BMJ Open Training patients to review scientific reports for the Patient-Centered Outcomes Research Institute: an observational study

Ilya Ivlev,^{1,2} Kelly J Vander Ley,¹ Jack Wiedrick,³ Kira Lesley,¹ Amy Forester,¹ Rebekah Webb,¹ Marina Broitman,⁴ Karen B Eden¹

To cite: Ivlev I, Vander Ley KJ, Wiedrick J, *et al.* Training patients to review scientific reports for the Patient-Centered Outcomes Research Institute: an observational study. *BMJ Open* 2019;**9**:e028732. doi:10.1136/bmjopen-2018-028732

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-028732>).

Received 20 December 2018
Revised 14 August 2019
Accepted 20 August 2019

© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Department of Medical Informatics & Clinical Epidemiology, Oregon Health & Science University, Portland, Oregon, USA

²Center for Health Research, Kaiser Permanente, Portland, Oregon, USA

³Biostatistics & Design Program, Oregon Health & Science University, Portland, Oregon, USA

⁴Patient-Centered Outcomes Research Institute, Washington, District of Columbia, USA

Correspondence to

Dr Ilya Ivlev;
ilya.ivlev@kpchr.org

ABSTRACT

Objective The peer review of completed Patient-Centered Outcomes Research Institute (PCORI) funded research includes reviews from patient reviewers (patients, caregivers, and patient advocates). Very little is known about how best to support these reviewers in writing helpful comments from a patient-centred perspective. This study aimed to evaluate the effect of a new training in peer review for patient reviewers.

Design Observational study.

Setting Online.

Participants Adults registered in the PCORI Reviewer Database as a patient stakeholder.

Intervention A new online training in peer review.

Main outcome measures Changes in reviewers' knowledge and skills; change in self-efficacy and attitudes, satisfaction with the training and perceived benefits and relevance of the training.

Results Before-after training survey data were analysed for 37 (29.4% of 126) patient reviewers invited to participate in an online training as part of a quality improvement effort or as part of a PCORI peer review. The reviewers improved their answers to the knowledge questions ($p<0.001$, median number of answers improved 4 (95% CI 3 to 5), large effect size (ES) Cohen's $w=0.94$) after the training, particularly in the questions targeting the specifics of PCORI peer review. Reviewers improved their skills in recognising helpful review comments, but those without peer-review background improved proportionally more ($p=0.008$, median number of answers improved 2 (95% CI 1 to 3), medium ES $w=0.60$). The use of training modestly increased reviewers' confidence in completing a high-quality peer review ($p=0.005$, mean increase in 5-point Likert rating 0.51 (95% CI 0.17 to 0.86), small-to-medium ES Cliff's $\delta=0.32$) and their excitement about providing a review slightly increased ($p=0.019$, mean increase in 5-point Likert rating 0.35 (95% CI 0.03 to 0.68), small ES $\delta=0.19$). All reviewers were satisfied with the training and would recommend it to other reviewers.

Conclusions Training improved knowledge, skills and self-efficacy and slightly increased enthusiasm for completing a PCORI peer review.

INTRODUCTION

Rigorous peer review is crucial to establish scientific validity and the high-quality of published research manuscripts.¹ In response

Strengths and limitations of this study

- This is the first known publication to evaluate peer review training for patient reviewers who will be asked to comment on the usefulness and relevance of healthcare-related scientific reports.
- We partnered with patients, caregivers and patient advocates to design this training.
- Training improved patient reviewers' knowledge, skills and self-efficacy, and slightly increased enthusiasm for completing a peer review for the Patient-Centered Outcomes Research Institute.
- Analyses were conducted on a small number of patient reviewers.

to its authorising law,² the Patient-Centered Outcomes Research Institute (PCORI) established an external peer review process for draft final research reports (DFRRs), which provide a complete and detailed description of each funded study. PCORI's peer review process^{3,4} was designed to assess these DFRRs for scientific integrity and adherence to PCORI's Methodology Standards, as well as the relevance and usefulness of the report content for patients, clinicians and other stakeholders.⁵ During peer review, each DFRR is assigned to an Associate Editor (scientist experienced in patient-centred outcomes research contracted to serve as editor for PCORI peer review) to identify needed reviewer expertise, ensure reviewer comments are adequately addressed in subsequent revisions and confirm the report meets PCORI's requirements.

In line with PCORI's overall goals of involving patients and stakeholders in all aspects of research,^{6,7} PCORI's peer review includes patients, caregivers and patient advocates, as well as stakeholders and scientists with subject matter and methodological expertise. However, there is no established

Figure 1 Content and structure of the training. DFRR, draft final research report; PCORI, Patient-Centered Outcomes Research Institute.

practice to ensure that published reports are useful to patients and reflect their priorities.⁸ By having patients serve as peer reviewers, PCORI aims to ensure the final research reports provide information helpful to clinicians and other decision-makers, as well as patients making healthcare decisions.³ PCORI's patient reviewers are invited to comment on the patient-centred aspects of the conducted study and the meaningfulness of the interventions for patients or caregivers.

