BMJ Open

Threats to grant peer review: A Qualitative Study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2024-091666
Article Type:	Original research
Date Submitted by the Author:	25-Jul-2024
Complete List of Authors:	Sims Gould, Joanie; University of British Columbia, Department of Family Practice Lasinsky, Anne; University of British Columbia Mota, Adrian; Canadian Institutes of Health Research Khan, Karim; University of British Columbia, Department of Family Practice; Institute of Musculoskeletal Health and Arthritis Ardern, Clare L.; University of British Columbia, Department of Family Practice; University of British Columbia, Department of Physical Therapy
Keywords:	Review, Decision Making, Knowledge

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Threats to grant peer review: A Qualitative Study

Authors

Joanie Sims Gould[1] joanie@joaniesimsgould.com

Anne Lasinsky[2] anne.lasinsky@ubc.ca

Adrian Mota[3] adrian.mota@cihr-irsc.gc.ca

Karim M. Khan[1,5] karim.khan@ubc.ca

Clare L. Ardern[1,6] clare.ardern@ubc.ca

Affiliations

[1] Department of Family Practice, University of British Columbia, Vancouver, British Columbia, Canada

[2] School of Kinesiology, University of British Columbia, Vancouver, British Columbia, Canada

[3] Canadian Institutes of Health Research, Ottawa, Ontario, Canada

[5] Canadian Institutes of Health Research-Institute of Musculoskeletal Health and Arthritis, Vancouver, British Columbia, Canada

[6] Department of Physical Therapy, University of British Columbia, Vancouver, British Columbia, Canada

Abstract

Background & Objectives: Peer review is ubiquitous in evaluating scientific research. While peer review of manuscripts submitted to journals has been widely studied, there has been relatively less attention paid to peer review of grant applications (despite how crucial peer review is to researchers having the means and capacity to conduct research). There is spirited debate in academic community forums (including on social media) about the perceived benefits and limitations of grant peer review. Yet we suggest that at least some aspects of the discussion are not adequately captured in the format of published academic work.

Methods: Therefore, we conducted qualitative interviews with 18 members of grant-review panels—the Chairs, peer reviewers and Scientific Officers of a national funding agency—that highlight threats to the integrity of grant peer review.

Results: We identified three threats: (1) lack of training and limited opportunities to learn, (2) challenges in differentiating and rating applications of similar strength, and (3) reviewers weighting reputations and relationships in the review process to differentiate grant applications of a similar strength. These threats were compounded by reviewers' stretched resources or lack of time. Our data pointed to two areas—virtual peer review and the role of the Chair—that have major influence on transparency and rigorous grant peer review.

Conclusions: As researchers continue to evaluate the threats to grant peer review, the reality of stretched resources and time must be considered. We call on funders to implement practices that reduce reviewer burden.

Keywords: peer review, research grants, training, time

Strengths and Limitations of this Study

- Qualitative interviews with leaders of grant-review panels—the chairs and scientific officers of a national funding agency—highlight threats to the integrity of grant peer review.
- We identified three threats: lack of training, challenges in differentiating and rating applications of similar strength and reputation and relationship-driven decisions.
- We identified two areas, virtual peer review and the role of the Chair that can facilitate equitable and inclusive grant peer review.
- Data was collected in the Canadian context so findings may not reflect the experience of grant peer reviewers in countries with different review systems and practices.

Background

There are threats to the integrity of the grant peer review process. The merit of grant peer-review—a fundamental element of science—has been questioned in many quarters (1, 2, 3). Researchers have identified bias in grant peer review, including: preference toward established applicants (4), certain areas of study (5), and applicants from prestigious institutions (6). There is bias against female scientists (7, 8), early-career researchers (9), and scientists from minority groups (8, 10).

Grant peer-review has limitations beyond the issue of reviewer bias. Under the concept of 'scientific rigor', grant peer reviewers often: (i) cannot agree on what constitutes good science (11), (ii) assign scores to applications in an arbitrary way (12, 13), (iii) have difficulty estimating future productivity of applicants (14), and (iv) struggle to differentiate between similarly meritorious applications (15, 16). In a study of NIH grant peer, while all reviewers received similar instructions on how to rate and provide feedback, there was no agreement about how reviewer critiques translated to numeric scores. The outcome of grant peer review may depend more on the reviewer than the merits of the proposed research (17). While there have been some suggestions for how to improve grant peer review and reduce potential bias, like lottery systems (see (16), the academic consensus is that there is room to improve the transparency and rigour of grant peer review.

Much of the reporting on issues in grant peer review is based on quantitative analysis of funding or scoring outcomes, often using data from funding agencies (6, 18). Empirical data *quantifies aspects* of grant peer review, but they do not *illuminate the experience* of grant peer review—from the perspective of peer review committee members. In the social sciences, peer reviewers described 5 decision dilemmas when contributing to grant peer review: whether to (1) accept the

review invitation, (2) rely solely on the information included in the application, (3) consider the prestige of the applicant's institution, (4) comment on areas outside their area of expertise, and (5) overlook shortcomings in the application (19). Each peer reviewer brought their own values, priorities and habits to the peer-review work, which influenced the trade-offs they made to resolve their dilemmas (19). We suspected that peer reviewers in health fields in Canada encountered similar decision dilemmas, and we were interested in exploring the trade-offs they made.

In 2009 and in 2016, RAND Europe (www.rand.org) reviewed the effectiveness and efficiency of peer review for grant funding. They also provided lessons and implications for the Canadian Institutes of Health Research (CIHR) grant peer review process, including suggestions to address effectiveness (bias), burden, efficiency, monitoring and evaluation, and improve the evidence base. Seven years later, our team was interested to examine if the key issues in grant peer review remained the same and if any strategies had been implemented to address key concerns (20).

Specifically, we explored the experiences of people who participated in grant peer review at CIHR. We were interested in the perspectives of people who served in different roles on grant peer review committees, their training/preparation for the role, how they handled issues of conflict and bias in the committee meeting, and their preferences for in person versus virtual review. Our overarching research questions were: What is the experience of those who participated in grant peer review panel? What are the challenges in grant peer review and are there strategies to address these challenges?

Context

Grant peer review takes different forms. Perhaps the most common are (i) an expert committee that reviews all grant applications and rates or ranks their quality, and (ii) each application being sent to a small review panel (1 or 2 reviewers) who may provide a final score or contribute to a larger expert panel's discussion and rating or ranking. Some funders use a randomised component once certain criteria are met (21, 22). The peer review committees to that contributed to CIHR's Project Grant Competition peer review operated as expert committees that reviewed all applications and rate or ranked them.

The CIHR Project Grant Competition awards approximately $650 million of CIHR's $1.3 billion annual funding budget. Researchers at any career stage, who wish to conduct health-related research, are eligible to apply. Approximately 60 Peer Review Committees adjudicate about 2000 grant applications across the breadth of the CIHR mandate which spans (i) biomedical, (ii) clinical, (iii) health systems and services, and (iv) population health research themes (now pillars). The committees meet online in spring and fall each year to evaluate and rate each application they are assigned (23). Typically, these meetings occurred in person; until the pandemic necessitated that they occur virtually. Until the fall 2020 Peer Review Committee meetings, CIHR peer review was conducted in-person. Since then, all peer review has been conducted virtually.

For the CIHR Project Grant Competition, a Peer Review Committee comprises up to 20 members (peer reviewers) plus three leaders—Chair and two Scientific Officers—who, with support from CIHR staff, assign applications to reviewers, lead the committee consensus discussion, and summarise the committee discussion in written feedback for peer reviewers. Members are recruited from the CIHR College of Reviewers, nominated by Chairs and/or Scientific Officers, or identified by Internet search (including Canada Research Chairholders list,

Fellows Directory for the Canadian Academy of Health Sciences, publications, conference invited speakers, institutions in regions that are historically under-represented on Committees). When adjudicating each application, the Peer Review Committees are asked to consider (1) the significance and impact of the research, (2) the approaches and methods, and (3) expertise, experience and resources available to deliver on the research project objectives.

Methods

Upon approval from the University of British Columbia (UBC) Research and Ethics Board (H21-03875), we recruited 18 individuals who had participated in a CIHR Project Grant Competition peer review panel at least once as a Committee Member (reviewer), Chair or Scientific Officer. Once a committee completes its work, CIHR posts the names and institutions of reviewers on its public website. CIHR staff identified a list of 50 potential participants who represented the four pillars of CIHR research (biomedical, clinical, health systems and services, population health). Names were selected randomly by several CIHR staff members. One of us (JSG) sent a recruitment email to potential participants. Interested individuals replied via email or telephone. We did not track why participants chose not to respond, 11 sent an email to indicate they did not have time to participate. All who agreed to participate in the study provided verbal informed consent at the beginning of the interview. As per standard ethics practice for qualitative research, participants were informed that their data would be kept anonymous and confidential and that only aggregate themes would be shared. Where quotes are used no attribution is assigned. JSG and CA recruited and interviewed participants on a rolling basis from February to August 2022. JSG, CA and other members of the research team met on a bi-weekly basis to review the transcripts. In keeping with common qualitative practices, the team made the decision to stop recruitment of participants when we determined that the study had reached

saturation (repetition of topics and themes). We used the Consolidated Criteria for Reporting Qualitative Research (COREQ) (24) in the conduct and writing of our study (Appendix 1).

Data collection

Guided by a generic approach to qualitative research (25), the interview guide was developed based on a priori concepts of peer review and the study team's experience with grant review. The interview guide included questions about participants' background, training in grant peer review, strengths and challenges of the review process (including experiences of in-person and virtual peer review), conflict, bias, equity, diversity, inclusion. The interview guide can be found in Appendix 2. JSG and CLA conducted semi-structured interviews with 18 participants via Zoom. Interviews lasted 30–65 minutes. The number of participants in this study is consistent with best practices for qualitative research (26).

Processing & analysis

In accordance with our pre-designed sample (e.g., those who had participated on a grant peer review panel) and a priori topics, we used framework analysis to achieve our objectives. Participants' original accounts anchored and guided our descriptions and observations (27, 28). For analysis, we sifted, charted and sorted data based on key issues and themes using five steps. First, using Zoom, each interview was transcribed verbatim. One team member read the transcripts to obtain a sense of the interviews (Step 1. *familiarize*). Then we combined inductive and deductive approaches to develop a thematic framework. To guide our initial framework, we

first identified themes of significance from the literature. To refine our framework, we incorporated topics that we recognized as frequently occurring in our data (Step 2. *identify a thematic framework*). We then coded all transcripts using the thematic framework established in Step 2. We used the software Nvivo 14 to manage the transcripts and analyse data *(*Steps 3 & 4. *index and chart)*. To compare and contrast themes within and across groups we adopted the constant comparison method; we explored similarities and differences across the data (Step 5. *map and interpret*) (27).

Trustworthiness

Four strategies reinforced the rigor of our study. We cross-checked full transcripts against original audio files for quality and completeness. JSG recorded reflexive memos during data generation and analysis. JSG and CLA met after the interviews to discuss emerging themes. Using NVivo, JSG applied our thematic framework to code full paragraphs of the interviews so that we did not lose contextual meaning. As a team, we discussed themes and those cases that did not "fit within themes". We replaced participants' names with pseudonyms to report results.

