Journal of Clinical Epidemiology 116 (2019) 26–35

ORIGINAL ARTICLE

Using the STROBE statement: survey findings emphasized the role of journals in enforcing reporting guidelines

Melissa K. Sharp[a,b,*], Lorenzo Bertizzolo[b], Roser Rius[c], Elizabeth Wager[d,e], Guadalupe Gómez[c], Darko Hren[a]

[a]*Faculty of Humanities and Social Sciences, Department of Psychology, University of Split, Split, Croatia*
[b]*Université de Paris, CRESS, INSERM, INRA, F-75004 Paris, France*
[c]*Departament d'Estadística i Investigació Operativa, Universitat Politècnica de Catalunya-BarcelonaTech, Barcelona, Spain*
[d]*School of Medicine, University of Split, Split, Croatia*
[e]*Sideview, Princes Risborough, UK*

Accepted 30 July 2019; Published online 6 August 2019

Abstract

Objectives: The objective of the study was to identify factors affecting the use of the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement, specifically authors' attitudes toward and experiences with it.

Study Design and Setting: An online survey was distributed to authors of observational studies recruited via social media, personal network snowballing, and mass mailings using targeted search strategies. Data on demographics, awareness, motivators, and usage were collected in conjunction with a modified Unified Theory of Acceptance and Use of Technology (UTAUT) scale on which confirmatory factor analysis (CFA) was performed.

Results: One thousand fifteen participants completed the survey. Of these, 185 (18.2%) indicated they had never heard of STROBE nor used it previously, 195 (19.2%) had heard of it but never used it, and 635 (62.6%) had used it. Journals promoting STROBE were both key motivators and awareness mechanisms; peers and educational workshops were also important influencing factors to a lesser degree. The internal consistency of the modified UTAUT scale was strong (Cronbach's alpha = 0.94). CFA supported a four-factor model with 23 questions.

Conclusion: The endorsement of STROBE by journals is key to authors' awareness and use of the guideline. We tested and validated our scale which can guide future research on reporting guidelines.

Keywords: Observational studies; Guidelines as topic; Epidemiologic research design; Information dissemination/methods; STROBE; Online survey; Scientific writing

Authors' contributions: All authors have made substantive intellectual contributions to the development of the protocol and this manuscript. M.K.S. conceptualized the study and led the writing of the manuscript. D.H. led the supervision of the manuscript preparation. M.K.S. and L.B. managed survey recruitment. M.K.S., G.G., and R.R. assisted with all analyses. M.K.S. performed analyses. All authors read and approved the final manuscript.

Source of funding: This work was supported by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Sklodowska-Curie grant agreement No 676207. Coauthor G.G. was supported by work from Ministerio de Economía y Competitividad, Grant/Award Numbers: MTM2015-64465-C2-1-R, MDM-2014-0445.

Ethics approval and consent to participate: Ethical approval was granted by the University of Split (2181-198-03-04-18-0010).

Availability of data and material: The final R Markdown code used for the current study will be made available on the Open Science Framework (https://osf.io/2fkny/) and in the Zenodo repository in the Methods in Research on Research (MiRoR) community (https://zenodo.org/communities/miror/).

Conflict of interest: M.K.S. works with the STROBE statement as a part of her doctoral studies. D.H., G.G., and E.W. provide support and mentoring as a part of the Methods in Research on Research (MiRoR) project. E.W. was a Fellow of the UK EQUATOR Centre which promotes the use of reporting guidelines (this was an unpaid position).

* Corresponding author. University of Split, Université de Paris, Poljička cesta 35, Split 21000, Croatia. Tel.: +1 248 701 8109.

E-mail addresses: msharp@unist.hr; melissa.sharp@etu.parisdescartes.fr (M.K. Sharp).

https://doi.org/10.1016/j.jclinepi.2019.07.019

What is new?

Key findings

- We used a health technology assessment framework to investigate authors' attitudes and experiences of the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline. Our survey captured over 1,000 authors' views on the topic.
- Respondents reported STROBE to be useful, easy to use, clear, and understandable. However, they were concerned about the time needed to use the tool and expressed apprehension about how supportive peers and the research environment were toward using STROBE. Nearly 200 respondents were introduced to STROBE for the first time, making the survey an educational intervention itself.

What this adds to what was known?

- We have tested and validated a scale to assess author interactions with and views toward STROBE. This scale can be used to inform the promotion and evaluation of other reporting guidelines.

What is the implication and what should change now?

- Journals are key to raising awareness of reporting guidelines and enforcing their use. The research climate surrounding authors (i.e., peers and educational workshops) is also an important secondary influencing factor. Interventions should focus on establishing incentive systems and a culture change with these actors.

1. Introduction

Reporting guidelines (RGs) provide a protective "cognitive net" against the fallibility of human memory and support the skills of expert professionals [1]. Authors of biomedical manuscripts are generally unaware of the existence or utility of RGs and those responding to peer reviewers often have problems adhering to the methodological standards proposed [2–4]. Many journals do not require a relevant RG checklist to be submitted with a manuscript therefore, there is often no incentive for authors to complete one [5].

Some authors reject RGs, claiming that RGs can be condescending and rigid [6,7]. It is unclear what maintains these attitudes. Therefore, it would be useful to understand factors affecting use [8]. This study was designed to explore how researchers view and interact with one RG, the Strengthening the Reporting of Observational studies in Epidemiology (STROBE) Statement. STROBE was created in 2007 to improve the reporting of observational studies (e.g., cross-sectional, cohort, case-control).

