicine and the World Association of Medical Editors) and 14 journals, including leading medical journals such as *The New England Journal of Medicine (NEJM)* and *The Lancet*. In addition, about 5000 affiliated journals follow the ICMJE recommendations.⁷ As this policy is now in place, it is important to monitor its implementation both in the ICMJE-member journals and in the ICMJE-affiliated journals. It is also important to assess intentions to share data among RCTs published in the journals implementing a data-sharing policy.

METHODS

The protocol was registered before the start of the research on the Open Science Framework (OSF) (<https://osf.io/n6whd/>). This study was divided into two parts: a survey of journal data-sharing policies and a survey of published RCTs.

Survey of journal data-sharing policies

Journal eligibility criteria

Two samples of journals were surveyed: the 14 ICMJE-member journals at the time of the present research (*Annals of Internal Medicine*, *British Medical Journal*, *Bulletin of the World Health Organization*, *German Medical Journal*, *Ethiopian Journal of Health Sciences*, *Iranian Journal of Medical Sciences*, *Journal of the American Medical Association*, *Journal of Korean Medical Science*, *New England Journal of Medicine*, *New Zealand Medical Journal*, *PLOS Medicine*, *The Lancet*, *Medical Journal of Chile* and *Danish Medical Journal*) and a sample of ICMJE-affiliated journals listed on the ICMJE website on 1 February 2019. Journals were included if they (1) had medical content; (2) had published at least one RCT in 2018; (3) had published articles in English, German, French, Spanish or Portuguese; (4) had an accessible online website; and (5) were not considered as 'predatory' journals according to Beall's list.⁸

Search strategy for journals

The ICMJE website was consulted to copy the list of all ICMJE-member journals and all 4892 ICMJE-affiliated journals. In cases where an affiliated journal changed its name after registration on the ICMJE website (eg, *Cancer Immunity* changed to *Cancer Immunology Research*), we checked whether the new name was also listed on the

ICMJE website. If this was the case, the journal was considered as non-eligible and was marked as 'discontinued'; otherwise it was included.

All 4892 ICMJE-affiliated journals were assessed for eligibility in random order obtained using the R statistical software.⁹ The results of the randomisation can be found in the supplementary material on the OSF page.¹⁰ The first 489 journals that met the selection criteria (10% of affiliated journals) were included, enabling us to estimate a proportion of 50% (the worst-case scenario for precision estimates) with a precision (boundaries of the 95% CI) of about $\pm 4.5\%$.

Journal selection and data extraction

A data extraction sheet and a data extraction explanatory document were developed.¹⁰ Three investigators in charge of data extraction (MS, LC, HG) had a 1-hour training session and completed a pilot data extraction on 10 journals. For each journal, pairs of these investigators independently assessed eligibility (with reasons in case of non-eligibility) and extracted the data on all the outcomes listed in the following sections from each included journal. Disagreements were resolved by consensus or in consultation with a third investigator (FN).

Outcomes describing journal data-sharing policies

Our primary outcome was the existence of a data-sharing policy (specific data-sharing policy, no data-sharing policy or a policy merely referring to ICMJE requirements) as reported on the journal website. This outcome had to be changed from our initial protocol due to non-response to our emails from the sample of ICMJE-affiliated journals and because some email addresses could not be identified. The change took place before any analysis. For journals mentioning a specific data-sharing policy on their website, 'the explicit statement and various features of these policies were collected: the start date of the data-sharing policy, the type of policy: ICMJE compliant, more stringent than required by ICMJE or less demanding than required by ICMJE (for instance, less demanding could mean that there was no obligation for a data-sharing statement, and more stringent could mean that data were to be shared with other researchers). We also noted whether the policy was limited to clinical trials, and furthermore the indication of one or more preferred data-sharing platform (and if so, which ones) and the existence of any sanctions in case of non-compliance with data-sharing (and if so, what they were). Any existing policy demanding trial registration was also extracted (and if there was one, we noted whether it mentioned prospective registration). The following features of the journals were also extracted: indexed on PubMed, International Standard Serial Number (ISSN or print ISSN), number of issues per year, 2017 journal impact factor (JIF), publisher or publishing group, gender of the editor in chief ('men', 'women' and 'both genders represented', if the coeditors in chief were men and women), country of the journal head office, wealth category of the country where the

editorial office is located as defined by the World Bank¹¹ and the research domain covered by the journal.

Survey of RCTs published in journals with a data-sharing policy

RCT eligibility criteria

Eligible studies were RCTs published after 1 January 2019 in a journal with an explicit data-sharing policy reported on its website and submitted after 1 July 2018. Any RCTs, including cluster trials and cross-over trials, non-inferiority designs and superiority designs, were included. No distinction was made in terms of patients, interventions, comparators or outcomes. We had originally planned to include only phase III studies but realised that this information was not always reported in the publications. Consequently, no distinction in terms of study phase was applied.

Search strategy for RCTs published in journals with a data-sharing policy

ICMJE-member journals were contacted to gather the list of RCTs they published after 1 January 2019. This approach was not used for ICMJE-affiliated journals due to non-response from most of the 14 member journals and because it was not possible to identify an email contact for all these journals. The following search strategy was applied to retrieve all RCTs. For journals indexed on PubMed/Medline, a search algorithm to identify RCTs was developed with the help of a librarian from Rennes 1 University using the Cochrane sensitivity maximising approach¹² and adding further keywords. The exact filter can be found in the supplementary material on the OSF page.¹⁰ For journals not indexed on PubMed, an investigator (MS) screened all articles published after 1 January 2019 to identify RCTs.

All identified RCTs published in ICMJE-member journals with a data-sharing policy were assessed in random order using R.¹⁰ The first 100 that met our selection criteria were included, enabling us to estimate a proportion of 50% (the worst-case scenario for precision estimates) with a precision (boundaries of the 95% CI) of $\pm 9.8\%$. We followed the same approach to include a second sample of 100 RCTs published in ICMJE-affiliated journals.