Incorporating patient reviewers in peer review is challenging because patient reviewers have no roadmap for critiquing a manuscript and no training to prepare

them, which may limit their useful and actionable feedback. Consequently, both scientific and patient reviewers expressed concern about patient reviewers being able to understand and comment on scientific and technical aspects of manuscripts or research proposals.^{6,9}

To prepare patient reviewers for peer review, PCORI considered past attempts of training aimed to improve the quality of scientific peer review. These attempts include: voluntary training workshops,^{10,11} the implementation of structured training interventions¹² and self-teach training packages.¹³ For example, the *British Medical Journal* (BMJ) offers reviewer training materials in the form of readily available online presentations, papers about peer review and an exercise where reviewers can practice their skills with a research article.¹⁴ Despite these efforts, there is no evidence that specialised training can improve peer review quality.^{13,15–18} However, training peer reviewers remains a promising intervention to enhance the quality of reviews^{17,19,20} and may provide additional support to patient peer reviewers. For example, patient and public reviewers for *The BMJ* and *Research Involvement and Engagement* would like access to additional online resources and training.⁹

We developed online Patient Reviewer Training to help PCORI patient reviewers write reviews that express their unique perspectives. The purpose of this study was to evaluate our training to determine if it would help future reviewers to gain the appropriate knowledge and skills to conduct a useful peer review, and to what extent the training affected reviewers' self-efficacy and attitudes towards reviewing a DFRR. We also studied reviewers' satisfaction, perceived benefits and relevance of the training. We used the results of this study to refine the training for future users.

METHODS

Data were collected and analysed at the Oregon Health & Science University (OHSU), which provides peer review services to PCORI under contract. The OHSU Institutional Review Board (IRB) determined that this activity is exempt from IRB review and approval (IRB ID: STUDY00016016).

Study design

This was an observational study designed to evaluate the effect of the training on prospective PCORI patient reviewers. Surveys were completed by patient reviewers who used our web-based peer review training between January and July 2017 prior to their first peer review. We employed SurveyMonkey to collect survey responses before and after the training.

Participants and settings

PCORI defines patients as 'Persons with current or past experience of illness or injury, family members or other unpaid caregivers of patients or members of advocacy organizations that represent patients or caregivers'.²¹ The

Figure 2 Flow diagram. *Three reviewers answered post-training survey twice. First answers from these reviewers were used for analysis. PCORI, Patient-Centered Outcomes Research Institute.

Table 1 Participants' characteristics

Characteristic	Overall, n=37	Reviewers without peer review background, n=22	Reviewers with peer review background, n=15
Age in years, mean (SD)	53.8 (12.7)	54.3 (12.6)	54.8 (11.5)
Gender, n (%)			
Female	32 (86.5)	18 (81.8)	14 (93.3)
Male	5 (13.5)	4 (18.2)	1 (6.7)
Race/ethnicity, n (%)			
White	29 (78.4)	14 (63.6)	15 (100)
Black or African American	4 (10.8)	4 (18.2)	0
Asian	2 (5.4)	2 (9.1)	0
Reported more than one	3 (8.1)	2 (9.1)	1 (6.7)
Other	1 (2.7)	1 (4.5)	0
Declined to answer	1 (2.7)	1 (4.5)	0
Hispanic, Latino or Spanish origin	3 (8.1)	2 (9.1)	1 (6.7)
Role in PCORI's peer review, n (%)			
Patient advocate	18 (48.6)	11 (50.0)	7 (46.7)
Patient	15 (40.5)	8 (36.4)	7 (46.7)
Caregiver	4 (10.8)	3 (13.3)	1 (6.7)
Degree, n (%)			
Some college but no degree	1 (2.7)	1 (4.5)	0
Associate degree	4 (10.8)	4 (18.2)	0
Bachelor degree	9 (24.3)	6 (27.3)	3 (20.0)
Master's degree	20 (54.1)	9 (40.9)	11 (73.3)
Professional degree (MD, JD and so on) or Doctoral (PhD)	3 (8.1)	2 (9.1)	1 (6.7)
Previous training in peer review, n (%)	12 (32.4)	NA	12 (80.0)
Performed a peer review of any medical literature, n (%)	13 (35.1)	NA	13 (86.7)
Grant, policy or guideline	11 (29.7)		11 (73.3)
Manuscript or research report	11 (29.7)		11 (73.3)
Other	1 (2.7)		1 (6.7)
Involved in biomedical research, n (%)	8 (21.6)	5 (22.7)	3 (20.0)
Coauthor of a paper in a peer-reviewed journal, n (%)	11 (29.7)	4 (18.2)	7 (46.7)
Received any training in epidemiology, n (%)	6 (16.2)	3 (13.6)	3 (20.0)