Many journals promote STROBE by requiring or recommending its use during the manuscript preparation process. However, endorsement rates are relatively low [9–13], and there is a diffusion of responsibility among journal editors, authors, and peer reviewers for RG compliance [3]. To better understand the current situation facing authors, we aimed to identify the personal and environmental facilitators, barriers, and motivators to using the STROBE statement. With this information, we hoped to extend the practical value of STROBE and perhaps other RGs.

2. Methods

2.1. Survey design

We followed the CHERRIES guideline for online surveys (Supplemental File 1) [14]. Before distribution, we piloted the survey within the Methods in Research on Research (MiRoR) network [15], allowing collaborators to give feedback on content and functionality [16]. The University of Split School of Medicine Ethical Review Committee granted ethical approval.

The survey flow is presented in Figure 1, and the survey is in Supplemental File 2. All questions were forced response except for one optional open-ended question and mistakenly, the question asking about the respondent's country. After consenting to participate, adaptive questioning branched the survey based on participant's level of awareness and use of STROBE (i.e., never heard of, never used; heard of, never used; heard of, have used). After branching, participants were presented with questions about their interactions with STROBE (e.g., real or theoretical timing of use: writing a grant or peer-reviewing an article).

Next, all participants were presented 25 questions informed by the Unified Theory of Acceptance and Use of Technology (UTAUT) scale [17,18]. UTAUT is an amalgamation of eight dominant psychological and health technology assessment (HTA) theories and models that attempts to explain one's intention to use a piece of technology and their subsequent use behavior. The scale aims to explain information system usage behavior by measuring: Performance Expectancy (PE), Effort Expectancy (EE), attitude toward using technology, Social Influence (SI), Facilitating Conditions (FCs), self-efficacy, anxiety, and behavioral intention to use the tool [17]. HTA systematically evaluates direct and indirect consequences of using a piece of health technology. It can tap into whether the technology works, for whom, and at what cost [19].

We rephrased questions to be relevant to STROBE and kept the scale's four core constructs (PE, EE, SI, and FCs) (Fig. 2). Each subscale contained several items to ensure reliability and validity. The final version contained nine Likert scale items from PE, six from EE, five from EE, four from FCs, and one assessing the intention to use STROBE. Respondents rated statements on seven-point

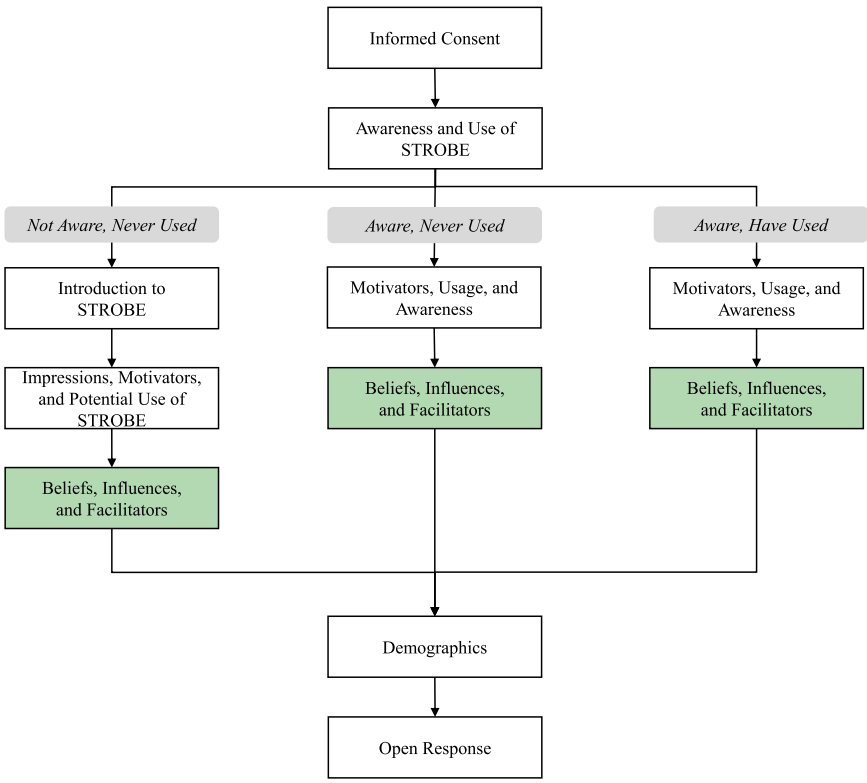

Fig. 1. Survey flow. STROBE, Strengthening the Reporting of Observational Studies in Epidemiology.

Likert-type scales from "strongly disagree" to "strongly agree" (Supplemental File 2).

2.2. Recruitment

Eligible participants were researchers involved in manuscript writing (within the past 10 years) reporting the results of observational studies. The survey was distributed from March 5 to August 31, 2018.