RCT selection and data extraction

As for the journal selection procedure, a data extraction sheet and a data extraction explanatory document were developed.¹⁰ Three investigators (MS, LC, JFG) had a 1-hour explanation and were trained via a pilot data extraction performed on 10 RCTs. For each RCT, two of these investigators independently assessed eligibility (giving reasons in case of non-eligibility) and extracted the characteristics listed in the following sections from each included published article. Disagreements were resolved by consensus or in consultation with a third investigator (FN).

Outcomes describing data-sharing statements in published RCTs

For this part of our survey, the primary outcome was the intention to share individual patient data (IPD) expressed by the authors in the data-sharing statement (yes/no/unclear). The latter of the three response options, 'unclear', was recorded if the statement was written in a general tone without specifically mentioning that IPD would be available. Secondary outcomes were trial registrations: the existence of trial registration, prospective trial registration and registration of a data-sharing plan. If the trial report mentioned the existence of a data-sharing plan, we checked whether there was an intention to share data or not. For data-sharing under the secondary outcomes, we checked whether a statement was included in the article, whether the statistical code was shared, whether other data sets than IPD were available and if not whether only parts were available, and lastly whether any other documents were available. Under the section data accessibility, we checked to see whether there was a time restriction for data access, whether it was freely accessible or with restrictions, whether the data could be used for any type of purpose and if not whether there was an aim for data use suggested in the proposal, whether there was a specific access mechanism, and whether data requests were reviewed by an independent committee.

Statistical analysis

All outcomes were reported and described by counts, percentages, means (or medians) and SD (or range) with all the corresponding 95% CIs. If available, verbatim quotes from journal policies were used (qualitative analysis). For ICMJE-affiliated journals, the features of the included journals were compared as part of an exploratory analysis between journals with and without a specific data-sharing policy using univariate logistic regression and multivariate logistic regression (which included covariates identified in univariate analyses at a threshold of $p < 0.25$). Due to complete separation observed in our data set, the 'brglm' package in R was used, implementing the bias reduction method developed by Firth.¹³ All analyses were conducted using R V.3.4.1. The supplementary material on the OSF page contains our statistical analysis plan.¹⁰

Changes to the initial protocol

The definition of our primary outcome for journal policies was changed. Indeed, we had initially planned to contact journals with no explicit policy on their website to ask them about the existence of a data-sharing policy. Due to non-response from some of the 14 member journals, and because it was not possible to identify an email contact for all ICMJE-affiliated journals, we decided to rely only on the information presented on the journal websites. Some minor changes were also made. Our selection criteria were simplified so that only one RCT in 2018 was necessary (instead of three over the last 3 years as initially planned). No distinction was made for RCTs in terms of clinical phase. Lastly, we added a secondary

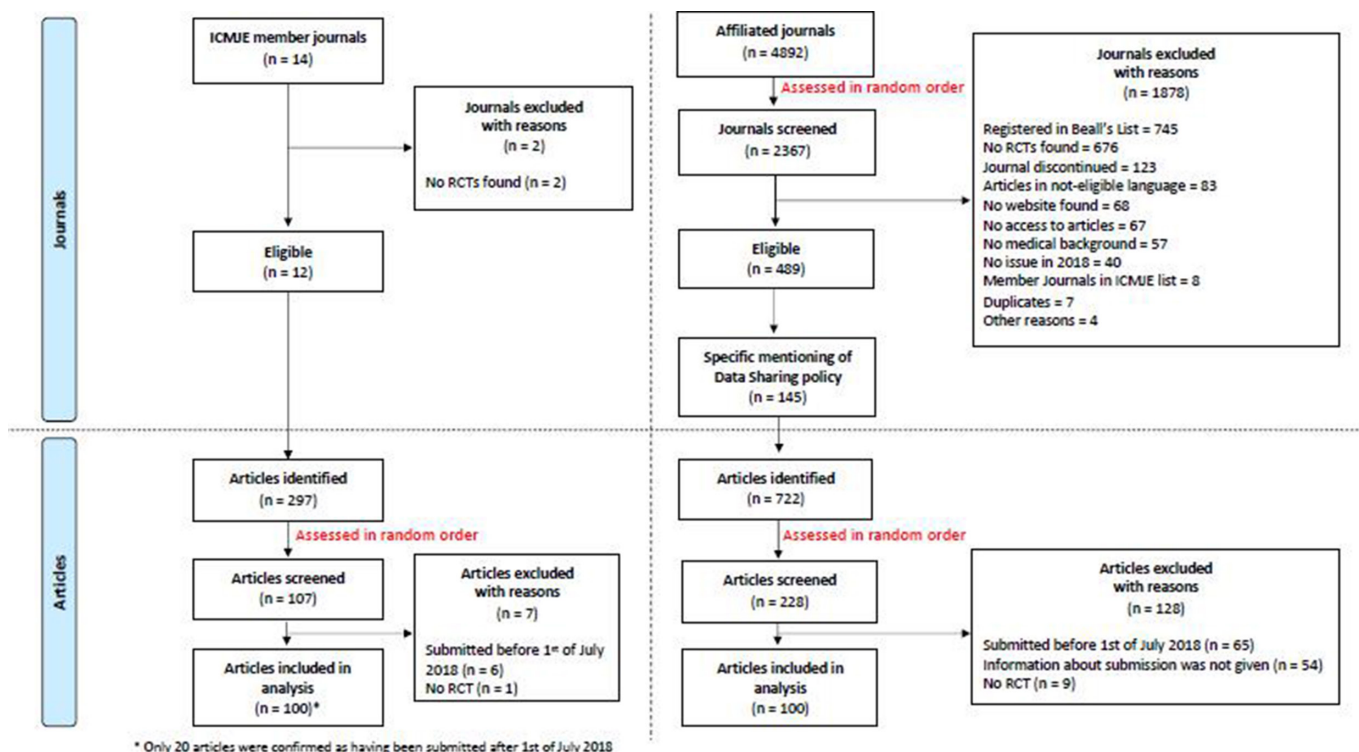


Figure 1 Flow chart of the selection and analysis process for journals and articles. ICMJE, International Committee of Medical Journal Editors; RCT, randomised controlled trial.

outcome, whether or not data requests were reviewed by an independent committee.