NA, not applicable; PCORI, Patient-Centered Outcomes Research Institute.

eligibility criteria for our study population were adults who voluntarily registered in the PCORI peer review online database (opened to the public in September 2016) as patient reviewers and had not been previously invited to review a DFRR or accessed the training prior to this study. We invited all 126 eligible people to take the online training and complete a pre/post-training survey. Seventy (55.6%) of those eligible were invited as a part of the PCORI patient peer review workflow. They had been newly assigned a

review but had not yet initiated it; the training was presented as a precursor to the review. The remaining 56 (44.4%) were invited as a part of this quality improvement effort and incentivised to take the training prior to any review assignment. We did this to increase response numbers within the study's limited timeframe. Those who completed only the pre/post-training survey as part of the quality improvement effort were offered US\$10. Those participants who completed the pre/post-training survey and submitted their

Table 2 Change in knowledge and skills

Reviewers' background	Change in answers after training	Outcomes							
		Knowledge				Skills			
		n (%)	Median number of answers improved (approximate 95% CI)	P value	*Cohen's w	n (%)	Median number of answers improved (approximate 95% CI)	P value	*Cohen's w
All reviewers, n=37	Improved	33 (89)	4 (95% CI 3 to 5)	<0.001	Large, 0.94	22 (59)	1 (95% CI 0 to 2)	0.002	Medium, 0.54
	Worsened	0				5 (14)			
	Remained the same	4 (11)				10 (27)			
Reviewers without peer review background, n=22	Improved	19 (86)	5 (95% CI 4 to 6)	<0.001	Large, 0.93	15 (68)	2 (95% CI 1 to 3)	0.008	Medium, 0.60
	Worsened	0				3 (14)			
	Remained the same	3 (14)				4 (18)			
Reviewers with peer review background, n=15	Improved	14 (93)	4 (95% CI 3 to 5)	<0.001	Large, 0.97	7 (47)	0 (95% CI -1 to 1)	0.180	Medium, 0.43
	Worsened	0				2 (13)			
	Remained the same	1 (7)				6 (40)			

*Effect size

review of a DFRR were offered US\$50. Money was administered through a debit card.

Training development

OHSU and PCORI cohosted a workshop with 68 patient attendees at Stanford University's 2016 Medicine X Conference. We asked attendees to participate in four discussion groups that each peer-reviewed a sample of DFRRs. While participants were pleased to be part of the peer-review process, they asked if peer-review training, resources on research design and a glossary of research terms were available. In addition to the feedback received at the Medicine X Conference, we used the Ethical Guidelines for Peer Reviewers by the Committee on Publication Ethics (COPE)²² to identify underlying principles and general rules that patient reviewers could adopt during peer review (eg, provide constructive and avoid hostile comments, provide recommendations that will help make the DFRR more useful and be specific in criticisms and recommendations). After we created the first version of the training, we performed three testing rounds with 16 patient reviewers, two rounds of testing with five PCORI ambassadors (patient volunteers already trained as advocates by PCORI and comfortable with scientific writing), and one testing round with the staff from the OHSU Editorial Office (Portland, OR, USA) which is under contract to manage peer review services for PCORI. We demonstrated the nearly-final version of the training on a conference call with 22 PCORI associate editors and collected their feedback.

Training evaluation framework

Our evaluation approach was guided by Kirkpatrick's hierarchical framework for assessing the success of complex training interventions.^{23 24} We used this approach to evaluate our new online training at the first

two levels—'Reaction' (Level 1) and 'Learning' (Level 2). The reaction level evaluates reviewers' attitudes towards the training, satisfaction with the training and perceived benefits and relevance. The learning level evaluates the gain of intended knowledge/skills and changes in enthusiasm and confidence in conducting peer reviews. In our study, the 'Reaction' training evaluation examined: (i) the patient reviewers' satisfaction with the training, including qualitative data (obtained from open-ended questions); (ii) perceived benefit of the training (change in self-reported improvement in knowledge of peer review) and (iii) perceived relevance and usefulness of the training (willingness to recommend the training to other reviewers). 'Learning'—changes in knowledge, skills, self-efficacy and attitudes—was measured by evaluating: (i) the pre-post-training change in answers to knowledge questions; (ii) the change in the reviewers' ability to identify examples of more helpful peer-review comments; (iii) change in self-efficacy (confidence in ability to complete a high-quality peer review) and (iv) changes in reviewers' self-reported excitement about performing a peer review. Additionally, we assessed the final quality of the peer-review reports and the time that the reviewers spent completing their peer reviews.