Survey recruitment used several snowball and purposive sampling routes. First, M.K.S. invited her professional network and those involved in the MiRoR consortium [15] to participate. Next, the survey was promoted through social media, primarily Twitter. We then emailed the editors of 257 biomedical journals identified in another study [9,20] and asked them to invite their authors to participate (e.g., via e-mail list-servs, Twitter, LinkedIn, and so forth). Up to three e-mails were sent if they did not respond. When initial recruitment methods failed to provide sufficient respondents, we used Python to scrape emails of corresponding authors from an observational study corpus which examined endorsement of seven STROBE extensions [9,20,21]. To broaden the scope, we also included other journals primarily focused in Epidemiology. We identified 75 English language journals from the "Epidemiology" Broad Subject Term in the National Library of Medicine [22], 122 endorsing journals from the STROBE statement website [23], and 98 top-ranked journals in the Scimago Journal and Country 2017 "Medicine" ranking [24]. We ran an Ovid MEDLINE observational study search filter from the same previous study [20] on all journals, deleted nonrelevant publication types (e.g., case summaries, editorials), and restricted the search to English language articles published within the past year (to reduce bounced emails). Supplementary File 3 details search strategies and journals searched. We deduplicated e-mails and sent up to two emails to each author.

2.3. Statistical analyses

General information on demographics, STROBE extension awareness, research stage usage, and awareness referral mechanisms is presented as counts and percentages

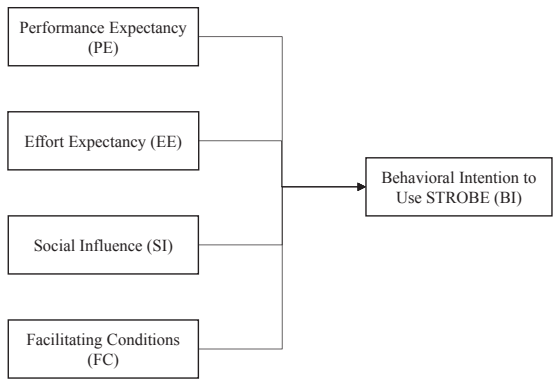

Fig. 2. Model and domain definitions. STROBE, Strengthening the Reporting of Observational Studies in Epidemiology.

Table 1. Sample demographics

Demographics	Total sample *N* (%) 1,015 (100)	Never heard of STROBE, never used (group 1) *N* (%) 195 (19)	Heard of STROBE, never used (group 2) *N* (%) 185 (18)	Heard of STROBE, have used (group 3) *N* (%) 635 (62)
Time spent in research				
1–10 years	332 (33)	57 (29)	65 (35)	210 (33)
11–30	362 (36)	107 (55)	95 (51)	372 (59)
31+	86 (10)	30 (15)	25 (14)	48 (8)
I do not work in research	3 (<1)	1 (0)	0 (0)	2 (<1)
Prefer not to say	3 (<1)	0 (0)	0 (0)	3 (<1)
Age				
18–34	185 (18)	36 (19)	38 (21)	111 (1)
35–54	589 (58)	101 (52)	83 (45)	405 (64)
55+	235 (23)	58 (30)	64 (35)	113 (18)
Prefer not to say	6 (<1)	0 (0)	0 (0)	6 (<1)
Gender				
Woman	469 (46)	97 (50)	82 (44)	289 (46)
Man	525 (52)	94 (48)	101 (55)	329 (52)
Trans	3 (<1)	0 (0)	0 (0)	3 (<1)
Prefer not to say	20 (2)	4 (2)	2 (1)	14 (2)
Region				
Africa	22 (2)	5 (3)	2 (1)	15 (2)
Asiatic region	31 (3)	7 (4)	4 (2)	20 (3)
Eastern Europe	33 (3)	12 (6)	5 (3)	16 (3)
Latin America	54 (5)	14 (7)	10 (5)	30 (5)
Middle East	26 (3)	11 (6)	6 (3)	9 (1)
Northern America	283 (28)	58 (30)	57 (31)	168 (27)
Pacific Region	54 (5)	4 (2)	10 (5)	40 (6)
Western Europe	465 (46)	69 (35)	83 (45)	313 (49)
Not reported	47 (5)	15 (8)	8 (4)	24 (4)

Abbreviation: STROBE, Strengthening the Reporting of Observational Studies in Epidemiology.

in the aggregate and per subgroup. Likert scale responses are reported as means and standard deviations. Completion/dropout rates were calculated overall and per group based on completion of the final forced-response question.

As we used a modified UTAUT scale (Table 3), we had a priori assumptions about our model and its latent factors (Fig. 1). Essentially, we were testing the HTA theory in our setting. Thus, confirmatory factor analysis (CFA) was used to test a four-factor model of intention to use STROBE in the overall sample and subgroups (Fig. 1). Rather than simply comparing average attitudes between groups, CFA allows us to test a theory and whether we captured relevant indicators and how they relate to each other (e.g., that we captured the key influencing factors that affect one's likelihood to use STROBE and furthermore, that we are comprehensive with our questioning and not redundant). All questions were scored from 1 to 7 and treated as continuous variables (Supplemental File 2). Three negatively worded questions from the EE scale were reverse-coded before calculating Cronbach's alpha and conducting CFA (see Table 4). For judging internal consistency, or the estimate of the reliability indicating the degree to which items measure different aspects of the same concept, we used Cronbach's alpha and considered $\geqq 0.7$ an acceptable value [25,26].