Patient and public involvement

We had no established contacts with specific patient groups in this project. No patients were involved in defining the research question or the outcome measures, nor were they involved in the design and implementation of the study. There are no plans to involve patients in the dissemination of results, nor will we disseminate results directly to patients.

RESULTS

Survey of journal data-sharing policies

Journal selection and data extraction

Search for and extraction of eligible journals started on 1 February 2019, ended with a consensus on 11 July 2019, and resulted in 14 ICMJE-member journals and 4892 ICMJE-affiliated journals. Of the affiliated journals, 2367 were randomly screened and 1878 (79%) were excluded, including 745 journals (31% of all screened journals) for which the journal and/or the publisher were listed as 'predatory' on Beall's list. Therefore, 489 ICMJE-affiliated journals were included in analyses as initially planned. The selection process is reported in figure 1.

ICMJE-member journals

The characteristics of the 14 ICMJE-member journals are detailed in table 1.

Twelve of 14 (86%) journals published an RCT in 2018. The *New Zealand Medical Journal* and the *Ethiopian Journal*

of Health Sciences did not publish an RCT in 2018. Eight (57%) journals had a specific data-sharing policy on their website: three (38%) of these had a more stringent policy than required by the ICMJE (IPD to be available for *The BMJ* and *PLOS Medicine*, or explicit demands of data for peer review for the *Annals of Internal Medicine*), four were ICMJE-compliant (50%; *NEJM*, *Danish Medical Journal*, *Journal of the American Medical Association* and *The Lancet*), and one (12%; *Journal of Korean Medical Science*) had a less demanding policy than required by the ICMJE that did not require a data-sharing statement but merely encouraged it. Of 14 medical journals, 5 (35%; *Bulletin of the World Health Organization*, *German Medical Journal*, *Ethiopian Journal of Health Sciences*, *Iranian Journal of Medical Sciences and Medical Journal of Chile*) referred to the ICMJE guidelines and 1 (8%; *The New Zealand Medical Journal*) did not have any policy mentioned on its website (its editorial office said that they had no time to clarify this point with us). Only three journals had a data-sharing policy before 2017. The earliest was presented in 2007 by the *Annals of Internal Medicine*, followed by *PLOS Medicine* and *The BMJ* in 2014 and 2015, respectively.

Three journals (out of 14, 21%; *The BMJ*, *PLOS Medicine* and *The Lancet*) indicated specific data-sharing platforms in their policy: Dryad and Mendeley.

For the eight journals with specific data-sharing statements, five referred specifically to clinical trial data, and for the three others it was for all research data submitted.

Sanctions were described in two journals, *PLOS Medicine* and the *Annals of Internal Medicine*: possible rejection of the manuscript if the data were not provided.

Table 1 Characteristics of journal policies for ICMJE-member and ICMJE-affiliated journals

	ICMJE-member journals (n=14)	ICMJE-affiliated journals (n=489)
Type of data-sharing policy		
Explicit	8 (57%)	145 (30%; 26% to 34%)
Not existing	1 (8%)	68 (14%; 11% to 17%)
ICMJE	5 (35%)	276 (56%; 52% to 61%)
Sanctions in non-compliance of data-sharing	2 (14%)	0
Trial registration demanded		
Yes, with a specification that it must be prospective	7 (50%)	178 (37%; 32% to 41%)
Yes, without specification	3 (21%)	142 (29%; 25% to 33%)
Referring to ICMJE	3 (21%)	114 (23%; 20% to 27%)
No	1 (8%)	55 (11%; 9% to 14%)
Issue/year*	16 (12–51)	6 (4–12)
Impact factor 2017†	11.7 (2.7–35.6)	2.4 (1.5–4)

*Not found for two journals; indicated in median and IQR.

†Not found for 258 journals; indicated in median and IQR.

ICMJE, International Committee of Medical Journal Editors.

Except for the *Ethiopian Journal of Health Sciences* and the *Iranian Journal of Medical Sciences*, all journals had their editorial office in high-income countries.

ICMJE-affiliated journals

The characteristics of the 489 ICMJE-affiliated journals are also presented in [table 1](#). Of these journals, 145 (30%; 95% CI 26% to 34%) had a specific data-sharing policy on their website. Two hundred and seventy-six journals (56%; 52% to 61%) merely referred to the ICMJE guidelines, without any specific mention of a data-sharing policy. Sixty-eight (14%; 11% to 17%) had no data-sharing policy and did not allude to the ICMJE in their recommendations. In contrast, 178 (37%; 32% to 41%) required prospective trial registration, 142 (29%; 25% to 33%) asked for trial registration without specifications, 114 (23%; 20% to 27%) referred to the ICMJE, and 55 (11%; 9% to 14%) did not refer to any trial registration.

Among the 145 journals with a specific data-sharing policy, 11 (7%; 4% to 13%) had a more stringent policy than that required by the ICMJE, 49 (34%; 26% to 42%) journals were ICMJE-compliant, and 85 (59%; 50% to 67%) had a less demanding policy than required by the ICMJE that did not explicitly require a data-sharing statement. Nineteen (out of 145; 13%) journals with data-sharing policies referred only to clinical trial data, while for the rest the statement comprised a more general statement. Out of 145 journals, 94 (65%) had no start date found for the policy, 25 (17%) had a policy starting in early 2018 (January and February), and 26 (18%) had a policy starting on 1 July 2018.