Results

Participants ranged in age from 42 to 77 years (mean 53.6 years). Those who identified as women made up 61% of the sample. All participants were either mid-career (5-15 years since their first faculty position) or late-career scholars (15+ years); 67% identified as Caucasian and

17% identified as South Asian. Participant numbers were balanced across all four pillars of CIHR research.

Consistent with findings in the literature on grant peer review, three main themes arose from the analysis of participants' responses: (i) on lack of training and opportunities to learn in particular related to scoring, (ii) differentiating and rating applications of similar strength because reviewers lacked guidelines to assess grants, and in particular those in the meritorious middle, and (iii) an emphasis on reputations and relationships in the review process as a mechanism to distinguish between equally meritorious grants . Two themes related to best practices were: (i) virtual review as a desirable approach to grant peer review, with specific mention of the value of reduced time and cost related to travel (and multiple time zones), and (ii) the essential (and important) role of the Chair in grant peer review. Table 1 shows the identified themes and examples.

Table 1: Study themes, descriptions and illustrative quotes

Theme	Description	Illustrative Quote
Lack of training and limited opportunities to learn	16 participants indicated they had informal or no training which made scoring grants difficult.	*"It [training] has been pretty much experience based."*

Challenges in differentiating and rating applications of similar strength	17 participants had trouble differentiating (and rating) applications that were of similar strength.	*"[there is a] challenge to reliably distinguish a swathe of excellent grants"*
An emphasis on reputations and relationships in the review process	12 participants indicated that when it is unclear how to rate grants, reviewers rely on the reputation of the applicant and/or personal relationships to fill in the blanks.	*"You hope that it's (grant review) based on merit, not who you are, but I have seen a degree of fascination with established career researchers who, in my opinion have not written the best grant proposal, get the benefit of the doubt—let's just call it that."* *"...networking and having the opportunity to learn and to be in a place in a space physically together, where you can get to know people [assists in grant review]."*

Virtual review	13 participants indicated that virtual review is more inclusive of those with caregiving, geographic and economic considerations.	*"I am inclined to a virtual review. For many reasons that also are related with equity -- particularly for women that are a single parent…you know parents of children have a tough time arranging …and that creates an inequity an invisible inequity."*
Role of the Chair	18 participants indicated that the role of the Chair is essential to ensure the integrity of the grant peer review process.	*"it really sort of helps if you have a really good chair".*

Lack of training, limited opportunities to learn creates challenges when assessing grants

In response to questions (e.g., "What training did you receive for your role as a reviewer/Chair/Scientific Officer"), participants drew on their own experiences as both a grant

reviewer and as a grant applicant. They spoke about the lack of formal training for grant peer review; at best it might be considered a "learn as you go" model. Participants drew on their own review philosophy and experiences as an applicant.

"I learned from, you know, some of some of my mentors and when I watched them as chairs and those who brought me into the system and then kind of learned from them."

Participants emphasized the lack of in-person training or systematic feedback for grant peer review but did mention CIHR written guides for reviewers (weblinks to text material).

"I did not receive training for any of those roles. Zero training."

"I mean I was given all the documents you know… the guides to review and so on."

For those who mentioned the availability of resource material, there was no mention of how they used the materials or how useful the materials were, and the lack of training was still emphasized. Participants mentioned that the volunteer role of grant peer reviewer added pressure to their already full list of academic and life commitments. Participants found it challenging to balance their desire to train well to do the peer review role with all their other commitments.

Challenges in differentiating and rating applications of similar strength

Participants indicated that they were challenged to differentiate between grants of a similar strength—the group of grants that take the majority of the Peer Review Committee's work time, which we have termed 'the meritorious middle' (differentiated from the bottom group of applications that are considered 'un-fundable' and the top group of applications that are

considered exceptional). Participants discussed how, without a scientific "fatal flaw" and lack of clarity on how to distinguish one fundable (deserving) grant from another, the decision on a grant's score might be influenced by how interesting the topic was to the reviewers, and not on scientific merit or any clear rating guidelines:

"You can have a lot of grants where there's nothing flawed and there's a solidly proposed piece of work. You know there's nothing wrong with [the] methods—there's nothing that you could pick apart in terms of the theory or the research question. But there's just another grant in the competition that is scored marginally higher because it catches the eye and the interest of the review committee, and it's that intangible kind of interest piece."

"Catching the eye and interest of the review committee" are not best practices described in review guidelines, nor are they a reproducible, equitable or inclusive practice. Similarly, review decisions might be made based on the topic of the research and not the merits of the [very good] application:

"...it's not always dependent on how good you are as a scientist, it's very much dependent on how fashionable your topic is."

While participants described decision making based on "interest" and "fashion" they did not explicitly state how the approach threatened the review process. Rather, participants focused on the lack of clarity and challenges associated in the review of mostly high-quality grants. Participants described a review process that was apt for rating or ranking the outstanding

applications and the weak applications (those considered as not fundable). Peer Review Committee members felt their challenging work was in how to reliably review and score the substantial proportion of grants that were considered 'fundable' (i.e. 'the meritorious middle')::

"... at that point, you may as well throw them down the stairs".

In addition to a sense of frustration, there was also a distinct sense of defeat. Participants felt that there was no clear way to distinguish between the fundable applications. In an exasperated tone, one participant shrugged and stated:

"That is really hard to grapple with in a peer review process……I honestly don't think that the review Committee does a better job than a lottery."

Participants discussed rating and ranking at length, in the context of challenges with the current rating system. Some suggested that the full range of scores are not used when Peer Review Committee members are rating applications. One participant described the problem as "the mushy middle".

"In the mushy middle [is the problem]. The exceptional ones, usually, you know, come through. But ones that are deeply, deeply flawed that really don't need just an edit or bit of a fix, but actually need to go back to the drawing board—we rarely give those really low rankings or really low scores, right? And so, the one thing that, you know, I tend to push for—encourage—is to make sure that the verbal description of the score that you are giving actually reflects your

opinion…we need to work with the full range of scores, so that we can better differentiate the few that are going to be funded."

Participants shared the sentiment that if a grant is not going to be funded, the consensus score (the score the committee decides at the meeting) and the comments must reflect that fact. The words "clarity" and "clear message" were used frequently throughout the interviews when speaking about not fundable grants. One participant exclaimed:

.

"I despise the "this is 3.5", and "that is 3.6" and then 3.7 …it's creeping in that middle range… we need to send a clear message here if this grant going to be funded, if no, then … it needs to be reflected in the score."

Calibration was raised as a strategy to provide clarity in rating grants. The responsibility for calibration landed solely on the Peer Review Committee Chairs.

"I think the Chairs need to quickly establish this is a (outstanding) grant where you've got three reviewers who are like, you know, this is a 4.5, 4.6, 4.7 .. this is where the bar is set, this is where people are agreeing and then maybe identify one grant that everyone agrees wasn't a good grant. And then work your way towards the middle…it's sort of you establishing a floor and a ceiling and I always think that that's a way to ***calibrate*** *people …I began to get a better appreciation (through the review process) that most people still are very uncomfortable with the full-scale concept. And I get it, right? Nobody likes to give anybody a bad score."*

Ranking instead of rating was also suggested as a strategy to improve the review process.

All participants wished for more clarity in the review process, especially when assessing grants from 'the meritorious middle' (although the definition of "middle" varied between participants).

An emphasis on reputations and relationships in the review process to resolve decision dilemmas

The role that personal relationships play within the grant peer review process also reflected a serious threat to grant peer review. Although there is training on bias in the review process, participants noted the absence of strict and clear guidelines for review. As a consequence, unconscious (and sometimes conscious) bias crept into the process. Established researchers (famous by name) could "*receive the benefit of the doubt*" in the review process:

"You hope that it's (grant review) is based on merit, not who you are. But I have seen a degree of fascination with established career researchers who, in my opinion have not written the best grant proposal, get the benefit of the doubt—let's just call it that."

Similarly, another participant described this as *"old school, new school stuff*" and suggested that the reputation of the applicant was prominent in the review process. Another participant reflected on the role that an applicant's curriculum vitae can have on the process in influencing decision making and the unfair (inequitable) advantage afforded to some applicants:

"I still see this happening, particularly with more senior career investigators, they get all excited about a CV that has 150 papers on and I'm like: "the research proposal doesn't make any sense" ...but they have 150 papers, so that must be good, right? ... that is a distinct conscious bias [and] it's persistent now."

"[it's] kind of a human nature that we are all biased in some form or shape ...and I think we do take that into consideration when it's core [to someone's work], because so-and-so is so well known in the field, or has been running this lab for [years]...But the methods aren't very good, you know, so people will say oh we're going to give them the benefit of the doubt so again, I think [the Chair is essential]."

In addition to attributes of the applicant influencing the review process, the use of social moments and "networking" among reviewers during in-person reviews may also serve as a threat to grant peer review because they preference those who are in the room. In discussion of in-person reviews, many participants noted that relationship building, during social times, were important rewards for people who volunteered their time to participate in the peer review process:

"*It's the side conversations sometimes away from the grant review that are enriching and rewarding as part of the process."*

Others noted the indirect benefits of participating in the in-person reviews as the informal networking that occurred:

"the honest truth is that the in-person experience was really as much around getting together with your colleagues, which is always enjoyable, in my opinion."

While some enjoyed the indirect benefits of in-person reviews, others questioned the need for in in-person review.

"Although I agree social connections are important, I'm not sure that our panel meetings should serve that purpose."

Virtual grant peer review

The benefits of virtual peer review were described by many participants as an important development to improving grant peer review. Participants discussed how virtual peer review helped to address the issue of time away from caregiving and for travel:

"I am inclined to a virtual review. For many reasons that also are related with equity – particularly for women that are a single parent…you know parents of children have a tough time arranging …and that creates an inequity, an invisible inequity."

Virtual review can support inclusion of those living in rural environments and promote geographic equity on the panels:

"So, you *know, like we were really not taking geography well into consideration when they form these panels, we're not taking rural people into consideration at all. [virtual reviews can help this]*."

Participants also noted that virtual review can save money, which could then be redistributed:

"We are reducing the cost, I mean, like all the money that is taken for these stupid trips to [city name where the grant reviews typically occur]… how many scholarships can be paid?".

"So if it's a choice between increasing or enhancing support for graduate student programs or postdocs or, you know, anything, then having money spent on bringing everybody to [city name]…[if you] reduce administrative burden, reduce the costs of doing business and put more money into, as I said, there's so many fundable things that don't get funded because there's been no money."

While the majority of participants favoured virtual peer review, some spoke about the value of in person connections. One participant highlighted the value of in person review for new investigators:

"I feel like the discussion was just very different. I feel like from the indirect benefits to the scientific community, as well as to individual investigators, especially new investigators, the opportunity to network and to find other researchers, that you can work with [this] does not happen in the virtual format…the magic happens when you come across somebody that is not in your field, not at your university, but that you think oh wow if we got together and did this. That's really cool and you don't really have that opportunity beyond the in-person panel."

Role of the Chair in clarifying how to assess equally meritorious grants

All participants noted the key role played by the Peer Review Committee Chair in grant peer review. The Chair is a researcher who manages the applications, ensures qualified reviewers are assigned to all applications and chairs the consensus meetings. The Chair role was described as "essential" and critical to grant review:

"it really sort of helps if you have a really good Chair".