All analyses were performed in R, version 3.4.0. The R Markdown file, containing code and output, is available on Open Science Framework [27]. The model was fit using lavaan, version 0.6–3 [28]. Maximum likelihood estimation (MLM specification) with robust standard errors was used to account for non-normality sample variance-covariance matrices and provide scaled test statistics. Latent factors were standardized, allowing for free estimation of all factor loadings. As suggested by Hu and Bentler [29], we considered Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) values $\geq$ 0.90 for acceptable and $\geq$0.95 for good fit, root mean squared error of approximation (RMSEA) values $\leq$ 0.06 (poor fit $>$ 0.10), and standardized root mean squared residual (SRMR) values $\leq$ 0.08 to indicate a good fit between the model and data. When conducting multiple-group CFA, convergence issues are common [30]. When they occurred, we investigated the model within subgroups to detect issues with modification indices, individual factor loadings, and covariances between latent factors.

Table 2. Motivators, usage, and awareness descriptives

Question	Never heard of STROBE, never used; group 1, $n = 195$	Heard of STROBE, never used; group 2, $n = 185$	Heard of STROBE, used; group 3, $n = 635$
Motivator of use (past/theoretical)[a,b]			
Self	128 (66)	55 (30)	308 (49)
Coauthors	57 (29)	72 (39)	116 (18)
Mentor/supervisor encouraged	40 (21)	-	-
Social norm	44 (23)	-	-
Journal submission process	104 (53)	134 (72)	376 (59)
Journal peer review	82 (42)	90 (49)	77 (12)
Incentivized in workplace	-	28 (15)	-
Immediate feedback	-	35 (19)	-
Free text	8 (4)	8 (4)	44 (7)
Reasons for not using[b]			
My writing would not benefit	-	26 (14)	-
Do not want strict rules	-	25 (14)	-
Hard to understand	-	11 (6)	-
Word count	-	20 (11)	-
Format is difficult	-	16 (9)	-
Coauthors do not use it	-	45 (24)	-
May result in more negative review	-	1 (<1)	-
Journals do not require it	-	98 (53)	-
Not applicable for study type	-	25 (14)	-
Other	-	35 (19)	-
Frequency of current use			
Do not currently use	-	-	5 (<1)
Less than a quarter of manuscripts	-	-	305 (48)
Roughly half of manuscripts	-	-	134 (21)
Roughly 75% of manuscripts	-	-	118 (19)
All applicable manuscripts	-	-	73 (12)
Research stage of use (past/theoretical)[b]			
Did not consider	22 (11)	10 (5)	NA
Protocol/design stage	126 (65)	110 (60)	239 (38)
Grant	76 (39)	61 (33)	89 (14)
Manuscript	122 (63)	126 (68)	451 (71)
After completing the article to check	98 (50)	80 (43)	439 (69)
Evaluating the article	70 (36)	69 (37)	243 (38)
Awareness mechanism			
Peer/colleague	-	31 (17)	97 (15)
Boss/mentor/supervisor	-	16 (9)	90 (14)
Journal	-	55 (30)	234 (37)
Course/workshop	-	32 (17)	105 (17)
Online	-	37 (20)	66 (10)
Other	-	14 (8)	43 (7)

[a] Columns/items are blank as not all questions were presented to all branches.
[b] Question allowed for multiple responses to be selected.

3. Results

Of the 257 editorial offices contacted, 65 (25.3%) responded after three attempts. Of those who responded, 20 (30.8%) reported that they would invite their authors to participate (via Twitter, LinkedIn, listserv, blog, etc.), 42 (64.6%) declined to participate, and 3 (4.6%) reported individual-level participation. Reasons for declining included no access to a list, no time, a desire to remain neutral, the inability to contact authors due to General Data Protection Regulation restrictions [31], a belief that the journal did not publish observational studies (although we contend that it did), and a belief that the survey was flawed.

Table 3. Comparisons of model fit

Model	Group (*n*)	χ^2	df	RMSEA (90% CI)[a]	SRMR[b]	TLI[c,d]	CFI[c,d]	AIC
Four-factor model	Overall (1,015)	–	–	–	–	–	–	–
	1 (195)	776.900	266	0.109 (0.100, 0.118)	0.087	0.818	0.838	12593.900
	2 (185)	730.552	266	0.108 (0.099, 0.117)	0.085	0.797	0.802	12305.731
	3 (635)	1,582.699	266	0.102 (0.097, 0.107)	**0.077**[b]	0.813	0.834	42,959.805
Final model: Four-factor model, No FC3, Method effects[e]	Overall (1,015)	1,931.539	717	0.078 (0.074, 0.082)	**0.072**[b]	0.895	**0.909**[d]	64,488.934
	1 (195)	489.527	239	0.079 (0.069, 0.089)	**0.077**[b]	**0.904**[d]	**0.917**[d]	11,915.867
	2 (185)	496.303	239	0.084 (0.074, 0.095)	**0.075**[b]	0.877	0.894	11,734.401
	3 (635)	927.172	239	0.076 (0.071, 0.081)	**0.070**[b]	0.897	**0.911**[d]	40,838.666

Abbreviations: χ^2, chi-squared; df, degree of freedom; RMSEA, root mean square error of approximation with 90% confidence intervals; SRMR, square root mean residual; TLI, Tucker-Lewis Index; CFI, Comparative Fit Index; AIC, Akaike Information Criterion.

Bold = reached pre-established cutpoint threshold.

[a] Within the range ($\leq$0.06), indicating a good fit between the model and the data.

[b] Within the range ($\leq$0.08), indicating a good fit between the model and the data.