One hundred and one (out of 145; 70%) journals indicated a preferred data-sharing platform, with Mendeley (81 journals), Figshare (79 journals) and Dryad (67 journals) being the three most often cited.

Except for the gender of the editor, all features explored in univariate analyses were associated ($p < 0.25$) with the explicit mention of a data-sharing policy on the journal website and were therefore used in the multivariate analyses. Publisher and wealth category of country of journal offices remained associated with the explicit mention of a data-sharing policy in multivariate analysis. The respective adjusted ORs can be found in [table 2](#).

Survey of RCTs published in journals with a data-sharing policy

RCT selection and data extraction

Search for and extraction of eligible RCTs started on 6 August 2019 and ended with a consensus on 26 September 2019. Among the 12 eligible ICMJE-member journals, the *New Zealand Journal of Medicine* and the *Ethiopian Journal of Health Sciences* did not present any RCT in 2018. *PLOS Medicine* and the *Bulletin of the World Health Organization* provided a list of published articles. Two hundred and ninety-seven RCTs published in member journals were found.¹⁰ We could only confirm for 20 articles that they had been submitted after 1 July 2018. For six articles without data-sharing statements in the *NEJM* and *The Lancet*, we were not sure if they were eligible with respect to the submission date. Authors were contacted and we were able to confirm for two journals that they had been submitted before 1 July 2019. These were replaced, as were the four others where doubt persisted. Among the affiliated journals 722 RCTs were identified and were randomly sorted and assessed for eligibility criteria. [Figure 1](#) details the selection process for both ICMJE-member and ICMJE-affiliated journals.

**Table 2** Journal characteristics associated with an explicit data-sharing policy

	All journals (n=489)	Journals with an explicit data-sharing policy (n=145)	Journals without an explicit data- sharing policy (n=344)	Univariate analysis OR (95% CI)	P value	Multivariate analysis aOR (95% CI)*	P value
Number of issues per year*							
More than 12 (reference)	17 (4%)	9 (6%)	8 (2%)	–	–	–	–
1–5	241 (49%)	45 (31%)	196 (57%)	0.21 (0.07 to 0.56)	0.002	1 (0.25 to 4.17)	0.83
6–12	229 (47%)	90 (73%)	139 (41%)	0.58 (0.21 to 1.56)	0.28	0.88 (0.25 to 3)	0.99
Journal impact factor							
First quartile (reference) (0.126–1.532)	58 (12%)	20 (14%)	38 (11%)	–	–	–	–
Second quartile (1.532– 2.388)	58 (12%)	25 (17%)	33 (10%)	1.43 (0.68 to 3.07)	0.35	0.9 (0.35 to 2.32)	0.82
Third quartile (2.388– 3.993)	57 (12%)	28 (19%)	29 (8%)	1.81 (0.87 to 3.92)	0.12	0.77 (0.28 to 2.05)	0.59
Fourth quartile (3.993– 20.871)	58 (12%)	32 (22%)	26 (8%)	2.3 (1.11 to 5.01)	0.03	1.7 (0.69 to 4.65)	0.26
No impact factor	258 (52%)	40 (28%)	218 (63%)	0.35 (0.19 to 0.67)	0.001	0.52 (0.2 to 1.3)	0.16
Publisher†							
Big output	194 (40%)	66 (45%)	128 (37%)	–	–	–	–
Medium output	39 (8%)	32 (22%)	7 (2%)	8.37 (3.77 to 22.86)	<0.001	3.23 (1.23 to 11.33)	0.02
Small output	39 (8%)	17 (12%)	22 (7%)	1.5 (0.74 to 3.01)	0.25	0.36 (0.08 to 1.09)	0.09
Other	217 (44%)	30 (21%)	187 (54%)	0.31 (0.19 to 0.51)	<0.001	0.35 (0.17 to 0.63)	0.001
Gender of editor							
Men (reference)	403 (82%)	120 (83%)	283 (82%)	–	–	–	–
Women	63 (13%)	16 (11%)	47 (14%)	0.82 (0.43 to 1.46)	0.51	–	–
Both genders represented	23 (5%)	9 (6%)	14 (4%)	1.54 (0.62 to 3.55)	0.33	–	–
Country of editorial office‡							
North America (reference)	114 (24%)	51 (40%)	63 (19%)	–	–	–	–
Asia/Middle East/ Oceania	211 (45%)	22 (17%)	189 (56%)	0.15 (0.08 to 0.26)	<0.001	0.96 (0.4 to 2.25)	0.92
Europe	71 (15%)	16 (12%)	55 (16%)	0.37 (0.18 to 0.7)	0.003	0.48 (0.21 to 1)	0.06
UK	51 (11%)	37 (29%)	14 (4%)	3.2 (1.6 to 6.86)	0.001	1.82 (0.83 to 4.4)	0.15
Other country/region	21 (4%)	2 (2%)	19 (5%)	0.16 (0.02 to 0.53)	0.01	3.01 (0.3 to 19.02)	0.23
Income band of country of editorial office‡							
High income (reference)	270 (58%)	120 (94%)	150 (44%)	–	–	–	–
Upper middle income	88 (19%)	4 (3%)	84 (25%)	0.07 (0.2 to 0.16)	<0.001	0.12 (0.02 to 0.44)	0.002
Lower middle income/ low income	110 (23%)	4 (3%)	106 (31%)	0.05 (0.01 to 0.13)	<0.001	0.09 (0.02 to 0.27)	<0.001
Research domain							
General and internal medicine (reference)	110 (22%)	19 (13%)	91 (26%)	–	–	–	–
Surgery specialty	71 (15%)	23 (16%)	48 (14%)	2.27 (1.14 to 4.67)	0.02	1.47 (0.6 to 3.84)	0.4
Dentistry	30 (6%)	3 (2%)	27 (8%)	0.6 (0.12 to 1.82)	0.41	0.43 (0.04 to 2.02)	0.33
Medical specialty	220 (45%)	75 (52%)	145 (43%)	2.43 (1.41 to 4.47)	0.002	1.12 (0.53 to 2.51)	0.76
Pharmacology and pharmacy	23 (5%)	12 (8%)	11 (3%)	5.1 (2 to 13.84)	<0.001	2.47 (0.66 to 11.29)	0.19
Other specialty	35 (7%)	13 (9%)	22 (6%)	2.82 (1.2 to 6.6)	0.02	1.35 (0.42 to 4.4)	0.6

*Missing data for 2 journals: 1 journal without explicit data-sharing policy and 1 with explicit data-sharing policy.