Participants noted that an effective Chair guided the conversation and provided much needed direction when disagreements occur. One participant noted, "*I remember that the Chair was very ... elegant in in bringing us back ...into a discussion."* The Chair role was described as that of a facilitator, a mediator and in some cases an arbitrator who makes a final decision. Participants acknowledged the "responsibility" of the Chair to manage conflicts:

"Sometimes discussions can get heated ..., especially if you have a reviewer that really just doesn't like something about the grant and they are going to stand firm, because they really don't think it should be funded,... like managing that—I think that's the responsibility of the Chair."

The role that a Chair plays in minimizing bias was also discussed. Participants noted that while everyone "has bias" ultimately it is the responsibility of the Chair to identify and address bias.

Participants also discussed the role of the Chair in managing more challenging applications, including resubmissions. Lack of clarity around the "mushy middle" was expressed, so too, lack of clarity and consistency in how resubmissions were handled. One participant discussed their role in managing resubmissions as a Chair:

"Most recently, I was Chair of one of the panels, and when resubmissions came up people gave them a regular review. But in their comments they might say "we saw this one before.". And sometimes I've heard comments and I had to had to intercede: they would say "well we've seen this one for the third or fourth time we need to either fund it or not, or give them a very strong message, like this is just not gonna do it.". So, sometimes the reviewer would be trying to push it over that funding line with no other reason than this is the fourth time we've seen this and I'm having to say as the Chair "that's not the reason to fund the grant"."

Chairs helped to clarify the peer review process for reviewers. The Chair was critical in promoting reproducibility and rigor. Beyond scientific skills, participants agreed that Chairs needed excellent interpersonal skills:

"Sometimes it's [the review process] managed well … and [it requires] a lot of its interpersonal skills, more so than scientific skills and how meetings are chairs and how individuals are coached."

Discussion

Grant peer review is inherently an imperfect process. Yet, the scientific community considers it an essential process for identifying the best science for granting agencies to fund. Seven years

after a comprehensive expert review of grant peer review in Canada, which identified key issues such as whether peer review funds the best science and whether it is a reliable process, members of Peer Review Committees continue to struggle with the same issues (20).

Given a crisis of trust in grant peer review (3), we suggest our work helps bring transparency to a process that for many applicants appears frustratingly opaque. In our qualitative study of the opinions of active grant Peer Review Committee members, three key threats to grant peer review surfaced. Participants' voices validated the 2018 experts commentators' review (20) that concluded grant peer review quality was limited by: (i) lack of reviewer and Chair training, (ii) the conundrum of differentiating and rating applications of similar strength, and, (iii) the emphasis on reputations and relationships in the review process to differentiate grant applications of a similar strength. Participants suggested how grant peer review could be improved and also shared potential "roadblocks" to these solutions. The biggest roadblock to improving the grant review process was reviewers' lack of time and the volunteer nature of the role.

Participants described their pathway to become a grant peer reviewer as "learn as you go". In grant peer review, participants drew on their own experiences as an applicant and personal philosophy to understand and navigate the process. Participants spoke at length about time constraints. There was little if any formal standardized training, and where there was discussion, participants highlighted their own time constraints. Where standardized materials had been provided (such as links and PDF documents), reviewers indicated not having reading them carefully or considering the materials as "training". While reporting a craving for standardized

training, many participants felt they did not have the time. Without standardized training participants tended to rely on their own knowledge (and biases) to make decisions.

When participants lacked clear guidance from training, the Chair, or in reference materials, they made their own best decisions about scoring grants. The "mushy middle" or what we refer to as 'the meritorious middle'—the applications that are considered 'fundable' if the funding pool was larger—were challenging to score. Instead, Peer Review Committee members rated applications based on interest, familiarity with the applicants or arbitrarily. The practice is exacerbated in a climate where funding is very constrained (as budgets are being cut or at least not keeping pace with inflation). It was strongly suggested that there needs to be a process to deal with grants that fall into this category. Random allocation of funds (sometimes called a partial lottery) might foster a fairer process (2, 21, 22). While partial lotteries are currently being implemented by other national funders to precisely address these issues, they are not yet implemented by CIHR. An important consideration in the future will be if partial lotteries indeed reduce the time demands on Peer Review Committees.

Participants were uncertain about how to rate grant applications. There is debate about the relative merits of rating (i.e. peer reviewers rate applications on an ordinal scale, e.g. poor to excellent, making an *absolute judgement* against the "ideal") and ranking applications (i.e. peer reviewers make a *relative judgement* to order applications from highest to lowest quality). We studied the reliability of both approaches in the CIHR peer review system, and found that ranking was more reliable, and less susceptible to reviewer expertise and experience (29).

Despite having access to a scoring rubric, participants were unconvinced that rating—especially with the small increments on an ordinal scale—was sufficient to distinguish the 'fundable' applications. There was inherent tension between the bluntness of rating as a tool for allocating funding, and the precision required of the task—ranking might overcome some of the problems. But there were uncertainties about how effective ranking was for addressing the shortcomings of rating. Some participants spoke of calibration, taking the top and bottom grants and using those as yardsticks for scoring (30). We suggest that the current scoring system requires improvements like having Committee members rank applications instead of rating (Tamblyn, Girard et al. 2023) or at least that Peer Review Committees would benefit from comprehensive training on how to use the rating system. Time commitments for training and for the task of reviewing must be considered. Peer Review Committees felt constrained by the amount of funding available: there are many more fundable grants than funds to go around; peer reviewers often described "splitting hairs" and described the extensive time it took to do this work (26).

Although participants highlighted the importance of limiting or eliminating bias in discussions about rating grants, the applicant's reputation was one area that was often considered. Participants tried to avoid bias (i.e. applicant A has 150 publications, so I'll give them the benefit of the doubt and rate the application higher than applicant B whose CV reports 80 publications), yet struggled because was difficult to *not* consider the reputation of applicants. It was a particular challenge when an applicant was considered "famous" in their field. This is the Matthew Effect in grant peer review, where the past success of an established researcher perpetuates future success (4). Early career researchers, researchers who are under-represented in science (e.g.

racialized scholars) and previously unsuccessful applicants are examples of cohorts who are penalised by the Matthew Effect (5, 6, 7, 10).

Participants raised the idea about the merit (or feasibility) of blinding reviewers to the identity of the applicants—a practice used by some funding agencies and in journal peer review—as a way of overcoming bias. In a recent study on journal peer review, when manuscript authors' identities and affiliations were blocked from peer reviewers, unconscious bias was less likely to influence peer review than when the information was available (31). At present in Canada, the applicants' CVs are included with the project information. This raises questions including whether double anonymization is possible in grant peer review, and whether distinguished scientists should be afforded some advantage in grant peer review or whether the research proposal should be judged on its merits alone. At a minimum, our data suggest that funders should continue to provide explicit guidance on whether Peer Review Committees are to consider an applicant's reputation when rating applications.

Our data suggest that—during in-person peer review—social moments and "networking" among reviewers preferences those in the room, and may influence the decisions they make—and this threatens grant peer review. Minoritized researchers often struggle to access mentoring, networking and career development opportunities to progress as independent researchers (32). Social interactions in the context of Peer Review Committee meetings, where reviewers publicly declare their ranking or rating (as occurs in CIHR's Project Grant Committee meetings), could influence peer reviewers' scores and introduce bias (33).

Given the opportunity, participants noted that some members of the Peer Review Committee, although forbidden to do so in the guidelines, would "chat" over dinner about applicants and applications. They would discuss teams that they knew and may also touch on some aspects of the science. Participants noted that the practice of discussing grants outside of the formal review process could influence how committee members might view a team or grant leading to bias; yet the discussions continued to occur. This finding calls into question the value and the potential for bias that is introduced when review committees enjoy social time. Community building through social engagement is important. We argue there are other ways to create those opportunities, without introducing bias in the grant peer review process. One way to eliminate opportunities for socializing during peer review is through virtual peer review.

A reviewer training conference or workshops could fulfill a dual purpose of training and community building. In many other sectors (e.g., jury deliberation) it is common to limit interaction outside of an adjudication process while it is in process. While there is some guidance from those guiding the review process on these informal interactions, it is clearly being breached.

Participants raised two issues that they felt had potential to improve grant peer review: the role of the Chair and virtual grant peer review. Peer-review authority, Professor Gallo, considers the Chair as pivotal to the quality of conversations about grants (34). In our study, the Chair was considered responsible for overseeing the entire process, identifying potential sources of bias and explaining processes and scoring as needed. Participants noted that Chairs did not necessarily have all the answers, and that there was a need for more comprehensive training. Again, time constraints were noted as important considerations for any additional training.

Virtual grant peer review was seen as a way to limit bias and avoid exclusion related to travel and caregiving responsibilities. It was also viewed as one way to eliminate bias associated with in-person, out-of-committee, social moments. Research has shown little difference in peer review outcomes or consistency between virtual and in person review (35). In our study, most participants viewed virtual grant peer review favorably, while some placed high value on the professional networking and socialization that occurred during in-person meetings. These social opportunities are not within the mandate or overall mission of the grant peer review process and could be accomplished elsewhere. We suggest virtual review has the potential to improve the grant peer review process by decreasing time and cost associated with travel, and by reducing opportunities for bias to creep in.

Our study focused on grant peer review in the Canadian context. While we believe many of the findings are likely universal, it is a limitation of the current study. Future research would benefit from the inclusion of other granting agencies.

Conclusions

In this study we highlight three threats to the integrity of grant peer review. In doing so, we underscore the dissonance between reviewers wanting to do better while being constrained by time. As researchers continue to evaluate the threats to grant peer review, the reality of stretched resources and time must be considered. We call on funders to implement practices that reduce reviewer burden, such as a lottery system. Future studies would benefit from a focus on the role of equity, diversity and inclusion practices in the grant peer review process. Processes that are equitable and inclusive for diverse people help to ensure transparency and rigour.

List of Abbreviations

BC: British Columbia

CIHR: Canadian Institutes of Health Research

UBC: University of British Columbia

Declarations

Ethics approval and consent to participate

We obtained approval from the University of British Columbia (UBC) Research and Ethics Board (H21-03875). All participants consented to participated and as per standard ethics practice for qualitative research, participants were informed that their data would be kept anonymous and confidential and that only aggregate themes would be shared. Where quotes are used no attribution is assigned.

Consent for publication

Not applicable.

Patient and Public Involvement

The study was guided by researchers with experience in grant peer review. Given the expertise of our team it was not necessary or appropriate to engage patients or the public in the development of the study.

Availability of data and materials

The datasets generated or analyzed during the current study are not publicly available due to confidentiality requirements for ethics. Data are available from the corresponding author upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

Funding for this study came from KK's personal funding stipend from the Canadian Institutes of Health Research.