[c] Within the range ($\geq$0.95), indicating a good fit between the model and the data.

[d] Within the range (0.90 $\leq$ x $\leq$ .95), indicating an acceptable fit between the model and the data.

[e] Method effects address the reverse-coded items EE4, EE5, and EE6 and the high covariance between PE4 and PE5.

After accounting for 2,304 invalid addresses, 14,621 e-mails were sent to authors; we sent a second reminder e-mail to nonrespondents. Over 100 authors (n = 109) informed us of participation, 23 declined (giving no reason or stating no time/interest), and 23 reported ineligibility (i.e., did not work in observational research). Another 145 were ineligible as they were unreachable during the recruitment period (e.g., family, sick, sabbatical leaves) or unreachable permanently (e.g., left job, retired, died).

As the survey was anonymous and recruitment methods used network snowballing, social media, and mass mailings, we cannot estimate the total number of people that read the survey invitation. However, we know that 1,293 visitors read the informed consent page and 1,265 (97.8%) agreed to participate. After evaluating free-text responses, seven indicated ineligibility (e.g., "I do not do observational research"). Of the 1,258 eligible participants, 1,015 (80.7%) completed the survey. Nearly 20% (n = 195) indicated they never heard of STROBE nor used it before the study (group 1), 18% (n = 185) had heard of it but never used it (group 2), and over half (63%, n = 635) had heard of and used it (group 3) (Fig. 3, Table 1). The completion rates were 67% for group 1, 81% for group 2, and 97% for group 3 (Fig. 3).

We found fairly equal distributions for demographic categories across groups (Table 1, Supplemental File 4). The top five countries responding were the United States (21.6%), United Kingdom (9.8%), Italy (6.8%), Canada (6.4%), and Australia (4.9%). To account for the multidisciplinary nature of research, we allowed up to three selections for area of work. Participants working in public health and epidemiology were well represented with 470 (46%) and 247 (24%), respectively, choosing those options as one of their primary fields of work.

3.1. Motivators of use

When asked about what factors would or have motivated use of STROBE, the journal submission process and mandatory RG use were the most frequently chosen options. After journal policies, self-motivation was among the top-ranked reported influences. Around half (53%) of those who were aware of STROBE but had not used it (group 2) reported that this was because journals did not require it. The next most frequently reported reason was that their coauthors did not use it (24.3%) (Table 2).

3.2. Usage timing and frequency

Participants who used STROBE (group three) most commonly did so during the manuscript writing process (n = 451) or after completing their draft to check that all relevant information had been reported (n = 439). Participants who had not used STROBE before (groups one and two; n = 380) most frequently reported that they would most likely use it during the manuscript writing process (62.6%; 68.1%) or during the protocol/study design stage (64.6%; 59.5%). For those who previously used STROBE (group 3), 48% used it for less than a quarter of their manuscripts, whereas 11.5% used it for all of their manuscripts (Table 2).

3.3. Awareness

Of those who were aware of STROBE before the survey (groups 2 and 3, n = 820), the most frequently reported route that made them aware of STROBE was a journal requiring or recommending it (group 2: n = 55, 29.7%; group 3: n = 234, 36.9%). The other options (peers, superiors, courses, or online) ranged from 12.6 to 16.7%. A

Table 4. Factor loadings of final model (n = 1,015)

Items		Performance Expectancy (PE)			Effort Expectancy (EE)			Social Influence (SI)			Facilitating Conditions (FC)		
		1[a]	2[b]	3[c]	1[a]	2[b]	3[c]	1[a]	2[b]	3[c]	1[a]	2[b]	3[c]
PE1	STROBE will be/is useful in my job	0.728	0.710	0.665									
PE2	Using STROBE will enable/enables me to write papers more quickly	0.868	0.821	0.818									
PE3	STROBE will increase/increases my productivity	0.865	0.817	0.775									
PE4	If I use STROBE, I (will) increase my chances of getting published	0.766	0.699	0.511									
PE5	If I use STROBE, I will get a more positive peer review of my paper	0.737	0.670	0.553									
PE6	Using STROBE will make/makes it easier for me to write papers	0.903	0.861	0.864									
PE7	Using STROBE will improve/ improves the quality of my manuscripts	0.803	0.675	0.767									
PE8	Using STROBE will make/makes my manuscript writing more efficient	0.849	0.867	0.870									
PE9	Using STROBE increases the quality of my output for the same amount of effort	0.850	0.831	0.804									
EE1	I think STROBE will be/is easy to use				0.841	0.771	0.887						
EE2	I think STROBE's content is clear and understandable				0.869	0.833	0.866						
EE3	I think that it will be/is easy for me to become skillful at using STROBE				0.793	0.797	0.693						
EE4	Using STROBE will take/takes too much time compared with my normal writing process*				0.437	0.464	0.604						
EE5	STROBE is so complicated, it will be/is difficult to understand what to do*				0.622	0.579	0.671						
EE6	Will take/takes too long to learn how to properly use STROBE to make it worth the effort*				0.569	0.542	0.598						
SI1	My peers will think/think that I should use STROBE							0.848	0.909	0.870			
SI2	My superiors will think/think that I should use it							0.621	0.562	0.639			
SI3	The research climate is helpful in promoting the use of reporting guidelines such as STROBE							0.887	0.890	0.831			
SI4	In general, I think that journals will support/support the use of STROBE							0.649	0.461	0.473			
SI5	I will use STROBE because a lot of scientists in my field are using it							0.531	0.553	0.549			
FC1	I have the knowledge necessary to use STROBE										0.599	0.490	0.567
FC2	STROBE is compatible with my current workflow										0.785	0.777	0.817
FC4	Using STROBE fits well with the way I like to work										0.852	0.878	0.843

Abbreviation: STROBE, Strengthening the Reporting of Observational Studies in Epidemiology.