†Missing data for 21 journals: 4 journals without explicit data-sharing policy and 17 with explicit data-sharing policy.

‡Journals that published over 15 journals in the medical domain: big output >1000 journals, medium output 250–1000 journals, small output <250 journals in publisher repertoire. Other: publishers that did publish under 15 journals in the medical domain.

aOR, adjusted Odds ratio.

Table 3 Characteristics of all published randomised controlled trials included

	ICMJE-member journals (n=100)	ICMJE-affiliated journals (n=100)
Data-sharing statement in article	98 (98%; 92% to 99%)	25 (25%; 17% to 35%)
Intentions to share individual patient data in statement		
Yes	67 (67%; 57% to 76%)	17 (17%; 10% to 26%)
No	21 (21%; 14% to 31%)	3 (3%; 0.1% to 9%)
Unclear	10 (10%; 5% to 18%)	5 (5%; 2% to 12%)
Not available	2 (2%; 0.3% to 8%)	75 (75%; 65% to 82%)
Type of registration		
Prospective	80 (80%; 71% to 87%)	50 (50%; 40% to 60%)
Retrospective	20 (20%; 13% to 29%)	22 (22%; 15% to 32%)
Unclear	–	28 (28%; 20% to 38%)
Registration of a data-sharing plan		
Yes	10 (10%; 5% to 18%)	8 (8%; 4% to 16%)
Yes, but not in original version	12 (12%; 7% to 20%)	5 (5%; 2% to 12%)
No	78 (78%; 68% to 85%)	87 (87%; 78% to 93%)

ICMJE, International Committee of Medical Journal Editors.

RCTs published in ICMJE-member journals

The results are displayed in [table 3](#) for the 100 selected articles. Among these, 30 were from *NEJM*, 28 from *The Lancet*, 17 from *JAMA*, 13 from *PLOS Medicine*, 5 from the *Annals of Internal Medicine*, 3 from *The BMJ*, 2 from the *Journal of Korean Medical Sciences* and 2 from the *German Medical Journal*.

Almost all the articles (98%; 92% to 99%) had a data-sharing statement. The two articles without data-sharing statements were from the *Journal of Korean Medical Science* and were confirmed as having been submitted after 1 July 2018. Of the statements 67% (57% to 76%) indicated an intention to share data, while the intention was unclear for an additional 10 (10%; 5% to 18%). The characteristics of the data-sharing plans are detailed in [table 4](#).

Of the 77 articles with data-sharing intentions, 7 (9%) mentioned access to other data from the study, besides IPD (eg, data frame for ‘unpublished data’/‘medical coding dataset’/‘non-patient-level data’). Of the 77 articles, 63 (82%) mentioned sharing for the following supplementary documents: study protocol (for 51), statistical analysis plan (for 37), informed consent form (for 15), data dictionary (for 14) and case report form (for 5). Time restriction for IPD was present in 34 of 77 (44%) data sets for either the start date of data-sharing, the end date or both. In two data-sharing statements, it was clear that data were to be available directly after approval of the drug in the European Union and in the USA. In 28 out of 30 other cases there was an embargo: 8 after 2 years, 2 after 18 months, 15 after 1 year, 1 after 9 months, and 2 after 3 months. A restricted access period was specified for 12 data sets: six of these specified restricted access for 2 years, three for 1 year, one for 5 years, one for 10 years,

and for one it was stated that the time would be defined by the committee. Of 77 data-sharing statements, 60 (78%) specified that data could only be used for specific reasons: 53 mentioned a scientific aim only, 6 indicated willingness to share data specifically for meta-analyses or individual meta-analyses, and 1 data-sharing statement specified that the aim of the reuse was to be focused on a particular disease (herpes zoster). A specific mechanism was detailed in 68 of 77 (88%) data-sharing statements. Thirty-five only mentioned the need to establish a data-sharing agreement and/or a formal data request, 20 indicated that an email contact was necessary, and 13 mentioned data-sharing platforms.

Twenty-two of 100 (22%) had registered a data-sharing plan on registers such as ClinicalTrials.gov. Of these, 14 specified IPD data-sharing, 6 did not, and for 2 it was unclear.

RCTs published in ICMJE-affiliated journals

The 100 selected RCTs were from 38 different journals (mean number of RCTs per journal=11 (±10)). We found 25 RCTs with data-sharing statements. Seventeen authors/teams (17%; 10% to 26%) declared an intention to share, while the intention was unclear for an additional five (5%; 2% to 12%). The characteristics of the data-sharing plans are detailed in [tables 3 and 4](#). Seven of 22 (32%) articles with a positive (or unclear) intention to share data expressed in a data-sharing statement indicated that other data sets, besides IPD, would be available. Regarding the sharing of any other documents, authors stated they would share study protocols (for seven studies), the statistical analysis plan and the study report (for four studies), and the case report form (for

Table 4 Characteristics of the data-sharing statements for articles with an intention to share individual patient data (including those with unclear intentions)