Authors' contributions

JSG and CA collected and analyzed the data regarding grant peer review. KK, AL, AM discussed the results and were major contributors in writing the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Bendiscioli S. The troubles with peer review for allocating research funding: Funders need to experiment with versions of peer review and decision-making. EMBO reports. 2019;20(12):e49472.
2. Guthrie S, Ghiga I, Wooding S. What do we know about grant peer review in the health sciences? F1000Res. 2018;6:1335.
3. Langfeldt L, Reymert I, Svartefoss SM. Distrust in grant peer review—reasons and remedies. Science and Public Policy. 2023;51(1):28-41.
4. Huber J, Inoua S, Kerschbamer R, König-Kersting C, Palan S, Smith VL. Nobel and novice: Author prominence affects peer review. PNAS. 2022;119(41):e2205779119.
5. Pina DG, Buljan I, Hren D, Marušić A. A retrospective analysis of the peer review of more than 75,000 Marie Curie proposals between 2007 and 2018. Elife. 2021;10.
6. Doyle JM, Baiocchi MT, Kiernan M. Downstream funding success of early career researchers for resubmitted versus new applications: A matched cohort. PLoS One. 2021;16(11):e0257559.
7. Ginther DK, Kahn S, Schaffer WT. Gender, Race/Ethnicity, and National Institutes of Health R01 Research Awards: Is There Evidence of a Double Bind for Women of Color? Acad Med. 2016;91(8):1098-107.
8. Sato S, Gygax PM, Randall J, Schmid Mast M. The leaky pipeline in research grant peer review and funding decisions: challenges and future directions. High Educ (Dordr). 2021;82(1):145-62.
9. Lauer M, Tabak L, Collins F. The Next Generation Researchers Initiative at NIH. Proceedings of the National Academy of Sciences. 2017;114(45):11801-3.
10. Biernat M, Carnes M, Filut A, Kaatz A. Gender, Race, and Grant Reviews: Translating and Responding to Research Feedback. Pers Soc Psychol Bull. 2020;46(1):140-54.
11. Hesselberg JO, Fostervold KI, Ulleberg P, Svege I. Individual versus general structured feedback to improve agreement in grant peer review: a randomized controlled trial. Res Integr Peer Rev. 2021;6(1):12.
12. Coveney J, Herbert DL, Hill K, Mow KE, Graves N, Barnett A. 'Are you siding with a personality or the grant proposal?': observations on how peer review panels function. Res Integr Peer Rev. 2017;2:19.
13. Graves N, Barnett AG, Clarke P. Funding grant proposals for scientific research: retrospective analysis of scores by members of grant review panel. Bmj. 2011;343:d4797.
14. Fang FC, Bowen A, Casadevall A. NIH peer review percentile scores are poorly predictive of grant productivity. Elife. 2016;5.
15. Li D, Agha L. Big names or big ideas: Do peer-review panels select the best science proposals? Science. 2015;348(6233):434-8.
16. Fang FC, Casadevall A. Research Funding: the Case for a Modified Lottery. mBio. 2016;7(2):e00422-16.
17. Pier EL, Brauer M, Filut A, Kaatz A, Raclaw J, Nathan MJ, et al. Low agreement among reviewers evaluating the same NIH grant applications. Proc Natl Acad Sci U S A. 2018;115(12):2952-7.
18. Batagelj V, Ferligoj A, Squazzoni F. The emergence of a field: a network analysis of research on peer review. Scientometrics. 2017;113(1):503-32.
19. Vallée-Tourangeau G, Wheelock A, Vandrevala T, Harries P. Peer reviewers' dilemmas: a qualitative exploration of decisional conflict in the evaluation of grant applications in the medical humanities and social sciences. Humanit Soc Sci Commun. 2022;9(1).
20. Guthrie S, Ghiga I, Wooding S. What do we know about grant peer review in the health sciences? An updated review of the literature and six case studies. Santa Monica, CA: RAND Corporation; 2018.
21. Horbach SPJM, Tijdink JK, Bouter LM. Partial lottery can make grant allocation more fair, more efficient, and more diverse. Science and Public Policy. 2022;49(4):580-2.

22. Liu M, Choy V, Clarke P, Barnett A, Blakely T, Pomeroy L. The acceptability of using a lottery to allocate research funding: a survey of applicants. Res Integr Peer Rev. 2020;5:3.
23. [Available from: https://cihr-irsc.gc.ca/e/51250.html. .
24. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. International Journal for Quality in Health Care. 2007;19(6):349-57.
25. Cooper S, Endacott R. Generic qualitative research: a design for qualitative research in emergency care? Emergency Medicine Journal. 2007;24(12):816-9.
26. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: A systematic review of empirical tests. Social Science & Medicine. 2022;292:114523.
27. Srivastava A, Thomson SB. Framework analysis: a qualitative methodology for applied policy research. 2009.
28. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. Analyzing qualitative data: Routledge; 2002. p. 187-208.
29. Tamblyn R, Girard N, Hanley J, Habib B, Mota A, Khan KM, et al. Ranking versus rating in peer review of research grant applications. PLoS One. 2023;18(10):e0292306.
30. Pier EL, Raclaw J, Kaatz A, Brauer M, Carnes M, Nathan MJ, et al. 'Your comments are meaner than your score': score calibration talk influences intra-and inter-panel variability during scientific grant peer review. Research Evaluation. 2017;26(1):1-14.
31. Fox CW, Meyer J, Aimé E. Funct Ecol. 2023;37.
32. Zea MC, Bowleg L. The Final Frontier-Transitions and Sustainability: From Mentored to Independent Research. AIDS Behav. 2016;20 Suppl 2(Suppl 2):311-7.
33. Pier EL, Raclaw J, Carnes M, Ford CE, Kaatz A. Laughter and the chair: Social pressures influencing scoring during grant peer review meetings. Journal of General Internal Medicine. 2019;34:513-4.
34. Gallo SA, Schmaling KB, Thompson LA, Glisson SR. Grant reviewer perceptions of the quality, effectiveness, and influence of panel discussion. Res Integr Peer Rev. 2020;5:7.
35. Vo NM, Quiggle GM, Wadhwani K. Comparative outcomes of face-to-face and virtual review meetings. International Journal of Surgery Open. 2016;4:38-41.

Appendix 1: COREQ Checklist

Consolidated criteria for reporting qualitative studies (COREQ): 32-item checklist

No	Item	Guide questions/description	Responses
Domain 1: Research team and reflexivity			
Personal Characteristics			
1.	Interviewer/facilitator	Which author/s conducted the interview or focus group?	JSG, CA
2.	Credentials	What were the researcher's credentials? *E.g. PhD, MD*	PhD, PhD
3.	Occupation	What was their occupation at the time of the study? Researcher, Researcher	Researcher, Researcher
4.	Gender	Was the researcher male or female?	Female, female
5.	Experience and training	What experience or training did the researcher have? Extensive 10 plus years	Extensive (20 years), moderate (3 years)
Relationship with participants			
6.	Relationship established	Was a relationship established prior to study commencement?	Yes
7.	Participant knowledge of the interviewer	What did the participants know about the researcher? e.*g. personal goals, reasons for doing the research*	Rationale for research

No	Item	Guide questions/description	Responses
8.	Interviewer characteristics	What characteristics were reported about the interviewer/facilitator? e.g. *Bias, assumptions, reasons and interests in the research topic*	Reasons and interest in the topic, biases and opinions not shared, but experience in peer review disclosed by interviewers
Domain 2: study design			
Theoretical framework			
9.	Methodological orientation and Theory	What methodological orientation was stated to underpin the study? *e.g. grounded theory, discourse analysis, ethnography, phenomenology, content analysis*	Generic qualitative research
Participant selection			
10.	Sampling	How were participants selected? *e.g. purposive, convenience, consecutive, snowball*	Convenience and purposive
11.	Method of approach	How were participants approached? e.*g. face-to-face, telephone, mail, email*	Email letter
12.	Sample size	How many participants were in the study?	18
13.	Non-participation	How many people refused to participate or dropped out? Reasons?	11 chose not to participate due to time constraints, others did not indicate
Setting			
14.	Setting of data collection	Where was the data collected? e.*g. home, clinic, workplace*	Workplace
15.	Presence of non-participants	Was anyone else present besides the participants and researchers?	No

No	Item	Guide questions/description	Responses
16.	Description of sample	What are the important characteristics of the sample? *e.g. demographic data, date*	Participants represented the different research pillars and ranged in career stage (experience in reviewing)
Data collection			
17.	Interview guide	Were questions, prompts, guides provided by the authors? Was it pilot tested?	Yes guides and prompts were provided. It was pilot tested with our research team members.
18.	Repeat interviews	Were repeat interviews carried out? If yes, how many?	No
19.	Audio/visual recording	Did the research use audio or visual recording to collect the data?	Yes, zoom recordings
20.	Field notes	Were field notes made during and/or after the interview or focus group?	Yes extensive field notes were made during and after the interviews
21.	Duration	What was the duration of the interviews or focus group?	
22.	Data saturation	Was data saturation discussed?	Yes
23.	Transcripts returned	Were transcripts returned to participants for comment and/or correction?	No
Domain 3: analysis and findings			
Data analysis			
24.	Number of data coders	How many data coders coded the data?	Two team members
25.	Description of the coding tree	Did authors provide a description of the coding tree?	Yes, we utilized framework analysis

No	Item	Guide questions/description	Responses
26.	Derivation of themes	Were themes identified in advance or derived from the data?	Both
27.	Software	What software, if applicable, was used to manage the data?	Nvivo
28.	Participant checking	Did participants provide feedback on the findings?	No
Reporting			
29.	Quotations presented	Were participant quotations presented to illustrate the themes / findings? Was each quotation identified? e.*g. participant number*	Yes, quotations presented with pseudonyms
30.	Data and findings consistent	Was there consistency between the data presented and the findings?	Yes
31.	Clarity of major themes	Were major themes clearly presented in the findings?	Yes
32.	Clarity of minor themes	Is there a description of diverse cases or discussion of minor themes?	Yes

683

684

685

Appendix 2: Interview Guide

Demographics

1. What is your primary area of research expertise?
2. What year did you receive your highest degree?
3. What is the year of your first faculty appointment?
4. How many years have you been a CM, SO, Chair?
5. In what year were you born?
6. How do you describe your ethnicity?

Training, Roles and Philosophy

7. What training did you receive to undertake your role [probe: formal, informal, historical]?
8. Do you have suggestions to improve training/support for reviewers? [probe: training for your role, training for chairs/SO's and CM's]
9. Can you please describe your chairing/review philosophy? [probe: how do you set the tone for a review? do you set ground rules, discuss process or conflict up front]
10. What do you consider are the key features of a successful review process?
11. How do you handle resubmissions? [probe: do you consider them as a new grants; what do you look for; how do you message your feedback]

Conflict/Bias

12. What types of conflicts have you experienced during the review process?
    a. How do/did you manage this conflict? [probe for each type of conflict identified]
13. How do you manage interpersonal conflicts?
14. How do you identify and manage micro-aggressions during the review process [probe type of micro-aggressions]?
15. Do you discuss bias in the review process? [probe: how is bias recognized and addressed?]