1[a] Subgroup one: Never heard of STROBE and never used it (n = 195).

2[b] Subgroup two: Heard of STROBE but never used it (n = 185).

3[c] Subgroup three: Heard of STROBE and have used it (n = 635).

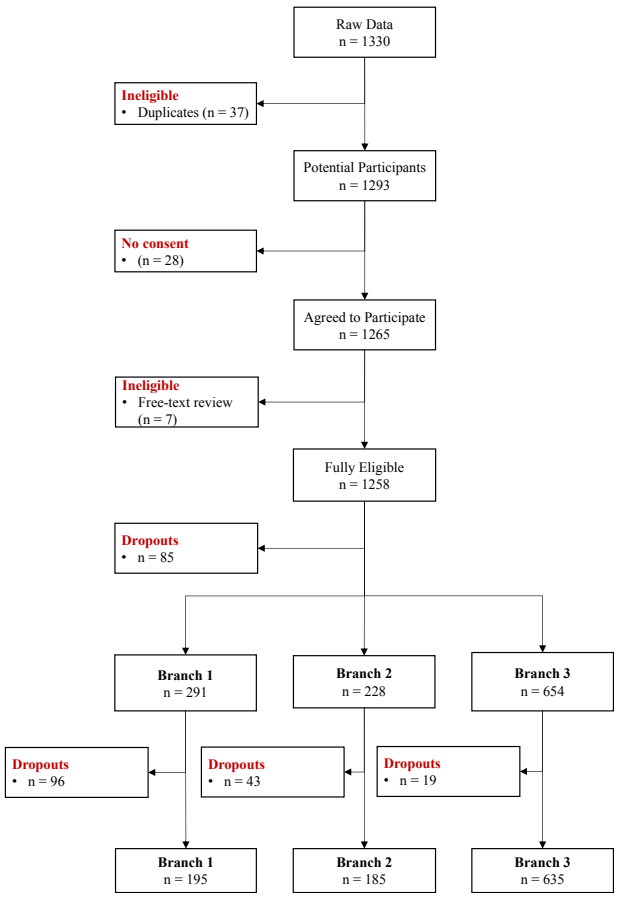

Fig. 3. Participant flow diagram.

majority of participants (70.7%, $n = 718$) indicated that they were not aware of any STROBE extension.

3.4. Confirmatory factor analysis on modified UTAUT scale

We attempted CFA on the overall sample, but it would not converge. Therefore, we investigated the model within subgroups to identify convergence issues; it converged in all subgroups (Table 3). There were three recurring issues across groups: (1) the third FC item (FC3) appeared to not belong to the FC scale; (2) the FC and EE covariance was very high (0.88–0.91); and (3) two pairs of items (EE4:EE5 and PE4:PE5) had significant shared variance, with the highest modification indices across all subgroups.

Model fit statistics and accompanying step-by-step descriptions are in Supplemental File 4, Table 6. Succinctly, the FC3 item phrasing was redundant with EE1. When FC3 was deleted, the model converged. This also reduced the high covariance between the FC and EE factors. The EE4 and EE5 items, along with EE6, were negatively worded, so we allowed them to covary to account for method effects [32]. Items PE4 and PE5 were also allowed to covary as they were both related to academic publishing, suggesting that they could covary for reasons other than the shared influence of the latent factor.

The four-factor model addressing these issues was the best fit model for our data (Table 3). The CFI (0.91) and TLI (0.90) reached the "acceptable" cut point of 0.90. The SRMR (0.07) was below its cut point of 0.08. The RMSEA (0.08) was not less than 0.06, however. All factor loadings were statistically significant (all $ps \leq 0.001$) and salient (0.437 to 0.909) (Table 4). The internal consistency reliability of all four subscales was strong (Cronbach's alpha ≥ 0.94 for all). Our items were parsimonious, functional, and internally consistent.

An overall pattern between groups was seen where those who had used STROBE before (group 3) had the highest scores, those who had never heard of STROBE before the survey (group 1) had second highest scores, and those who had heard of STROBE but never used it (group 2) most often had the lowest scores (Additional File 4, Table 4; Fig. 3).

4. Discussion

To the best of our knowledge, this is the first project to ask authors about their attitudes toward and experiences with STROBE, especially using HTA framework. Our project used a broad and multifaceted sampling strategy which created a diverse sample of observational study authors. We also engaged nearly 200 participants who previously had never heard of STROBE, making our survey an awareness intervention itself.

The large sample enabled us to test a modified UTAUT scale on our entire sample and within subgroups. With an acceptable fit between our model and the data, we expect that this instrument may be useful for evaluating interactions with other RGs. Our results confirm the applicability of an HTA approach to RGs, reveal important factors impacting STROBE use, and highlight a unique additional aspect of use, which may separate it from other pieces of technology—the academic publishing environment. Because our model-data fit was only "acceptable" and we needed to address shared error variance of two publishing-related items (PE4 and PE5), we believe that these two PE items might signal an unaccounted latent factor related to publishing.