	Articles in ICMJE-member journals (n=77)	Articles in ICMJE-affiliated journals (n=22)
Intention to share code		
Yes	9 (12%; 6% to 22%)	0
No	17 (22%; 14% to 33%)	0
Unclear	51 (66%; 54% to 76%)	22 (100%)
Intention to share other research data		
Yes	6 (8%; 3% to 17%)	7 (32%; 15% to 55%)
No	14 (18%; 11% to 29%)	0
Unclear	57 (74%; 63% to 84%)	15 (68%; 45% to 85%)
Intention to share any other documents		
Yes	63 (82%; 71% to 89%)	10 (46%; 25% to 67%)
No	1 (1%; 0.1% to 8%)	0
Unclear	13 (17%; 10% to 28%)	12 (54%; 33% to 75%)
Restriction of time for availability		
Yes	34 (44%; 33% to 56%)	10 (46%; 25% to 67%)
No	22 (29%; 19% to 40%)	8 (36%; 18% to 59%)
Unclear	21 (27%; 18% to 39%)	4 (18%; 6% to 41%)
Free access		
Yes	7 (9%; 4% to 18%)	1 (4%; 0.2% to 25%)
No	69 (90%; 80% to 95%)	21 (96%; 75% to 99%)
Unclear	1 (1%; 0.1% to 8%)	0
Possibility to use data for any type of purpose		
Yes	7 (9%; 4% to 18%)	1 (4%; 0.2% to 25%)
No	60 (78%; 67% to 86%)	14 (64%; 41% to 82%)
Unclear	10 (13%; 7% to 23%)	7 (32%; 15% to 55%)
Specific kind of access mechanism		
Yes	68 (88%; 78% to 94%)	21 (96%; 75% to 99%)
No	7 (9%; 4% to 18%)	1 (4%; 0.2% to 25%)
Unclear	2 (3%; 0.5% to 10%)	0
Reviewed by a committee that is independent of the sponsor/author?		
Yes	18 (23%; 15% to 35%)	4 (18%; 6% to 41%)
No	22 (29%; 19% to 40%)	0
Unclear	37 (48%; 37% to 60%)	18 (82%; 59% to 94%)

ICMJE, International Committee of Medical Journal Editors.

two studies). Time restriction was present for 10 out of 22 (46%) data sets. Three data sets had a limitation for the start date of data-sharing, ranging from 12 months to 18 months and up to 3 years. For the end date of data availability, the following time frames were collected: 5 years in two cases, 3 years for one, 2 years for one, 1 year for three and 3 months for one. For the question as to whether data could be used for any type of purpose, 14 out of 22 (64%) eligible data sets were only available for specific purposes (ie, research). For 10 of these cases the scientific aim was mentioned but not detailed, and in four statements no aim was specified at all. A specific kind of access mechanism was cited in 21 of 22 (96%) statements.

Six of them mentioned a data-sharing agreement, one referred to a data platform, and fourteen data sets could be requested by email. For the 13 out of 100 (13%) trials with registration of their data-sharing plan, 2 planned to share the data, 6 did not, and for 5 it was unclear.

DISCUSSION

Statement of principal findings

In our survey we found that 57% of ICMJE-member journals had an explicit data-sharing policy on their website and that approximately a third of the ICMJE-affiliated journals had one. Slightly more than a third

of the member journals and most of the affiliated journals (around 56%) referred to the ICMJE guidelines without specifying a specific data-sharing policy. In addition, nearly 60% of the affiliated journals with an explicit policy had a less demanding policy than that required by the ICMJE. In contrast, the former ICMJE policy of trial registration was better implemented, with more than 71% of member journals and 66% of the affiliated journals explicitly requiring it as part of their policies.

For journals with a data-sharing policy, a data-sharing statement was frequent among member journals (98%), with rates of intention to share data of around 77%. These rates are in line with the intention to share previously reported in the *Annals of Internal Medicine*.¹⁴ In contrast, among ICMJE-affiliated journals with a data-sharing policy, data-sharing statements were not frequent (25%), and the intention to share data was only found in 22% of RCTs published in journals with an explicit data-sharing policy. Importantly, the statements often refer to data-sharing on request, and rarely to a specific repository or to fully available data sets. We already know that, even under a strict data-sharing policy such as the policy in place at *The BMJ* and *PLOS Medicine*, data availability is suboptimal, even when researchers express an intention to share.¹⁵ And indeed, in a recent scoping review¹⁶ we found that while the willingness to share data was generally high across trials, actual data-sharing rates were generally lower. In addition, there was considerable heterogeneity in data-sharing statements, with a focus on IPD data, and with very inconsistent information related to statistical codes and other documents (eg, the study protocol or the study report), which are key elements for reproducible research.¹⁷ Our results therefore question whether the new ICMJE policy as implemented by journals adequately supports clinical trial data-sharing, and they underline the need for efforts towards more reproducible research. Although data-sharing is only one aspect among others (eg, registration/best practices in reporting), without data-sharing, reproducibility is not possible.

Few characteristics were found to be associated with an explicit data-sharing policy. All were related to the publishers and the World Bank wealth category of the country/region of the editorial offices. As observed in previous research,¹⁸ a positive association between the JIF and data-sharing was found in univariate analysis, but it did not survive in the multivariate analysis. While the JIF is often (incorrectly) thought to be a surrogate for journal quality, our study suggests that professionalism and characteristics of the publisher and the editorial office resources could be better markers of quality and the implementation of reproducible research policies.

Findings in relation to other studies

A similar survey conducted in 2019 by our team also identified lack of implementation of basic data-sharing instructions in surgery journals with a JIF over 2. Only 50% of the journals had a data-sharing policy on their website,¹⁹

and in general these policies were not as demanding as those required by the ICMJE.

Furthermore, research done by Dal-Ré and Marušić²⁰ found independently an alike number of predatory journals in the list of journals which claim to follow the ICMJE recommendations.