Other

16. What are the pros and cons of virtual versus in person review versus other models of review? [probe: which is preferred and why; how can we establish connections in a virtual environment (i.e. replace a social dinner)]
17. Thinking more generally, is there anything that CIHR could do to further support you in your role [probe: examples from other review processes you have been involved in]

BMJ Open

Threats to grant peer review: A Qualitative Study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2024-091666.R1
Article Type:	Original research
Date Submitted by the Author:	10-Jan-2025
Complete List of Authors:	Sims Gould, Joanie; The University of British Columbia Department of Family Practice, Department of Family Practice Lasinsky, Anne; University of British Columbia Mota, Adrian; Canadian Institutes of Health Research Khan, Karim; University of British Columbia, Department of Family Practice; Institute of Musculoskeletal Health and Arthritis Ardern, Clare L.; University of British Columbia, Department of Family Practice; University of British Columbia, Department of Physical Therapy
Primary Subject Heading:	Medical publishing and peer review
Secondary Subject Heading:	Qualitative research
Keywords:	Review, Decision Making, Knowledge, Capacity Building, QUALITATIVE RESEARCH

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Threats to grant peer review: A Qualitative Study

Authors

Joanie Sims Gould[1] joanie@joaniesimsgould.com

Anne Lasinsky[2] anne.lasinsky@ubc.ca

Adrian Mota[3] adrian.mota@cihr-irsc.gc.ca

Karim M. Khan[1,5] karim.khan@ubc.ca

Clare L. Ardern[1,6] clare.ardern@ubc.ca

Affiliations

[1] Department of Family Practice, University of British Columbia, Vancouver, British Columbia, Canada

[2] School of Kinesiology, University of British Columbia, Vancouver, British Columbia, Canada

[3] Canadian Institutes of Health Research, Ottawa, Ontario, Canada

[5] Canadian Institutes of Health Research-Institute of Musculoskeletal Health and Arthritis, Vancouver, British Columbia, Canada

[6] Department of Physical Therapy, University of British Columbia, Vancouver, British Columbia, Canada

Abstract

Background & Objectives: Peer review is ubiquitous in evaluating scientific research. While peer review of manuscripts submitted to journals has been widely studied, there has been relatively less attention paid to peer review of grant applications [despite how crucial peer review is to researchers having the means and capacity to conduct research]. There is spirited debate in academic community forums [including on social media] about the perceived benefits and limitations of grant peer review. The aim of our study was to understand the experiences and challenges faced by grant peer reviewers.

Methods: Therefore, we conducted qualitative interviews with 18 members of grant-review panels—the Chairs, peer reviewers and Scientific Officers of a national funding agency—that highlight threats to the integrity of grant peer review.

Results: We identified three threats: (1) lack of training and limited opportunities to learn, (2) challenges in differentiating and rating applications of similar strength, and (3) reviewers weighting reputations and relationships in the review process to differentiate grant applications of a similar strength. These threats were compounded by reviewers' stretched resources or lack of time. Our data also highlighted the essential role of the Chair in ensuring transparency and rigorous grant peer review.

Conclusions: As researchers continue to evaluate the threats to grant peer review, the reality of stretched resources and time must be considered. We call on funders and academic institutions to implement practices that reduce reviewer burden.

Keywords: peer review, research grants, training, time

Strengths and Limitations of this Study

- Strengths:
  - Qualitative interviews with leaders of grant-review panels illuminate the experience of grant peer review.
  - Results provide insight into opportunities to improve the rigour of grant peer view.
- Limitations:
  - Data were collected in the Canadian context with one health funding agency.
  - Participants predominantly had grant peer review experience with one funder.

Background

There are threats to the integrity of the grant peer review process. The merit of grant peer-review—a fundamental element of science—has been questioned in many quarters [1, 2, 3]. Researchers have identified bias in grant peer review, including: preference toward established applicants [4], certain areas of study [5], and applicants from prestigious institutions [6]. There is bias against female scientists [7, 8], early-career researchers [9], and scientists from minority groups [8, 10].

Grant peer-review has limitations beyond the issue of reviewer bias. Under the concept of 'scientific rigor', grant peer reviewers often: (i) cannot agree on what constitutes good science [11], (ii) assign scores to applications in an arbitrary way [12, 13], (iii) have difficulty estimating future productivity of applicants [14], and (iv) struggle to differentiate between similarly meritorious applications [15, 16]. In a study of NIH grant peer, while all reviewers received similar instructions on how to rate and provide feedback, there was no agreement about how reviewer critiques translated to numeric scores. The outcome of grant peer review may depend more on the reviewer than the merits of the proposed research [17]. While there have been some suggestions for how to improve grant peer review and reduce potential bias, like lottery systems [see [16], the academic consensus is that there is room to improve the transparency and rigour of grant peer review.

Much of the reporting on issues in grant peer review is based on quantitative analysis of funding or scoring outcomes, often using data from funding agencies [6, 18]. Empirical data *quantifies aspects* of grant peer review, but they do not *illuminate the experience* of grant peer review—from the perspective of peer review committee members. In the social sciences, peer reviewers described 5 decision dilemmas when contributing to grant peer review: whether to (1) accept the

review invitation, (2) rely solely on the information included in the application, (3) consider the prestige of the applicant's institution, (4) comment on areas outside their area of expertise, and (5) overlook shortcomings in the application [19]. Each peer reviewer brought their own values, priorities and habits to the peer-review work, which influenced the trade-offs they made to resolve their dilemmas [19]. We suspected that peer reviewers in health fields in Canada encountered similar decision dilemmas, and we were interested in exploring the trade-offs they made.

In 2009 and in 2016, RAND Europe (www.rand.org) reviewed the effectiveness and efficiency of peer review for grant funding. They also provided lessons and implications for the Canadian Institutes of Health Research (CIHR) grant peer review process, including suggestions to address effectiveness (bias), burden, efficiency, monitoring and evaluation, and improve the evidence base. Seven years later, our team was interested to examine if the key issues in grant peer review remained the same and if any strategies had been implemented to address key concerns [20].

Specifically, we explored the experiences of people who participated in grant peer review at CIHR. We were interested in the perspectives of people who served in different roles on grant peer review committees, their training/preparation for the role, and how they handled issues of conflict and bias in the committee meeting. Our overarching research questions were: What is the experience of those who participated in grant peer review panel? What are the challenges in grant peer review and are there strategies to address these challenges?

Context

Grant peer review takes different forms. Perhaps the most common are (i) an expert committee that reviews all grant applications and rates or ranks their quality, and (ii) each application being

sent to a small review panel (1 or 2 reviewers) who may provide a final score or contribute to a larger expert panel's discussion and rating or ranking. Some funders use a randomised component once certain criteria are met [21, 22]. The peer review committees that contributed to CIHR's Project Grant Competition peer review operated as expert committees that reviewed all applications and rated or ranked them.

The CIHR Project Grant Competition awards approximately $650 million Canadian dollars of CIHR's $1.3 billion (Canadian dollars) annual funding budget. Researchers at any career stage, who wish to conduct health-related research, are eligible to apply. For each Competition, approximately 60 Peer Review Committees adjudicate about 2000 grant applications across the breadth of the CIHR mandate which spans (i) biomedical, (ii) clinical, (iii) health systems and services, and (iv) population health research themes (now pillars). The committees meet online in spring and autumn each year to evaluate and rate each application they are assigned [23]. Typically, these meetings occurred in person; until the pandemic necessitated that they occur virtually. Until the fall 2020 Peer Review Committee meetings, CIHR peer review was conducted in-person. Since then, all peer review has been conducted virtually.

For the CIHR Project Grant Competition, each Peer Review Committee comprises up to 20 members (peer reviewers) plus three leaders—Chair and two Scientific Officers—who, with support from CIHR staff, assign applications to reviewers, lead the committee consensus discussion, and summarise the committee discussion in written feedback for applicants. Members are recruited from the CIHR College of Reviewers, nominated by Chairs and/or Scientific Officers, or identified by Internet search (including Canada Research Chairholders list, Fellows Directory for the Canadian Academy of Health Sciences, publications, conference invited speakers, institutions in regions that are historically under-represented on Committees).

When adjudicating each application, the Peer Review Committees are asked to consider (1) the significance and impact of the research, (2) the approaches and methods, and (3) expertise, experience and resources available to deliver on the research project objectives.

Peer review occurs in 2 stages. First, all submitted applications are initially reviewed and scored (rated) by a primary reviewer and 2 secondary reviewers, who provide a rating (on a 0-4.9 point rating scale) and written feedback. The second stage of the review process occurs at the Peer Review Committee meeting. Because only about 20% of the applications to the Project Grant Competition are ultimately funded, a streamlining process is first used to eliminate non-competitive applications so that the Committee has the maximum time available to discuss competitive applications. An application is streamlined (i.e. receives 3 ratings and written feedback by is not discussed by the Peer Review Committee or considered for funding) if (i) the average of the reviewers' ratings places the application in the bottom 60% of all applications that the Committee is considering, (ii) at least 1 reviewer has identified the application as non-competitive, and (iii) no Committee member objects to streamlining the application.

For applications that are discussed at the Committee meeting, the 3 reviewers are asked to reach a consensus rating (usually approximately the mean of the reviewers' ratings) after the Committee discussion. Once the consensus rating is announced to the Committee, all Committee members are asked to rate the application [final rating] within +/- 0.5 of the consensus rating. Ultimately, applicants whose applications are discussed receive (i) the final rating (collated by CIHR staff after the Committee meeting), (ii) written feedback from the Scientific Officer capturing the key elements that the Peer Review Committee considered during their discussion, and (iii) the written feedback and ratings from the reviewers.

There are no interviews with applicants and no opportunity for applicants to rebut the Peer Review Committee's feedback during the peer review/grant selection process. Applicants may submit a 2-page Response to Previous Reviews if they choose to re-submit their application to a subsequent Project Grant Competition round.

Methods

Upon approval from the University of British Columbia (UBC) Research and Ethics Board (H21-03875), we recruited 18 individuals who had participated in a CIHR Project Grant Competition peer review panel at least once as a Committee Member (reviewer), Chair or Scientific Officer. Once a committee completes its work, CIHR posts the names and institutions of reviewers on its public website. CIHR staff identified a list of 50 potential participants who represented the four pillars of CIHR research (biomedical, clinical, health systems and services, population health). Names were selected randomly by several CIHR staff members. One of us (JSG) sent a recruitment email to potential participants. Interested individuals replied via email or telephone; the response rate was 36%. We did not track why participants chose not to respond, although 11 people sent an email to indicate they did not have time to participate.

All participants provided verbal informed consent at the beginning of the interview. As per standard ethics practice for qualitative research, participants were informed that their data would be kept anonymous and confidential and that only aggregate themes would be reported. Where quotes are used, no attribution is assigned. JSG and CLA recruited and interviewed participants on a rolling basis from February to August 2022. JSG, CLA and other members of the research team met on a bi-weekly basis to review the transcripts. In keeping with common qualitative practices, the team made the decision to stop recruitment of participants when we determined

that the study had reached saturation (repetition of topics and themes). We used the Consolidated Criteria for Reporting Qualitative Research (COREQ) [24] in the conduct and writing of our study (Appendix 1).

Data collection

Guided by a generic approach to qualitative research [25], the interview guide was developed based on a priori concepts of peer review and the study team's experience with grant review. The interview guide included questions about participants' background, training in grant peer review, strengths and challenges of the review process [including experiences of in-person and virtual peer review], conflict, bias, equity, diversity, inclusion. The interview guide can be found in Appendix 2. JSG and CLA conducted semi-structured interviews with 18 participants via Zoom. Interviews lasted 30–65 minutes. The number of participants in this study is consistent with best practices for qualitative research [26].