Our CFA should be considered complementary to the descriptive results which emphasize the key role that journals have in raising awareness, motivating, and enforcing use. Journals were the most typical medium by which participants originally became aware of STROBE. Moreover, journals not requiring STROBE were the top reason why authors did not use it. We recommend that future work explores this concept more deeply. We suggest building on the most parsimonious model (Table 3), not including FC3, accounting for method effects on the EE scale, and addressing the shared error variance of PE4 and PE5.

With regard to limitations, estimating a sample size was not tenable as there is no clearly defined participant pool. In

addition, we used mass mailings, thus, potentially, some e-mails were likely blocked by spam filters [33]. In addition, we had differential dropout rates between groups which is expected as is conceptually harder to think in theoretical terms (e.g., when would you consider using STROBE which you were just introduced to vs. when have you used it). The differences in participation rates between groups and nonresponse and self-selection biases also could have skewed our responses to be more positive toward STROBE. In addition, the introduction to STROBE may have not been detailed enough and/or the participants may have not spent enough time on it.

Despite these limitations, overall, participants reported positive views toward STROBE, considering it useful, clear, and relatively easy to use. They also thought it would increase manuscript quality and the chances of getting published. However, they were not as positive regarding time requirements, reporting effects on productivity and speed and ease of writing. Our results should be reassuring to journal editors who fear losing authors to other journals with less-strict requirements for publication [3]. These fears may be unfounded as participants indicated that, despite time costs, there are benefits of using an RG such as increasing the quality of their manuscripts and the chances of being published. Furthermore, they thought that the publishing environment (i.e., journals) would or do support its use.

Despite this perceived benefit of an increase in quality, we caution that, empirically speaking, the research in this regard is mixed. Recent work (2019) demonstrated that having a methodological reviewer dedicated to looking for missing RG items (not only STROBE) increased the number of citations that an article received by 43% [34]. This could be perceived as a proxy for higher quality or impact. Conversely, other authors have found no effect on the reporting of confounding [35] or insufficient evidence to determine an impact on overall completeness of reporting [4,36]. To further assuage editors' (and authors') concerns, more research is needed in this area which focuses on a broad range of journals (i.e., not only high impact) and which takes endorsement type (i.e., requiring vs. recommending use) into account.

A 2019 scoping review complements our results, highlighting the complexity of RG adherence and highlighting the need to implement interventions with different stakeholders throughout the research process [37]. Their review showed that most of the evaluated interventions to improve RG adherence have been conducted in journals. There have been mixed results but promising ones for more active implementation efforts (i.e., requiring a checklist with submission), including editorial assistants trained on reporting issues, and automatic peer review tools.

Widespread interventions are needed to improve RG adherence. Efforts to target research clusters, not just individuals, to foster broader support are needed. With increased uptake among coauthors completing reporting checklists, the time required may be reduced further, thus making using STROBE more appealing. When RGs become an expected part of the research process, self-regulation can occur and formal journal and institutional policies can be more fruitful as well [38]. Targeted and widespread promotion of RGs is needed to improve the completeness of reporting and reduce research waste [39].

CRediT authorship contribution statement

Melissa K. Sharp: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Visualization, Writing - original draft, Writing - review & editing. **Lorenzo Bertizzolo:** Data curation, Methodology, Writing - review & editing. **Roser Rius:** Formal analysis, Methodology, Writing - review & editing, Supervision. **Elizabeth Wager:** Writing - review & editing. **Guadalupe Gómez:** Formal analysis, Methodology, Writing - review & editing, Supervision, Resources. **Darko Hren:** Supervision, Methodology, Formal analysis, Visualization, Writing - review & editing, Resources.

Acknowledgments

The authors would like to acknowledge the Methods in Research on Research (MiRoR) network for their support and guidance. The authors would also like to thank Dr. Michail Kovanis for assistance gathering author emails and Ms. Alice Biggane for her assistance with the recruitment strategy.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jclinepi.2019.07.019.

References

[1] Gawande A. The checklist manifesto: How to get things right. New York, NY: Metropolitan Books; 2010.
[2] Costa BR da, Cevallos M, Altman DG, Rutjes AWS, Egger M. Uses and misuses of the STROBE statement: bibliographic study. BMJ Open 2011;1:e000048.
[3] Grindlay DJ, Dean RS, Christopher MM, Brennan ML. A survey of the awareness, knowledge, policies and views of veterinary journal Editors-in-Chief on reporting guidelines for publication of research. BMC Vet Res 2014;10:10.
[4] Cobo E, Cortés J, Ribera JM, Cardellach F, Selva-O'Callaghan A, Kostov B, et al. Effect of using reporting guidelines during peer review on quality of final manuscripts submitted to a biomedical journal: masked randomised trial. BMJ 2011;343:d6783.
[5] Larson EL, Cortazal M. Publication guidelines need widespread adoption. J Clin Epidemiol 2012;65:239−46.
[6] MacMahon B, Weiss NS. Is there a dark phase of this STROBE? Epidemiology 2007;18:791.
[7] Rothman KJ, Poole C. Some guidelines on guidelines: they should come with expiration dates. Epidemiology 2007;18:794.