Using only the first two levels of the Kirkpatrick model ensured a rapid assessment of the training and supported future quality improvement. Levels 3 and 4 of the model are focused on 'Behaviour' and 'Results'. The 'Behaviour' level would look at how patients use new knowledge from the training in drafting their reviews. The 'Results' level would study the extent to what our training improved usefulness and relevance of patient peer review for the authors of DFRRs and PCORI's associate editors. A larger study involving additional levels of the Kirkpatrick model would require a longer timeframe to collect necessary peer

Table 3 Change in self-efficacy and level of excitement

All reviewers, n=37					Reviewers without peer review background, n=22					Reviewers with peer review background, n=15					
Outcome	Before training, n (%)	After training, n (%)	Mean increase in Likert rating (95%CI)	P value	*Cliff's delta	Before training, n (%)	After training, n (%)	Mean increase in Likert rating (95%CI)	P value	*Cliff's delta	Before training, n (%)	After training, n (%)	Mean increase in Likert rating (95% CI)	P value	*Cliff's delta
Response to the statement 'I am confident I can complete a high-quality peer review of a draft final research report for PCORI.'															
Strongly agree	16 (43)	26 (70)	0.51 (0.17 to 0.86)	0.005	Small-to-medium, 0.32	9 (41)	18 (82)	0.59 (0.13 to 1.05)	0.011	Medium, 0.41	7 (47)	8 (53)	0.40 (−0.16 to 0.96)	0.212	Small, 0.20
Agree	15 (41)	10 (27)				10 (45)	3 (14)				5 (33)	7 (47)			
Undecided	4 (11)	1 (3)				2 (9)	1 (5)				2 (13)	0			
Disagree	0	0				0	0				0	0			
Strongly disagree	2 (5)	0				1 (5)	0				1 (7)	0			
Response to the statement 'I am excited to provide a peer review of a draft final research report for PCORI.'															
Strongly agree	23 (62)	29 (78)	0.35 (0.03 to 0.68)	0.019	Small-to-medium, 0.32	14 (64)	18 (82)	0.36 (−0.05 to 0.78)	0.057	Medium, 0.41	9 (60)	11 (73)	0.33 (−0.23 to 0.89)	0.158	Small, 0.13
Agree	11 (30)	8 (22)				6 (27)	4 (19)				5 (33)	4 (27)			
Undecided	1 (3)	0				1 (5)	0				0	0			
Disagree	0	0				0	0				0	0			
Strongly disagree	2 (5)	0				1 (5)	0				1 (7)	0			
*Effect size PCORI, Patient-Centered Outcomes Research Institute.															

*Effect size
PCORI, Patient-Centered Outcomes Research Institute.

Table 4 Reaction to the training

Outcome	All reviewers, n=37	Reviewers without peer review background, n=22	Reviewers with peer review background, n=15
Response to the statement 'This training enhanced my knowledge of the peer-review process.'			
Strongly agree or Agree, n (%)	36 (97)	21 (95)	15 (100)
Undecided, n (%)	1 (3)	1 (5)	0
Disagree or Strongly disagree, n (%)	0	0	0
Response to the statement 'I was satisfied with the training overall.'			
Strongly agree or Agree, n (%)	37 (100)	22 (100)	15 (100)
Undecided, Disagree or Strongly disagree, n (%)	0	0	0
Response to the statement 'I would recommend this training to others patient or caregiver reviewers.'			
Strongly agree or Agree, n (%)	37 (100)	22 (100)	15 (100)
Undecided, Disagree or Strongly disagree, n (%)	0	0	0

reviews and to evaluate whether our training improved quality of peer reviews.

Intervention

We designed our web-based training to introduce patient reviewers to the PCORI peer-review process, help them to understand the value of their feedback and enhance their skills and abilities to complete a review. We intended the training to require no more than 90 min. The asynchronous training site was hosted by the Sakai Learning Management System (V.11)²⁵ and included videos, learning activities and review resources (figure 1). The videos provided the reasoning behind patient reviews as well as directions and guidance for completing a peer review. The training also included five learning activities: viewing a tailored review form for patient reviewers, reading a deidentified DFRR annotated by an associate editor, writing a practice review and reading more and less helpful review comments. Reading materials were offered in both Microsoft Word and PDF formats. Reviewers were also provided with screenshots of the online Editorial Manager system that manages peer review. The training could be accessed from any operating system on a personal computer, tablet or smartphone. Just-in-time support was provided via phone or email by the OHSU Teaching and Learning Center Help Desk.

Variables and data sources/measurement

The reviewers' use of the training materials (time between presurveys and postsurveys) and dose of exposure to the training (number of training elements accessed and number of times each training element accessed) were verified in Sakai Analytics (OHSU's learning management system that allows tracking learners' activities on the training website).²⁵

Outcomes: The reviewers' knowledge was measured using 14 statements (online supplementary files A and B) about what they should or should not do as PCORI patient reviewers when reviewing a DFRR. The reviewers were offered three response options: yes, no and unsure.

These knowledge questions were developed using the COPE recommendations and specific PCORI expectations for patient reviewers to cover the main knowledge elements relevant to PCORI's peer review. The reviewers' skills were measured as the ability to identify more and less helpful peer review comments. Each reviewer was asked to rate eight peer review comments as more or less helpful; the reviewers could also answer 'unsure'. These knowledge questions and peer-review statements were piloted among three associate editors and revised based on their feedback. The question and answer order in SurveyMonkey was randomly presented to each reviewer to alleviate order bias.