Processing & analysis

In accordance with our sample (e.g., those who had participated on a grant peer review panel) and a priori topics, we used framework analysis to achieve our objectives. Participants' original accounts anchored and guided our descriptions and observations [27, 28]. For analysis, we sifted, charted and sorted data based on key issues and themes using five steps. First, using Zoom, each interview was transcribed verbatim. One team member read the transcripts to obtain a sense of the interviews (Step 1. *Familiarize*). Then we combined inductive and deductive approaches to

develop a thematic framework. To guide our initial framework, we first identified themes of significance from the literature. To refine our framework, we incorporated topics that we recognized as frequently occurring in our data (Step 2. *identify a thematic framework*). We then coded all transcripts using the thematic framework established in Step 2. We used the software Nvivo 14 to manage the transcripts and analyse data *(*Steps 3 & 4. *index and chart)*. To compare and contrast themes within and across groups we adopted the constant comparison method; we explored similarities and differences across the data (Step 5. *map and interpret*) [27].

Trustworthiness

Four strategies reinforced the rigor of our study. We cross-checked full transcripts against original audio files for quality and completeness. JSG recorded reflexive memos during data generation and analysis. JSG and CLA met after the interviews to discuss emerging themes. Using NVivo, JSG applied our thematic framework to code full paragraphs of the interviews so that we did not lose contextual meaning. As a team, we discussed themes and those cases that did not "fit within themes". Where there were disagreements [there were very few], we reviewed and discussed the original transcripts, to reach consensus on the theme. We replaced participants' names with pseudonyms to report results.

Results

Participants ranged in age from 42 to 77 years (mean 53.6 years). Those who identified as women made up 61% of the sample. All participants were either mid-career (5-15 years since

their first faculty position) or late-career scholars (15+ years); 67% identified as Caucasian and 17% identified as South Asian. Seven participants, in addition to being a reviewer, had served in the role of Chair. Participant numbers were balanced across all four pillars of CIHR research.

Consistent with findings in the literature on grant peer review, three main themes arose from the analysis of participants' responses: (i) on lack of training and opportunities to learn in particular related to scoring, (ii) differentiating and rating applications of similar strength because reviewers lacked guidelines to assess grants, and in particular those in the meritorious middle, and (iii) an emphasis on reputations and relationships in the review process as a mechanism to distinguish between equally meritorious grants . One theme related to best practices was the essential [and important] role of the Chair in grant peer review. Table 1 shows the identified themes and examples.

[INSERT TABLE 1 here please]

Lack of training, limited opportunities to learn creates challenges when assessing grant applications

In response to questions (e.g., "What training did you receive for your role as a reviewer/Chair/Scientific Officer"), participants drew on their own experiences as both a grant reviewer and as a grant applicant. They spoke about the lack of formal training for grant peer review; at best it might be considered a "learn as you go" model. Participants drew on their own review philosophy and experiences as an applicant.

"I learned from, you know, some of some of my mentors and when I watched them as chairs and those who brought me into the system and then kind of learned from them."

Participants emphasized the lack of in-person training or systematic feedback for grant peer review, but did mention CIHR's written guides for reviewers, which were provided by CIHR to reviewers as weblinks to text material.

"I did not receive training for any of those roles. Zero training."

"I mean I was given all the documents you know… the guides to review and so on."

For those who mentioned the availability of resource material, there was no mention of how they used the materials or how useful the materials were, and the lack of training was still emphasized. Participants mentioned that the volunteer role of grant peer reviewer added pressure to their already full list of academic and life commitments. Participants found it challenging to balance their desire to train well to do the peer review role with all their other commitments.

Challenges in differentiating and rating applications of similar strength

Participants indicated that they were challenged to differentiate between grants of a similar strength—the group of grants that take the majority of the Peer Review Committee's work time, which we have termed 'the meritorious middle' [differentiated from the bottom group of applications that are considered 'un-fundable' and the top group of applications that are considered exceptional]. Participants discussed how, without a scientific "fatal flaw" and lack of clarity on how to distinguish one fundable [deserving] grant from another, the decision on a grant's score might be influenced by how interesting the topic was to the reviewers,:

"You can have a lot of grants where there's nothing flawed and there's a solidly proposed piece of work. You know there's nothing wrong with [the] methods—there's nothing that you could

pick apart in terms of the theory or the research question. But there's just another grant in the competition that is scored marginally higher because it catches the eye and the interest of the review committee, and it's that intangible kind of interest piece."

"Catching the eye and interest of the review committee" are not best practices described in review guidelines, nor are they a reproducible, equitable or inclusive practice. Similarly, review decisions might be made based on the topic of the research and not the merits of the [very good] application:

"...it's not always dependent on how good you are as a scientist, it's very much dependent on how fashionable your topic is."

While participants described decision making based on "interest" and "fashion" they did not explicitly state how the approach threatened the review process. Rather, participants focused on the lack of clarity and challenges associated in the review of mostly high-quality grant applications. Participants described a review process that was apt for rating or ranking the outstanding applications and the weak applications [those considered as not fundable]. Peer Review Committee members felt their challenging work was in how to reliably review and score the substantial proportion of grant applications that were considered 'fundable' [i.e. 'the meritorious middle']::

"... at that point, you may as well throw them down the stairs".

In addition to a sense of frustration, there was also a distinct sense of defeat. Participants felt that there was no clear way to distinguish between the fundable applications. In an exasperated tone, one participant shrugged and stated:

"That is really hard to grapple with in a peer review process……I honestly don't think that the review Committee does a better job than a lottery."

Participants discussed rating and ranking at length, in the context of challenges with the current rating system. Some suggested that the full range of scores are not used when Peer Review Committee members are rating applications. One participant described the problem as "the mushy middle".

"In the mushy middle [is the problem]. The exceptional ones, usually, you know, come through. But ones that are deeply, deeply flawed that really don't need just an edit or bit of a fix, but actually need to go back to the drawing board—we rarely give those really low rankings or really low scores, right? And so, the one thing that, you know, I tend to push for—encourage—is to make sure that the verbal description of the score that you are giving actually reflects your opinion…we need to work with the full range of scores, so that we can better differentiate the few that are going to be funded."

Participants shared the sentiment that if a grant is not going to be funded, the consensus score [the score the committee decides at the meeting] and the comments must reflect that fact. The

words “clarity” and “clear message” were used frequently throughout the interviews when speaking about not fundable grants. One participant exclaimed:

.

“I despise the “this is 3.5”, and “that is 3.6” and then 3.7 …it's creeping in that middle range… we need to send a clear message here if this grant going to be funded, if no, then … it needs to be reflected in the score.”

Calibration [i.e. members of the Committee reaching common ground and tuning [by consensus discussion] their individual interpretations of the application rating system to promote consistency and fairness in how the Committee rated each grant application [29]. For example, the Committee might discuss and agree on what would constitute a rating of 3.5 as opposed to a rating of 4.1. Individual reviewer scores are not re-calculated as z-scores to compensate for systematic differences between reviewers in CIHR’s Project Grant Competition] was raised as a strategy to provide clarity in rating grants. The responsibility for calibration landed solely on the Peer Review Committee Chairs.

“I think the Chairs need to quickly establish this is a [outstanding] grant where you've got three reviewers who are like, you know, this is a 4.5, 4.6, 4.7 .. this is where the bar is set, this is where people are agreeing and then maybe identify one grant that everyone agrees wasn’t a good grant. And then work your way towards the middle…it's sort of you establishing a floor and a ceiling and I always think that that's a way to ***calibrate*** *people …I began to get a better appreciation [through the review process] that most people still are very uncomfortable with the full-scale concept. And I get it, right? Nobody likes to give anybody a bad score.”*

Ranking instead of rating was also suggested as a strategy to improve the review process.

An emphasis on reputations and relationships in the review process to resolve decision dilemmas

The role that personal relationships played within the grant peer review process also reflected a serious threat to grant peer review. Although there was training on bias in the review process, participants noted the absence of strict and clear guidelines for review. As a consequence, unconscious (and sometimes conscious) bias crept into the process. Established researchers (famous by name) could "*receive the benefit of the doubt*" in the review process:

"You hope that it's [grant review] based on merit, not who you are. But I have seen a degree of fascination with established career researchers who, in my opinion have not written the best grant proposal, get the benefit of the doubt—let's just call it that."

Similarly, another participant described this as *"old school, new school stuff*" and suggested that the reputation of the applicant was prominent in the review process. Another participant reflected on the role that an applicant's curriculum vitae can have on the process in influencing decision making and the unfair [inequitable] advantage afforded to some applicants:

"I still see this happening, particularly with more senior career investigators, they get all excited about a CV that has 150 papers on and I'm like: "the research proposal doesn't make any sense"

...but they have 150 papers, so that must be good, right? ... that is a distinct conscious bias [and] it's persistent now."

"[it's] kind of a human nature that we are all biased in some form or shape ...and I think we do take that into consideration when it's core [to someone's work], because so-and-so is so well known in the field, or has been running this lab for [years]...But the methods aren't very good, you know, so people will say oh we're going to give them the benefit of the doubt so again, I think [the Chair is essential]."

In addition to attributes of the applicant influencing the review process, the use of social moments and "networking" among reviewers during in-person reviews may also serve as a threat to grant peer review because they preference those who are in the room. In discussion of in-person reviews, many participants noted that relationship building, during social times, were important rewards for people who volunteered their time to participate in the peer review process:

"*It's the side conversations sometimes away from the grant review that are enriching and rewarding as part of the process."*

Others noted the indirect benefits of participating in the in-person reviews as the informal networking that occurred:

"the honest truth is that the in-person experience was really as much around getting together with your colleagues, which is always enjoyable, in my opinion."

While some enjoyed the indirect benefits of in-person reviews, others questioned the need for in in-person review.

"Although I agree social connections are important, I'm not sure that our panel meetings should serve that purpose."

Role of the Chair in clarifying how to assess equally meritorious grants

All participants noted the key role played by the Peer Review Committee Chair in grant peer review. The Chair is a researcher who manages the applications, ensures qualified reviewers are assigned to all applications and chairs the consensus meetings. The Chair role was described as "essential" and critical to grant review:

"it really sort of helps if you have a really good Chair".

Participants noted that an effective Chair guided the conversation and provided much needed direction when disagreements occur. One participant noted, "*I remember that the Chair was very ... elegant in in bringing us back ...into a discussion."* The Chair role was described as that of a facilitator, a mediator and in some cases an arbitrator who makes a final decision. Participants acknowledged the "responsibility" of the Chair to manage conflicts:

"Sometimes discussions can get heated ..., especially if you have a reviewer that really just doesn't like something about the grant and they are going to stand firm, because they really

don't think it should be funded,… like managing that—I think that's the responsibility of the Chair."

The role that a Chair plays in minimizing bias and ensuring trustworthiness and rigor was also discussed. Participants noted that while everyone "has bias" ultimately it is the responsibility of the Chair to identify and address bias to ensure a rigorous grant review process.

Participants also discussed the role of the Chair in managing more challenging applications, including resubmissions. Lack of clarity around the "mushy middle" was expressed, so too, lack of clarity and consistency in how resubmissions were handled. One participant discussed their role in managing resubmissions as a Chair:

"Most recently, I was Chair of one of the panels, and when resubmissions came up people gave them a regular review. But in their comments they might say "we saw this one before.". And sometimes I've heard comments and I had to had to intercede: they would say "well we've seen this one for the third or fourth time we need to either fund it or not, or give them a very strong message, like this is just not gonna do it.". So, sometimes the reviewer would be trying to push it over that funding line with no other reason than this is the fourth time we've seen this and I'm having to say as the Chair "that's not the reason to fund the grant"."