It is important to note here that almost all the 'big' publishing houses have different data-sharing policies for their different journals (eg, the BMJ Group has three levels of data-sharing policy, and Taylor & Francis have five different types).^{21 22} A related analysis was conducted by Mellor²³ on the various data-sharing policies of the four big publishing houses, Elsevier, Springer Nature, Taylor & Francis and Wiley. This survey compared the Transparency and Openness (TOP) guidelines with the data-sharing policies of the different journals.²⁴ Similar definitions to the ones we used to define more or less stringent requirements were adopted, and the authors found that most of the basic or level 1 data-sharing policies were not even TOP-compliant. This confirms our impression that even if policies are in place, they are not sufficiently demanding to be liable to change the data-sharing culture. A harmonisation needs to be established, and previous successful experiences should be taken into account.^{25 26}

Limitations of this study

In our study, we had to rely on online information, as it proved difficult to contact editors and ascertain the existence of data-sharing policies. While all included journals had an electronic format, we cannot exclude that some may have implemented a policy without mentioning its enforcement explicitly on its website. In addition, we are studying a moving target in a changing environment, and it is likely that some journals that had no explicit policy when we performed our search have now implemented one. Repeated monitoring of the implementation of data-sharing policies therefore seems necessary. Another limitation was that we did not check specifically whether the data sets were actually made available when the authors indicated availability in the statement, nor did we request any data to ascertain data availability. Data availability rates could indeed be lower than suggested in an intention to share, as observed in *The BMJ* overall,²⁷ and more particularly for clinical trial data, even after communicating with the study authors.¹⁵ Moreover, it would be interesting to see how many funders or academic institutions really share their data after expiry of the time restrictions indicated.

A further limitation was the language filter. Due to lack of resources we were not able to include every language. This might have caused bias, as for instance Russian journals might have presented a different data-sharing policy from journals that publish in English.

A large range of journals were included, especially in terms of quality. We tried to limit the inclusion of 'predatory journals' using Beall's list. In this matter there is no real gold standard, as no exact definition existed when we



planned our study. Other lists such as Cabell's blacklist show an overlap with Beall's list.²⁸ Recently, a new definition has been proposed²⁹ and it could help to better identify predatory journals. On the one hand, we were surprised by the large number of ICMJE-affiliated journals referenced on the ICMJE website and listed in Beall's list. On the other hand, the investigators performing the data extraction had the impression that some of the selected journals had very poor editorial standards (in cases where instructions for authors were not clear and information was not given for all the steps of the editorial process) and could also fit the definition of predatory journals.

Finally, our identification of factors associated with an explicit data-sharing policy is only exploratory. Several unmeasured confounders, for instance the journal's income and/or the numbers of RCTs published by a given journal, could account for some of the associations found. Other unmeasured confounders may exist and great caution is warranted in interpreting these results, which naturally cannot be considered as reflecting any causal relationship.

Perspectives

It appears that data-sharing policies are infrequent and poorly enforced in most ICMJE-affiliated journals. Perhaps the journals do not know how best to implement the policy, or they may be worried they will lose submissions if the policies are implemented. Other explanations could be the costs resulting from the process or the greater labour intensity. It is also possible that some authors, researchers and indeed editors may be opposed to data-sharing policies. In addition, there is no specific enforcement for an affiliated journal to follow the ICMJE guidelines. It can be noted that the ICMJE states on its website that they 'cannot verify the completeness or accuracy of this list' and that 'there may be some listed journals that do not follow all of the many recommendations and policies in the document'.⁷ Furthermore, the large proportion of presumed predatory journals we found as well as the small proportion of journals enforcing the new policy are of concern for the impact and credibility of the ICMJE. We suggest that journals provide audits and feedback (to readers), especially as the number of ICMJE-affiliated journals is growing very fast, with 4725 in November 2018³⁰ and already 5504 in November 2019⁷ (+16% in 1 year). Without such checks, journals with poor editorial practices could present affiliation with the committee as an endorsement of a sort of quality label in biomedical journals, while this is not the case. The ICMJE affiliation could be indeed perceived as a guarantee, since these standards have of course been endorsed by more than three-quarters of the most prominent journals in biomedicine, as illustrated by Shamseer and colleagues in 2016.³¹

In addition, continuous audits of journal policies and their enforcement could be used as a better indicator of journal quality than the current exclusive focus on the JIF. There is room for development of new responsible metrics in this area, encompassing other aspects of reproducible research practices, such as registration policies and the use

of reporting guidelines.³² And indeed, data-sharing is only one facet of reproducible research policies.

Steps in the right direction have already been taken, such as the uniform guidelines for data-sharing in journals that have been developed by the Data Policy Standardisation and Implementation Interest Group of the Research Data Alliance.³³ This could help to reach the goal of full transparency and data-sharing for clinical trial results, since the implementation of the current ICMJE policy seems suboptimal.

Correction notice Size of the figure 1 has been increased.

Twitter Maximilian Siebert @ReiTheR_RCT and David Moher @dmoher

Acknowledgements We would like to thank Chloé Rousseau from CHU Rennes for her statistical advice and Clémence Belvéze from the university library Rennes 1 for her help with the search algorithm.

Contributors MS, JFG, AD, DM and FN conceived and designed the experiments. MS, JFG, LC and HG performed the experiments. MS and FN analysed the data. MS and FN interpreted the results. MS wrote the first draft of the manuscript. DM and FN contributed to the writing of the manuscript. MS, JFG, LC, HG, AD, DM and FN agreed with the results and conclusions of the manuscript. All authors have read them and confirm that they meet the ICMJE criteria for authorship. All authors had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis. MS is the guarantor.

Funding FN received funding from the French National Research Agency for the ReiTheR (Reproducibility in Therapeutic Research) project, which this article was part of (project number: ANR-17-CE36-0010).