Reviewers rated their (i) self-efficacy (confidence) in completing a high-quality peer review (before and after the training) and (ii) level of excitement in providing a peer review (before and after the training) using a 5-point Likert scale (strongly agree, agree, undecided, disagree, strongly disagree; coded 5 through 1 sequentially). Using the same scale, we also asked reviewers to rate their overall satisfaction with the training, perception of whether the training enhanced their knowledge of peer-review and willingness to recommend the training to other reviewers.

Additional variables and data sources/measurement: At the end of the survey, we asked participants to (i) provide their demographic data, (ii) report any prior training in peer-review or epidemiology, (iii) report experience peer-reviewing any biomedical literature, (iv) report whether they coauthored a paper published in a peer-reviewed journal and/or (v) report whether they had ever been involved in biomedical research. As part of the peer review workflow, associate editors provided a review quality score of every review (all PCORI reviewer categories: patients, stakeholders and scientists) on a single 5-point global scale in the Editorial Manager: 1—poor (unacceptable effort and content), 2—fair (unacceptable effort or content), 3—average (acceptable quality of the review), 4—good (commendable; of use to both editor and author), 5—excellent (exceptional; hard to

improve (expected to describe no more than 10%–15% of reviews)). The scale has been used by peer-reviewed journals^{16 18 26 27} and the National Institutes of Health, but to our knowledge, has not been used to rate the quality of patient peer reviews. The reliability and validity of this scale were previously assessed.^{10 26} Reviewers also reported how many hours they spent to complete their review.

Statistical methods

This was an observational study. Data included before and after training surveys completed by PCORI patient peer reviewers between January and July 2017.

Outcome analyses: We used a sign test to compare differences in responses to knowledge and skills questions by counting the numbers of answers that improved (coded +1), worsened (coded -1) or remained the same (coded zero) before and after the training. The goodness of fit of the response categorisation to a null distribution with equal probability on the numbers of improved and worsened answers (discounting answers that remained the same) was assessed and expressed as Cohen's *w*,²⁸ an effect size metric that measures arbitrary associations among categorical responses. Mean absolute deviation quantile regression was used to calculate CIs for the median change in all responses. Changes in the number of correct answers for each knowledge or skills question before and after the training were evaluated using McNemar's test for paired dichotomous data. Detailed analysis of each of the 14 knowledge questions is presented in online supplementary files A and B.

The Wilcoxon signed-rank test was used to examine the before-and-after differences in the reviewers' self-efficacy and level of excitement. Effect sizes for the differences were expressed as Cliff's *delta*,²⁹ a metric appropriate for ordinal data that characterises the extent to which values from one distribution (ratings after the training) are larger than values from another distribution (ratings before the training).

Subanalyses/Exploratory analyses: Additionally, exploratory subgroup analyses were performed to study to what extent the above outcomes were influenced by the reviewers' self-reported peer review background (previous training or experience in peer review not including PCORI). The purpose of the follow-up analysis was to see whether those without peer review background took more time to perform their first review, if first time reviews received a lower review quality score compared with those with peer review background, and whether the two groups devoted differing amounts of effort to the training in preparation for their first assigned review. Differences between outcomes (perceived improvement in knowledge, satisfaction with the training, willingness to recommend the training to other reviewers, review quality global score) among reviewers with and without a peer-review background were examined using the Mann-Whitney U test. Assessment of the quality of the reviewers' first assigned peer review was considered a follow-up analysis.

All participants who accessed training and completed before and after training surveys were included in the analysis. All data elements for these reviewers were complete (ie, there was no missing data for main outcomes).

All calculations were carried out using IBM SPSS Statistics (V.25) and Stata (V.15.1). Multiple-testing correction of the *p* values from the detailed analysis of the knowledge questions (online supplementary file A) employed a Holm-Bonferroni sequential procedure.³⁰

Patient and public involvement

Sixty-eight patients attended our workshop at Stanford University's 2016 Medicine X Conference. They helped us to identify a need for a training in peer-review. Our study was designed without patient involvement. We did not invite patients to contribute to the writing or editing of this manuscript; however, Amy Price, Ph.D. (Patient Editor, *The BMJ*) has reviewed the draft manuscript.

RESULTS

Participants

All 126 eligible reviewers from the PCORI Patient Reviewer Database were invited to participate in the online training as part of a quality improvement effort or as part of the PCORI patient peer review between January and July 2017. Of these 126, 65 (51.6%) entered the study and 61 (48.4%) did not respond to the invitation (*figure 2*). Sakai Analytics and Survey Monkey identified 37 patient reviewers who accessed the training and completed pretraining and post-training surveys (29.4% response rate). Of the 28 reviewers who were not included in the analysis, 19 dropped out (14 completed the pretraining survey only and five completed the pretraining survey and accessed the training), seven accessed the training and completed the post-training survey only, one started the pretraining survey after completing the post-training survey and one never accessed any of the training elements.