Chairs helped to clarify the peer review process for reviewers. The Chair was critical in promoting reproducibility and rigor. Beyond scientific skills, participants agreed that Chairs needed excellent interpersonal skills:

"Sometimes it's [the review process] managed well … and [it requires] a lot of its interpersonal skills, more so than scientific skills and how meetings are chairs and how individuals are coached."

Discussion

Grant peer review is inherently an imperfect process. Yet, the scientific community considers it essential for identifying the best science for granting agencies to fund. Seven years after a comprehensive expert review of grant peer review in Canada, which identified key issues such as whether peer review funds the best science and whether it is a reliable process, members of Peer Review Committees continue to struggle with the same issues [20].

Given a crisis of trust in grant peer review [3], we describe the challenges of a process that for many applicants appears frustratingly opaque. In our qualitative study of the opinions of active grant Peer Review Committee members, three key threats to grant peer review surfaced. Participants' voices validated the 2018 experts commentators' review [20] that concluded grant peer review quality was limited by: (i) lack of reviewer and Chair training, (ii) the conundrum of differentiating and rating applications of similar strength, and, (iii) the emphasis on reputations and relationships in the review process to differentiate grant applications of a similar strength. Participants suggested how grant peer review could be improved and also shared potential "roadblocks" to these solutions. The biggest roadblock to improving the grant review process was reviewers' lack of time and the volunteer nature of the role.

Participants described their pathway to become a grant peer reviewer as "learn as you go". In grant peer review, participants drew on their own experiences as an applicant and personal philosophy to understand and navigate the process. Participants spoke at length about time constraints. There was little, if any, formal and/or standardized training; where there was discussion, participants highlighted their own time constraints. Where standardized materials had been provided by CIHR to reviewers [such as links and PDF documents], reviewers indicated not having read them carefully or considering the materials as "training". While reporting a craving for standardized training, many participants felt they did not have the time to prioritise completing the training. Without training, participants tended to rely on their own knowledge [and biases] to make decisions.

When participants lacked clear guidance from training, the Chair, or in reference materials, they made their own best decisions about scoring grants. The "mushy middle" or what we refer to as 'the meritorious middle'—the applications that are considered 'fundable' if the funding pool was larger—were challenging to score. Instead, Peer Review Committee members rated applications based on interest, familiarity with the applicants, or arbitrarily. The practice was exacerbated in a climate where funding is very constrained [as budgets being cut or at least not keeping pace with inflation]. It was strongly suggested that a process is needed to deal with grants that fall into the 'meritorious middle' category. Random allocation of funds (sometimes called a partial lottery) might foster a fairer process [2, 21, 22]. While partial lotteries are currently being implemented by other national funders to precisely address these issues, they are not yet implemented by CIHR. An important consideration in the future will be if partial lotteries reduce the time demands on Peer Review Committees.

Participants were uncertain about how to rate grant applications. There is debate about the relative merits of rating (i.e. peer reviewers rate applications on an ordinal scale, e.g. poor to excellent, making an *absolute judgement* against the "ideal") and ranking applications (i.e. peer reviewers make a *relative judgement* to order applications from highest to lowest quality). We studied the reliability of both approaches in the CIHR peer review system, and found that ranking was more reliable, and less susceptible to reviewer expertise and experience [30].

Despite having access to a scoring rubric, participants were unconvinced that rating—especially with the small increments on an ordinal scale—was sufficient to distinguish the 'fundable' applications. There was inherent tension between the bluntness of rating as a tool for allocating funding, and the precision required of the task—ranking might overcome some of the problems. But there were uncertainties about how effective ranking was for addressing the shortcomings of rating. Some participants spoke of calibration, taking the top and bottom grants and using those as yardsticks for scoring [31]. We suggest that the current scoring system requires improvements like having Committee members rank applications instead of rating [30] or at least that Peer Review Committees would benefit from comprehensive training on how to use the rating system. Time commitments for training and for the task of reviewing must be considered by funders and academic institutions. Peer Review Committees felt constrained by the amount of funding available: there are many more fundable grants than funds to go around; peer reviewers often described "splitting hairs" and described the extensive time it took to do this work [26].

Although participants highlighted the importance of limiting or eliminating bias in discussions about rating grants, the applicant's reputation was one area that was often considered. Participants tried to avoid bias (i.e. "*applicant A has 150 publications, so I'll give them the benefit of the doubt and rate the application higher than applicant B whose CV reports 80 publications*") yet struggled because was difficult to *ignore* the reputation of applicants. It was a particular challenge when an applicant was considered "famous" in their field. This is the Matthew Effect in grant peer review, where the past success of an established researcher perpetuates future success [4]. Early career researchers, researchers who are under-represented in science [e.g. racialized scholars] and previously unsuccessful applicants are examples of cohorts who are penalised by the Matthew Effect [5, 6, 7, 10].

Participants raised the idea about the merit (or feasibility) of blinding reviewers to the identity of the applicants—a practice used by some funding agencies and in journal peer review—as a way of overcoming bias. In journal peer review, when manuscript authors' identities and affiliations were blocked from peer reviewers, unconscious bias was less likely to influence peer review than when the information was available [32] thus fostering a less biased review. At present in Canada, the applicants' CVs are included with the project information. This raises questions including whether double anonymization is possible in grant peer review, and whether distinguished scientists should be afforded some advantage in grant peer review, or whether the research proposal should be judged on its merits alone. At a minimum, our data suggest that funders should continue to provide explicit guidance on whether Peer Review Committees are to consider an applicant's reputation when rating applications.

Our data suggest that during in-person peer review, social moments and "networking" among reviewers preferences those in the room, and may influence the decisions they make—and this threatens grant peer review. Minoritized researchers often struggle to access mentoring, networking and career development opportunities to progress as independent researchers [33]. Social interactions in the context of Peer Review Committee meetings, where reviewers publicly declare their ranking or rating [as occurs in CIHR's Project Grant Committee meetings], could influence peer reviewers' scores and introduce bias [34].Given the opportunity, participants noted that some members of the Peer Review Committee, although forbidden to do so in the guidelines, would "chat" over dinner about applicants and applications. They would discuss teams that they knew and may also touch on some aspects of the science. Participants noted that the practice of discussing grants outside of the formal review process could influence how committee members might view a team or grant leading to bias; yet the discussions continued to occur. This finding calls into question the value and the potential for bias that is introduced when review committees enjoy social time. Community building through social engagement is important. We argue there are other ways to create those opportunities, without introducing bias in the grant peer review process. A reviewer training conference or workshops could fulfill a dual purpose of training and community building. In many other sectors (e.g., jury deliberation) it is common to limit interaction outside of an adjudication process while it is in process. While there is some guidance from those guiding the review process on these informal interactions, it is clearly being breached.

To improve the review process, participants noted the essential role of the Chair. Peer-review authority, Professor Gallo, considers the Chair as pivotal to the quality of conversations about

grants [35]. In our study, the Chair was considered responsible for overseeing the entire process, identifying potential sources of bias and explaining processes and scoring as needed to ensure rigour. Participants noted that Chairs did not necessarily have all the answers, and that there was a need for more comprehensive training. Again, time constraints were noted as important considerations for any additional training.

Limitations

Our study focused on grant peer review by one health agency in the Canadian context with 18 reviewers. Most of the reviewers had experience in reviewing with only CIHR. While we believe many of the findings are likely universal, these are limitations of the current study. Future research would benefit from the inclusion of other granting agencies in other countries. Future research would also benefit from interviews with reviewers with experience from other granting agencies.

Conclusions

We highlight three threats to the integrity of grant peer review: (i) lack of training and opportunities to learn in particular related to scoring, (ii) differentiating and rating applications of similar strength because reviewers lacked guidelines to assess grants, and in particular those in the meritorious middle, and (iii) an emphasis on reputations and relationships in the review process as a mechanism to distinguish between equally meritorious grants. We underscore the dissonance between reviewers wanting to do better while being constrained by time. As researchers continue to evaluate the threats to grant peer review, the reality of stretched resources

and time must be considered. We call on funders to implement practices that reduce reviewer burden, such as a lottery system. We also suggest that academic institutions could (i) do more to ensure that researchers have protected time for peer review tasks and opportunities to refine and develop their skills as reviewers, and (ii) make peer reviewer training a mandatory part of the curriculum for PhD students and postdoctoral researchers. Future studies would benefit from a focus on the role of equity, diversity and inclusion practices in the grant peer review process. Processes that are equitable and inclusive for diverse people help to ensure transparency and rigour.

List of Abbreviations

BC: British Columbia

CIHR: Canadian Institutes of Health Research

UBC: University of British Columbia

Declarations

Ethics approval and consent to participate

We obtained approval from the University of British Columbia (UBC) Research and Ethics Board (H21-03875). All participants consented to participated and as per standard ethics practice for qualitative research, participants were informed that their data would be kept anonymous and confidential and that only aggregate themes would be shared. Where quotes are used no attribution is assigned.

Consent for publication

Not applicable.

Patient and Public Involvement

The study was guided by researchers with experience in grant peer review. Given the expertise of our team it was not necessary or appropriate to engage patients or the public in the development of the study.

Availability of data and materials

The datasets generated or analyzed during the current study are not publicly available due to confidentiality requirements for ethics. Data are available from the corresponding author upon reasonable request. We will consider requests for data in an aggregate form [i.e. the coded or themed data], and any requests must identify the specific area of interest for which the data request is made.

Competing interests

The authors declare that they have no competing interests.

Funding

Funding for this study came from KK's personal funding stipend from the Canadian Institutes of Health Research.