[8] Hopewell S, Boutron I, Altman DG, Barbour G, Moher D, Montori V, et al. Impact of a web-based tool (WebCONSORT) to improve the reporting of randomised trials: results of a randomised controlled trial. BMC Med 2016;14:199.
[9] Sharp MK, Tokalić R, Gómez G, Wager E, Altman DG, Hren D. A cross-sectional bibliometric study showed suboptimal journal endorsement rates of STROBE and its extensions. J Clin Epidemiol 2019;107:42−50.
[10] Hua F, Walsh T, Glenny A-M, Worthington H. Surveys on reporting guideline usage in dental journals. J Dent Res 2016;95:1207−13.
[11] Kunath F, Grobe HR, Rücker G, Engehausen D, Antes G, Wullich B, et al. Do journals publishing in the field of urology endorse reporting guidelines? A survey of author instructions. Urol Int 2011;88:54−9.
[12] Meerpohl JJ, Wolff RF, Niemeyer CM, Antes G, von Elm E. Editorial policies of pediatric journals: survey of instructions for authors. Arch Pediatr Adolesc Med 2010;164:268−72.
[13] Toews I, Binder N, Wolff RF, Toprak G, von Elm E, Meerpohl JJ. Guidance in author instructions of hematology and oncology journals: a cross sectional and longitudinal study. PLoS One 2017;12:e0176489.
[14] Eysenbach G. Improving the quality of web surveys: the checklist for reporting results of internet E-surveys (CHERRIES). J Med Internet Res 2004;6:e34.
[15] Projet MiRoR | An innovative and ambitious joint doctoral training programme. Available at http://miror-ejd.eu/. Accessed July 16, 2018.
[16] Welcome to SurveyMonkey!. Available at https://www.surveymonkey.com/home/. Accessed March 17, 2017.
[17] Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27:425−78.
[18] Williams MD, Rana NP, Dwivedi YK. The unified theory of acceptance and use of technology (UTAUT): a literature review. J Enterp Inf Manag 2015;28:443−88.
[19] Technology (U.S.) NIC on HSR& HC. HTA 101: Introduction to Health Technology Assessment. Available at https://www.nlm.nih.gov/nichsr/hta101/ta10103.html. Accessed July 2, 2019.
[20] Sharp MK, Utrobičić A, Gómez G, Cobo E, Wager E, Hren D. The STROBE extensions: protocol for a qualitative assessment of content and a survey of endorsement. BMJ Open 2017;7:e019043.
[21] Sharp MK. Methods in research on research (MiRoR). STROBE guidelines (ESR9). 2017. Available at osf.io/u75gb. Accessed July 1, 2019.
[22] Broad subject terms for indexed journals. Available at https://wwwcf.nlm.nih.gov/serials/journals/index.cfm. Accessed July 16, 2018.
[23] STROBE statement: endorsement. Available at https://www.strobe-statement.org/index.php?id=strobe-endorsement. Accessed July 16, 2018.
[24] Journal rankings on agricultural and biological sciences. Available at https://www.scimagojr.com/journalrank.php?area=2700. Accessed July 16, 2018.
[25] Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951;16:297−334.
[26] Bland JM, Altman DG. Statistics notes: Cronbach's alpha. BMJ 1997;314:572.
[27] Sharp MK. Use of reporting guidelines as an educational intervention for teaching research methods and writing. *Open Sci Framew* https://doi.org/10.17605/OSF.IO/HP623. Accessed November 16, 2017.
[28] Rosseel Y. Lavaan: an R package for structural equation modeling. J Stat Softw 2012;1(2):2012.
[29] Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling Multidiscip J 1999;6:1−55.
[30] Brown TA. Confirmatory factor analysis for applied research. 2nd ed. New York, NY: The Guilford Press; 2015.
[31] EU GDPR Information Portal. EU GDPR portal. Available at http://eugdpr.org/eugdpr.org-1.html. Accessed August 10, 2018.
[32] Maul A. Method effects and the meaning of measurement. Front Psychol 2013;4:169.
[33] Fan W, Yan Z. Factors affecting response rates of the web survey: a systematic review. Comput Human Behav 2010;26:132−9.
[34] Vilaró M, Cortés J, Selva-O'Callaghan A, Urrutia A, Ribera J-M, Cardellach F, et al. Adherence to reporting guidelines increases the number of citations: the argument for including a methodologist in the editorial process and peer-review. BMC Med Res Methodol 2019;19:112.
[35] Pouwels KB, Widyakusuma NN, Groenwold RH, Hak E. Quality of reporting of confounding remained suboptimal after the STROBE guideline. J Clin Epidemiol 2016;69:217−24.
[36] Stevens A, Shamseer L, Weinstein E, Yazdi F, Turner L, Thielman J, et al. Relation of completeness of reporting of health research to journals' endorsement of reporting guidelines: systematic review. BMJ 2014;348:g3804.
[37] Blanco D, Altman D, Moher D, Boutron I, Kirkham JJ, Cobo E. Scoping review on interventions to improve adherence to reporting guidelines in health research. BMJ Open 2019;9:e026589.
[38] Fuller T, Pearson M, Peters J, Anderson R. What affects authors' and editors' use of reporting guidelines? Findings from an online survey and qualitative interviews. PLoS One 2015;10:e0121585.
[39] Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, et al. Reducing waste from incomplete or unusable reports of biomedical research. Lancet 2014;383:267−76.