Of the 37 reviewers (*table 1*) included in the analysis, 15 (40.5%) had performed a peer review (but not a DFRR review) prior to this study and/or had previous training in peer review (but not PCORI training), 11 (29.7%) had coauthored a paper published in a peer-reviewed journal, 8 (21.6%) were involved in biomedical research and 6 (16.2%) reported training in epidemiology. The majority of the participants self-identified as female, 32 (86.5%), and white, 29 (78.4%), and a range of educational levels were reported. Reviewers had unlimited access to the training and spent a median of 14.9 (IQR 109.8) hours between pretraining and post-training surveys. The 35 reviewers for whom Sakai data were available accessed a median of six (of seven, IQR 2) training elements three (IQR 2) times each (online supplementary file C).

Outcomes: knowledge and skills

Across the responses from all the reviewers, 33 (of 37, 89%) improved their answers to the knowledge questions

($p < 0.001$, large effect size Cohen's $w = 0.94$) after the training (table 2, online supplementary file A). The reviewers demonstrated a high level of knowledge before the training in five questions (online supplementary files A and B)—before undergoing the training many of them already knew that they should not contact authors directly (question 1; 35 (95%) vs 37 (100%)), that they needed to provide constructive comments and avoid hostile comments (question 3; 37 (100%) vs 36 (97%)), that they needed to report possible conflicts of interest (question 8; 35 (95%) vs 37 (100%)) and that they should not share peer-review materials (question 9; 37 (100%) before and after). The biggest improvement in knowledge was observed when reviewers answered correctly that they are expected to: (i) evaluate the potential influence of the research results on them as patients (question 6; 26 (70%) vs 36 (97%)); (ii) bring a sceptical mind to the task (question 2; 13 (35%) vs 32 (86%)); (iii) not comment on accepting or rejecting the DFRR (question 12; 8 (22%) vs 25 (68%)); (iv) not comment on appropriateness of statistical methods (question 13; 18 (49%) vs 35 (95%)) and (v) apply personal experience to provide recommendations (question 14; 23 (62%) vs 36 (97%)). Within the group of peer reviewers without any peer-review background, a higher number of reviewers correctly answered three questions after the training (questions 2, 13, 14) about the need for using personal expertise, bringing a sceptical mind to the review, and the lack of a need for commenting on statistical methods. The most prominent improvements in answers by reviewers with peer review background were exhibited in three questions that targeted the aspects of PCORI peer review that differ from journal peer review (questions 6, 12 and 13)—the lack of a need for commenting on the rejection/acceptance of DFRRs (2 (13%) vs 11 (73%) of 15 reviewers); lack of the need for evaluating the correctness of statistical methods (8 (53%) vs 15 (100%)) and need to evaluate the potential impact of the research results on themselves as patients, caregivers or patient advocates (9 (60%) vs 15 (100%)). Reviewers both with and without peer-review background improved their skills in recognising less helpful review comments, but those without peer-review background improved proportionally more (table 2; $p = 0.008$ for without, medium effect size $w = 0.60$; $p = 0.180$ for with, medium effect size $w = 0.43$).

Outcomes: self-efficacy, excitement, perceived benefits and relevance of the training

After taking the training, reviewers were more confident that they could complete a high-quality peer review (table 3; $n = 37$, $p = 0.005$, small-to-medium effect size Cliff's $\delta = 0.32$). The reviewers' level of excitement in providing a review after using the training also increased somewhat ($n = 37$, $p = 0.019$, small effect size $\delta = 0.19$). Before the training, 34 (92%) of the reviewers strongly agreed or agreed, 1 (3%) was undecided and 2 (5%) strongly disagreed with the statement that they were excited to provide a peer review of a DFRR (table 3).

However, after using the training, all 37 (100%) reviewers agreed or strongly agreed that they were excited about providing a peer review.

All 37 reviewers, regardless of whether they had prior background in peer review, were satisfied with the training and would recommend it to other reviewers (table 4). Although one reviewer without experience or prior training was unsure about whether the training enhanced his or her knowledge in peer review, the other 36 (97%) reviewers strongly agreed or agreed that the training improved their knowledge (table 4).

Reviewers also responded to open-ended questions on what they liked about the training and what they would improve. The reviewers especially appreciated the clarity of the training and commented that it was extremely informative. Particularly, they liked the examples of real annotated DFRRs with the comments from associate editors (12 reviewers commented), examples of more and less helpful review comments (eight reviewers commented), the 'write a sample review' learning activity and the variety of training materials and module formats (11 reviewers commented).