Authors' contributions

JSG are responsible for the overall content as guarantor. JSG and CA collected and analyzed the data regarding grant peer review. KK, AL, AM discussed the results and were major contributors in writing the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Bendiscioli S. The troubles with peer review for allocating research funding: Funders need to experiment with versions of peer review and decision-making. EMBO reports. 2019;20[12]:e49472.
2. Guthrie S, Ghiga I, Wooding S. What do we know about grant peer review in the health sciences? F1000Res. 2018;6:1335.
3. Langfeldt L, Reymert I, Svartefoss SM. Distrust in grant peer review—reasons and remedies. Science and Public Policy. 2023;51[1]:28-41.
4. Huber J, Inoua S, Kerschbamer R, König-Kersting C, Palan S, Smith VL. Nobel and novice: Author prominence affects peer review. PNAS. 2022;119[41]:e2205779119.
5. Pina DG, Buljan I, Hren D, Marušić A. A retrospective analysis of the peer review of more than 75,000 Marie Curie proposals between 2007 and 2018. Elife. 2021;10.
6. Doyle JM, Baiocchi MT, Kiernan M. Downstream funding success of early career researchers for resubmitted versus new applications: A matched cohort. PLoS One. 2021;16[11]:e0257559.
7. Ginther DK, Kahn S, Schaffer WT. Gender, Race/Ethnicity, and National Institutes of Health R01 Research Awards: Is There Evidence of a Double Bind for Women of Color? Acad Med. 2016;91[8]:1098-107.
8. Sato S, Gygax PM, Randall J, Schmid Mast M. The leaky pipeline in research grant peer review and funding decisions: challenges and future directions. High Educ [Dordr]. 2021;82[1]:145-62.
9. Lauer M, Tabak L, Collins F. The Next Generation Researchers Initiative at NIH. Proceedings of the National Academy of Sciences. 2017;114[45]:11801-3.
10. Biernat M, Carnes M, Filut A, Kaatz A. Gender, Race, and Grant Reviews: Translating and Responding to Research Feedback. Pers Soc Psychol Bull. 2020;46[1]:140-54.
11. Hesselberg JO, Fostervold KI, Ulleberg P, Svege I. Individual versus general structured feedback to improve agreement in grant peer review: a randomized controlled trial. Res Integr Peer Rev. 2021;6[1]:12.
12. Coveney J, Herbert DL, Hill K, Mow KE, Graves N, Barnett A. 'Are you siding with a personality or the grant proposal?': observations on how peer review panels function. Res Integr Peer Rev. 2017;2:19.
13. Graves N, Barnett AG, Clarke P. Funding grant proposals for scientific research: retrospective analysis of scores by members of grant review panel. Bmj. 2011;343:d4797.
14. Fang FC, Bowen A, Casadevall A. NIH peer review percentile scores are poorly predictive of grant productivity. Elife. 2016;5.
15. Li D, Agha L. Big names or big ideas: Do peer-review panels select the best science proposals? Science. 2015;348[6233]:434-8.
16. Fang FC, Casadevall A. Research Funding: the Case for a Modified Lottery. mBio. 2016;7[2]:e00422-16.
17. Pier EL, Brauer M, Filut A, Kaatz A, Raclaw J, Nathan MJ, et al. Low agreement among reviewers evaluating the same NIH grant applications. Proc Natl Acad Sci U S A. 2018;115[12]:2952-7.
18. Batagelj V, Ferligoj A, Squazzoni F. The emergence of a field: a network analysis of research on peer review. Scientometrics. 2017;113[1]:503-32.
19. Vallée-Tourangeau G, Wheelock A, Vandrevala T, Harries P. Peer reviewers' dilemmas: a qualitative exploration of decisional conflict in the evaluation of grant applications in the medical humanities and social sciences. Humanit Soc Sci Commun. 2022;9[1].
20. Guthrie S, Ghiga I, Wooding S. What do we know about grant peer review in the health sciences? An updated review of the literature and six case studies. Santa Monica, CA: RAND Corporation; 2018.
21. Horbach SPJM, Tijdink JK, Bouter LM. Partial lottery can make grant allocation more fair, more efficient, and more diverse. Science and Public Policy. 2022;49[4]:580-2.

22. Liu M, Choy V, Clarke P, Barnett A, Blakely T, Pomeroy L. The acceptability of using a lottery to allocate research funding: a survey of applicants. Res Integr Peer Rev. 2020;5:3.
23. [Available from: https://cihr-irsc.gc.ca/e/51250.html. .
24. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research [COREQ]: a 32-item checklist for interviews and focus groups. International Journal for Quality in Health Care. 2007;19[6]:349-57.
25. Cooper S, Endacott R. Generic qualitative research: a design for qualitative research in emergency care? Emergency Medicine Journal. 2007;24[12]:816-9.
26. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: A systematic review of empirical tests. Social Science & Medicine. 2022;292:114523.
27. Srivastava A, Thomson SB. Framework analysis: a qualitative methodology for applied policy research. 2009.
28. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. Analyzing qualitative data: Routledge; 2002. p. 187-208.
29. Pier EL, Raclaw J, Kaatz A, Brauer M, Carnes M, Nathan MJ, et al. 'Your comments are meaner than your score': score calibration talk influences intra- and inter-panel variability during scientific grant peer review. Res Eval. 2017;26[1]:1-14.
30. Tamblyn R, Girard N, Hanley J, Habib B, Mota A, Khan KM, et al. Ranking versus rating in peer review of research grant applications. PLoS One. 2023;18[10]:e0292306.
31. Pier EL, Raclaw J, Kaatz A, Brauer M, Carnes M, Nathan MJ, et al. 'Your comments are meaner than your score': score calibration talk influences intra-and inter-panel variability during scientific grant peer review. Research Evaluation. 2017;26[1]:1-14.
32. Fox CW, Meyer J, Aimé E. Funct Ecol. 2023;37.
33. Zea MC, Bowleg L. The Final Frontier-Transitions and Sustainability: From Mentored to Independent Research. AIDS Behav. 2016;20 Suppl 2[Suppl 2]:311-7.
34. Pier EL, Raclaw J, Carnes M, Ford CE, Kaatz A. Laughter and the chair: Social pressures influencing scoring during grant peer review meetings. Journal of General Internal Medicine. 2019;34:513-4.
35. Gallo SA, Schmaling KB, Thompson LA, Glisson SR. Grant reviewer perceptions of the quality, effectiveness, and influence of panel discussion. Res Integr Peer Rev. 2020;5:7.

Table 1: Study themes, descriptions and illustrative quotes

Theme	Description	Illustrative Quote
Lack of training and limited opportunities to learn	16 participants indicated they had informal or no training which made scoring	*"It [training] has been pretty much experience based."*

	grants difficult.	
Challenges in differentiating and rating applications of similar strength	17 participants had trouble differentiating (and rating) applications that were of similar strength.	*"[there is a] challenge to reliably distinguish a swathe of excellent grants"*
An emphasis on reputations and relationships in the review process	12 participants indicated that when it is unclear how to rate grants, reviewers rely on the reputation of the applicant and/or personal relationships to fill in the blanks.	*"You hope that it's (grant review) based on merit, not who you are, but I have seen a degree of fascination with established career researchers who, in my opinion have not written the best grant proposal, get the benefit of the doubt—let's just call it that."* *"...networking and having the opportunity to learn and to be in a place in a space physically together, where you can get to know people [assists in grant review]."*

Role of the Chair	18 participants indicated that the role of the Chair is essential to ensure the integrity of the grant peer review process.	*"it really sort of helps if you have a really good chair".*

Appendix 1: COREQ Checklist

Consolidated criteria for reporting qualitative studies (COREQ): 32-item checklist

No	Item	Guide questions/description	Responses
Domain 1: Research team and reflexivity			
Personal Characteristics			
1.	Interviewer/facilitator	Which author/s conducted the interview or focus group?	JSG, CA
2.	Credentials	What were the researcher's credentials? *E.g. PhD, MD*	PhD, PhD
3.	Occupation	What was their occupation at the time of the study? Researcher, Researcher	Researcher, Researcher
4.	Gender	Was the researcher male or female?	Female, female
5.	Experience and training	What experience or training did the researcher have? Extensive 10 plus years	Extensive (20 years), moderate (3 years)
Relationship with participants			
6.	Relationship established	Was a relationship established prior to study commencement?	Yes
7.	Participant knowledge of the interviewer	What did the participants know about the researcher? e.*g. personal goals, reasons for doing the research*	Rationale for research

No	Item	Guide questions/description	Responses
8.	Interviewer characteristics	What characteristics were reported about the interviewer/facilitator? e.g. *Bias, assumptions, reasons and interests in the research topic*	Reasons and interest in the topic, biases and opinions not shared, but experience in peer review disclosed by interviewers
Domain 2: study design			
Theoretical framework			
9.	Methodological orientation and Theory	What methodological orientation was stated to underpin the study? *e.g. grounded theory, discourse analysis, ethnography, phenomenology, content analysis*	Generic qualitative research
Participant selection			
10.	Sampling	How were participants selected? *e.g. purposive, convenience, consecutive, snowball*	Convenience and purposive
11.	Method of approach	How were participants approached? e.*g. face-to-face, telephone, mail, email*	Email letter
12.	Sample size	How many participants were in the study?	18
13.	Non-participation	How many people refused to participate or dropped out? Reasons?	11 chose not to participate due to time constraints, others did not indicate
Setting			
14.	Setting of data collection	Where was the data collected? e.*g. home, clinic, workplace*	Workplace
15.	Presence of non-participants	Was anyone else present besides the participants and researchers?	No

No	Item	Guide questions/description	Responses
16.	Description of sample	What are the important characteristics of the sample? *e.g. demographic data, date*	Participants represented the different research pillars and ranged in career stage (experience in reviewing)
Data collection			
17.	Interview guide	Were questions, prompts, guides provided by the authors? Was it pilot tested?	Yes guides and prompts were provided. It was pilot tested with our research team members.
18.	Repeat interviews	Were repeat interviews carried out? If yes, how many?	No
19.	Audio/visual recording	Did the research use audio or visual recording to collect the data?	Yes, zoom recordings
20.	Field notes	Were field notes made during and/or after the interview or focus group?	Yes extensive field notes were made during and after the interviews
21.	Duration	What was the duration of the interviews or focus group?	
22.	Data saturation	Was data saturation discussed?	Yes
23.	Transcripts returned	Were transcripts returned to participants for comment and/or correction?	No
Domain 3: analysis and findings			
Data analysis			
24.	Number of data coders	How many data coders coded the data?	Two team members
25.	Description of the coding tree	Did authors provide a description of the coding tree?	Yes, we utilized framework analysis

No	Item	Guide questions/description	Responses
26.	Derivation of themes	Were themes identified in advance or derived from the data?	Both
27.	Software	What software, if applicable, was used to manage the data?	Nvivo
28.	Participant checking	Did participants provide feedback on the findings?	No
Reporting			
29.	Quotations presented	Were participant quotations presented to illustrate the themes / findings? Was each quotation identified? e.*g. participant number*	Yes, quotations presented with pseudonyms
30.	Data and findings consistent	Was there consistency between the data presented and the findings?	Yes
31.	Clarity of major themes	Were major themes clearly presented in the findings?	Yes
32.	Clarity of minor themes	Is there a description of diverse cases or discussion of minor themes?	Yes

Appendix 2: Interview Guide

Demographics

1. What is your primary area of research expertise?
2. What year did you receive your highest degree?
3. What is the year of your first faculty appointment?
4. How many years have you been a CM, SO, Chair?
5. In what year were you born?
6. How do you describe your ethnicity?

Training, Roles and Philosophy

7. What training did you receive to undertake your role [probe: formal, informal, historical]?
8. Do you have suggestions to improve training/support for reviewers? [probe: training for your role, training for chairs/SO's and CM's]
9. Can you please describe your chairing/review philosophy? [probe: how do you set the tone for a review? do you set ground rules, discuss process or conflict up front]
10. What do you consider are the key features of a successful review process?
11. How do you handle resubmissions? [probe: do you consider them as a new grants; what do you look for; how do you message your feedback]

Conflict/Bias

12. What types of conflicts have you experienced during the review process?
    a. How do/did you manage this conflict? [probe for each type of conflict identified]
13. How do you manage interpersonal conflicts?
14. How do you identify and manage micro-aggressions during the review process [probe type of micro-aggressions]?
15. Do you discuss bias in the review process? [probe: how is bias recognized and addressed?]

Other

16. What are the pros and cons of virtual versus in person review versus other models of review? [probe: which is preferred and why; how can we establish connections in a virtual environment (i.e. replace a social dinner)]
17. Thinking more generally, is there anything that CIHR could do to further support you in your role [probe: examples from other review processes you have been involved in]