Competing interests FN, MS and JFG received funding from the French National Research Agency for the submitted work.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. All data and all materials are available on the OSF website: <https://osf.io/n6whd/>.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Maximilian Siebert <http://orcid.org/0000-0003-4385-5773>

David Moher <http://orcid.org/0000-0003-2434-4206>

REFERENCES

- 1 Taichman DB, Sahni P, Pinborg A, *et al*. Data sharing statements for clinical trials: a requirement of the International Committee of medical Journal editors. *Ann Intern Med* 2017;167:63–5.
- 2 Laine C, Goodman SN, Griswold ME, *et al*. Reproducible research: moving toward research the public can really trust. *Ann Intern Med* 2007;146:450–3.
- 3 Groves T. Managing UK research data for future use. *BMJ* 2009;338:b1525
- 4 Godlee F, Groves T. The new BMJ policy on sharing data from drug and device trials. *BMJ* 2012;345:e7888.
- 5 Loder E, Groves T. The BMJ requires data sharing on Request for all trials. *BMJ* 2015;350:h2373.
- 6 Bloom T, Ganley E, Winker M. Data access for the open access literature: PLOS's data policy. *PLoS Biol* 2014;12:e1001797.
- 7 International Committee of Medical Journal Editors. Journals stating that they follow the ICMJE Recommendations, 2019. Available: <http://www.icmje.org/journals-following-the-icmje-recommendations/> [Accessed 9 Dec 2019].
- 8 Beall J. Beall's list of predatory journals and publishers, 2020. Available: <https://beallslist.weebly.com/> [Accessed 15 Jan 2020].

- 9 R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2018.
- 10 Open Science Framework. Maximilian Siebert, 2020. Available: <https://osf.io/b5nq7/>
- 11 The World Bank. World bank country and lending groups, 2020. Available: <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups> [Accessed 28 Jan 2020].
- 12 Lefebvre C, Manheimer E, Glanville J. Searching for studies. In: *Cochrane Handbook for systematic reviews of interventions*. The Cochrane Collaboration, 2011. www.training.cochrane.org/handbook
- 13 Kosmidis I. *brglm: bias reduction in Binary-Response generalized linear models*. R package version 0.6.2 ED, 2019. <https://cran.r-project.org/package=brglm>
- 14 Laine C, Berkwits M, Mulrow C, et al. *Reproducible research: biomedical researchers' willingness to share information to enable others to reproduce their results*. Vancouver: International Congress on Peer Review and Scientific Publication, 2009.
- 15 Naudet F, Sakarovitch C, Janiaud P, et al. Data sharing and reanalysis of randomized controlled trials in leading biomedical journals with a full data sharing policy: survey of studies published in *The BMJ* and *PLOS Medicine*. *BMJ* 2018;360:k400.
- 16 Ohmann C, Moher D, Siebert M, et al. *Status, use and impact of sharing individual participant data from clinical trials: a scoping review*, 2019.
- 17 Munafò MR, Nosek BA, Bishop DVM, et al. A manifesto for reproducible science. *Nat Hum Behav* 2017;1.
- 18 Vasilevsky NA, Minnier J, Haendel MA, et al. Reproducible and reusable research: are Journal data sharing policies meeting the mark? *PeerJ* 2017;5:e3208.
- 19 Lombard N, Gasmi A, Sulpice L. Research transparency promotion by surgical journals publishing randomised controlled trials: a survey. *medRxiv* 2019:19002386.
- 20 Dal-Ré R, Marušić A. Potential predatory journals are colonizing the ICMJE recommendations list of followers. *Neth J Med* 2019;77:92–6.
- 21 BMJ Author Hub. Data sharing, 2018. Available: <https://authors.bmj.com/policies/data-sharing/> [Accessed 24 Mar 2020].
- 22 Author Services. Understanding our data sharing policies, 2020. Available: <https://authorservices.taylorandfrancis.com/data-sharing-policies/> [Accessed 24 Mar 2020].
- 23 Mellor D. The landscape of open data policies, 2018. Available: <https://cos.io/blog/landscape-open-data-policies/>
- 24 Nosek BA, Alter G, Banks GC, et al. Promoting an open research culture. *Science* 2015;348:1422–5.
- 25 Hardwicke TE, Mathur MB, MacDonald K, et al. Data availability, reusability, and analytic reproducibility: evaluating the impact of a mandatory open data policy at the journal *Cognition*. *R Soc Open Sci* 2018;5:180448.
- 26 Giofrè D, Cumming G, Fresc L, et al. The influence of Journal submission guidelines on authors' reporting of statistics and use of open research practices. *PLoS One* 2017;12:e0175583.
- 27 Rowhani-Farid A, Barnett AG. Has open data arrived at the British medical Journal (BMJ)? an observational study. *BMJ Open* 2016;6:e011784.
- 28 Strinzel M, Severin A, Milzow K, et al. Blacklists and Whitelists to tackle predatory publishing: a cross-sectional comparison and thematic analysis. *MBio* 2019;10:mBio.00411-19.
- 29 Grudniewicz A, Moher D, Cobey KD, et al. Predatory journals: no definition, no defence. *Nature* 2019;576:210–2.
- 30 International Committee of Medical Journal Editors. Journals stating that they follow the ICMJE recommendations, 2018. Available: <http://www.icmje.org/journals-following-the-icmje-recommendations/> [Accessed 9 Dec 2019].
- 31 Shamseer L, Hopewell S, Altman DG, et al. Update on the endorsement of CONSORT by high impact factor journals: a survey of journal "Instructions to Authors" in 2014. *Trials* 2016;17:301.
- 32 Moher D, Naudet F, Cristea IA, et al. Assessing scientists for hiring, promotion, and tenure. *PLoS Biol* 2018;16:e2004089.
- 33 Hrynaskiewicz I, Simons N, Hussain A. Developing a research data policy framework for all journals and publishers. *Data Sci J* 2019;19:dsj-2020-005.