Additional analyses: review performance

All reviewers were able to access the training again when reviewing an assigned DFRR. Twenty-nine reviewers (of 37; 78%) completed their first peer review between January 2017 and February 2018. Twenty-four reviewers (9 with peer review background; 15 without peer review background) received one global review quality score from the associate editor assigned to the DFRR in peer review (online supplementary file C). The reviewers' median global review quality score was equal to 4 (of 5; good (commendable; of use to both editor and author)) (IQR 1) points; this result did not differ among the reviewers with different peer-review backgrounds. The reviewers reported spending a median of five (IQR 4) hours to complete a peer review. Those who were more experienced in peer reviewing reported spending a median of an hour less compared with the reviewers without a peer-review background (4 hours (IQR 4) $n = 11$ vs 5 hours (IQR 7) $n = 18$, respectively); however, this difference was not large relative to the overall variance ($p = 0.415$).

DISCUSSION

To the best of our knowledge, this is the first study evaluating novel peer-review training for patient reviewers of healthcare-related scientific reports. This observational study showed that the online training improved reviewers' knowledge and skills in PCORI peer review. The reviewers with prior peer review or training experience enhanced their understanding of the specifics of PCORI peer-review. All reviewers were satisfied with the training and perceived the training as relevant. The training enhanced reviewers' confidence in performing a peer review.

This study has a few limitations. First, the participation level was relatively low—37 (29.4%) of 126 met the

inclusion criteria. Peer review background is a characteristic that could have an important effect on the reviewers' decision to participate in the study. We were not able to estimate the influence of the reviewers' peer review background on participation because this characteristic was not collected among people who did not answer a pretraining survey. Second, although the patient peer reviews completed during the study window were rated as 'good' (median score) by associate editors, the study design does not allow us to attribute the perceived usefulness of the reviews to the training. Third, the associate editors were responsible for peer reviewer selection and were not blinded to the number of times a patient peer-reviewed for PCORI, and other demographic information. Fourth, this peer-review training was developed for PCORI and may not have a similar effect on patient reviewers in scientific journals where manuscripts tend to be shorter and more field-specific. Last, PCORI's peer review programme is relatively new (begun in late 2016), so we are unable to say whether the current population of peer reviewers surveyed in this study are representative of the population as it may evolve.

In the future, the training developers may consider diversifying the training to account for patient reviewers' prior peer-review experience. Patient reviewers with prior training or peer-review experience might benefit from a shortened version of the training that could emphasise the differences between PCORI's peer review and expectations from those of scientific journal peer review.

Twitter Ilya Ivlev @DRivlev

Acknowledgements The authors would like to thank Ed Reid, MS, MAT, MFA for his help with retrieving reviewers' records from the PCORI Patient Reviewer Database and reviewing the draft manuscript; Amy Price, Ph.D. (Patient Editor, the BMJ) for reviewing the draft manuscript. The authors also would like to thank Camber Hansen-Karr, BA and Lauren Saxton, MS for their help with developing training. We would like to thank more than 100 patients, caregivers, patient advocates, associate editors and other contributors who partnered with us to develop this training.

Contributors All four International Committee of Medical Journal Editors (ICMJE) criteria for authorship met by AF, II, JW, KBE, KJVL, KL, MB, RW. Wrote the first draft of the paper: II. Agree with the manuscript's results and conclusions: AF, II, JW, KBE, KJVL, KL, MB, RW. Contributed to the writing of the paper: AF, II, JW, KBE, KJVL, KL, MB, RW.

Funding This study was conducted under a contract to provide peer review services for the PCORI DFRRs (OGA Project number: GSMMI0235B). It was additionally supported by the National Library of Medicine of the National Institutes of Health under Award #T15LM007088.

Disclaimer The views, statements, opinions presented in this article are solely the responsibility of the authors and do not necessarily represent the views of the Patient-Centered Outcomes Research Institute (PCORI), its Board of Governors or Methodology Committee. The National Library of Medicine had no role in design or conduct of this study.

Competing interests Authors have received support from PCORI for developing the training (AF, II, KE, KJVL, KL, RW), serving as an Associate Editor (KE, KJVL) or supporting PCORI's peer review process (KE, KJVL, MB, RW). MB is an Associate Director for Peer Review at the PCORI.

Patient consent for publication Not required.

Ethics approval The OHSU Institutional Review Board determined that this activity is exempt from IRB review and approval (ID: STUDY00016016).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement The datasets analyzed during the current study will be made available by the corresponding author on reasonable request.

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/>.

REFERENCES

1. Benda WGG, Engels TCE. The predictive validity of peer review: a selective review of the judgmental forecasting qualities of Peers, and implications for innovation in science. *Int J Forecast* 2011;27:166–82.
2. Compilation of patient protection and Affordable care act: extracted sections concerning patient-centered outcomes research and the authorization of the patient-centered outcomes research Institute (PCORI), 2010Subtitle D of Title VI-Sec. 6301. Enacted March 23, 2010. Amended May 1. Available: https://www.pcori.org/sites/default/files/PCORI_Authorizing_Legislation.pdf [Accessed 13 Aug 2019].
3. Patient-Centered Outcomes Research Institute (PCORI). PCORI's Process for Peer Review of Primary Research and Public Release of Research Findings, 2015. Available: <https://www.pcori.org/sites/default/files/PCORI-Peer-Review-and-Release-of-Findings-Process.pdf> [Accessed 13 Aug 2019].
4. Broitman M, Sox HC, Slutsky J. A model for public access to trustworthy and comprehensive reporting of research. *JAMA* 2019;321:1453–4.
5. Patient-Centered Outcomes Research Institute. PCORI methodological standards, 2018. Available: <https://www.pcori.org/sites/default/files/PCORI-Methodology-Standards.pdf> [Accessed 13 Aug 2019].
6. Fleurence RL, Forsythe LP, Lauer M, et al. Engaging patients and stakeholders in research proposal review: the patient-centered outcomes research Institute. *Ann Intern Med* 2014;161:122–30.
7. Patient-Centered Outcomes Research Institute (PCORI). Peer review of our studies, 2016. Available: <https://www.pcori.org/research-results/peer-review-our-studies> [Accessed 14 Jun 2019].
8. Ioannidis JPA. Why most clinical research is not useful. *PLoS Med* 2016;13:e1002049.
9. Schroter S, Price A, Flemmyng E, et al. Perspectives on involvement in the peer-review process: surveys of patient and public reviewers at two journals. *BMJ Open* 2018;8:e023357.
10. Callaham ML. Reliability of editors' subjective quality ratings of peer reviews of manuscripts. *JAMA* 1998;280.
11. Callaham ML, Schriger DL. Effect of structured workshop training on subsequent performance of Journal peer reviewers. *Ann Emerg Med* 2002;40:323–8.
12. Houry D, Green S, Callaham M. Does mentoring new peer reviewers improve review quality? A randomized trial. *BMC Med Educ* 2012;12:83.
13. Schroter S, Black N, Evans S, et al. Effects of training on quality of peer review: randomised controlled trial. *BMJ* 2004;328.
14. The BMJ. Guidance for BMJ patient and public reviewers, 2019. Available: <https://www.bmj.com/about-bmj/resources-reviewers/guidance-patient-reviewers> [Accessed 14 Jun 2019].
15. Bruce R, Chauvin A, Trinquart L, et al. Impact of interventions to improve the quality of peer review of biomedical journals: a systematic review and meta-analysis. *BMC Med* 2016;14:85.
16. Callaham ML, Tercier J. The relationship of previous training and experience of Journal peer reviewers to subsequent review quality. *PLoS Med* 2007;4:e40–40.
17. Mulligan A, Hall L, Raphael E. Peer review in a changing world: an international study measuring the attitudes of researchers. *J Am Soc Inf Sci Tec* 2013;64:132–61.
18. Callaham M, McCulloch C. Longitudinal trends in the performance of scientific peer reviewers. *Ann Emerg Med* 2011;57:141–8.
19. Patel J. Why training and specialization is needed for peer review: a case study of peer review for randomized controlled trials. *BMC Med* 2014;12:128.
20. Gough NR. Training for peer review. *Sci Signal* 2009;2:tr2.
21. Patient-Centered outcomes research Institute (PCORI). Engagement. PCORI's Stakeholders, 2014. Available: <https://www.pcori.org/engagement/what-we-mean-engagement/pcoris-stakeholders> [Accessed 13 Aug 2019].
22. Committee on publication ethics. Guidelines. principles of transparency and best practice in scholarly publishing, 2015.

- Available: <https://publicationethics.org/resources/guidelines-new/principles-transparency-and-best-practice-scholarly-publishing> [Accessed 14 Jun 2019].
23. Lovato C, Wall D. Programme evaluation: Improving practice, influencing policy and decision-making. In: Swanwick T, ed. *Understanding medical education: evidence, theory and practice*. John Wiley & Sons, Incorporated, 2013: 385–99.
 24. Kirkpatrick D, Kirkpatrick J. *Evaluating training programs: the four levels*. 3rd ed. San Francisco, CA: Berrett-Koehler Publishers, 2006.
 25. Caidin N. Sakai 11 release notes, 2017. Available: <https://confluence.sakaiproject.org/display/DOC/Sakai+11+Release+Notes> [Accessed 5 March 2018].
 26. van Rooyen S, Black N, Godlee F. Development of the review quality instrument (RQI) for assessing peer reviews of manuscripts. *J Clin Epidemiol* 1999;52:625–9.
 27. Callaham ML, Wears RL, Waeckerle JF. Effect of attendance at a training session on peer reviewer quality and performance. *Ann Emerg Med* 1998;32:318–22.
 28. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, NJ: Routledge, 1988.
 29. Cliff N. Dominance statistics: ordinal analyses to answer ordinal questions. *Psychol Bull* 1993;114:494–509.
 30. Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat* 1979;6:65–70.