



Peer Review ORD Response Letter

Executive Summary

We received approximately 150 responses from VA researchers throughout the country regarding the new proposed peer review process. The AAMC was asked and subsequently provided additional insight into this topic. The Office of Research and Development (ORD) would like to thank all of the VA researchers and the AAMC who took time out of their hectic schedules to provide input and guidance. I have personally, along with several of my ORD senior staff, read through all of the suggestions/responses. We firmly believe that many minds are better than a few. Thus, we are highly supportive of an open dialogue to solicit field input that will move us together toward the goal of attaining the best possible VA Research Program.

We know that some of the responses from the field have been widely circulated. However, we thought it was important to provide excerpts of some poignant comments that were provided to us. These selected statements demonstrate the quality and breadth of the feedback, which heavily guided the final decisions concerning the peer review process within ORD. In selecting the quotes, we tried to choose ones that accurately and best portrayed the major sentiments of the field.

Selected Responses from the Field on the Peer Review Process

Issue 1

1) Should productivity be assessed in the awarding of VA Merit Review projects? And if so, *WHO* should perform the evaluation?

Sample responses from the field on the assessment of productivity included:

“... At this point, I think the message should be strongly communicated with investigators how important it is to demonstrate productivity. I remember long ago being told, ‘research not published is research not done.’ Currently, this discussion occurs in some VA peer reviews, but inconsistently.”

“I think productivity should be part of the review process ...Ottenback states that it’s not research until it is disseminated to the scientific community so it can further knowledge development.”

“Peer review is best done by experts in the particular field of study, and it is those experts who are most capable of fairly and reliably judging an investigator’s productivity and contribution to the field of study.”

“All researchers need to be engaged in fashioning productivity standards. It is not reasonable to expect that there will be no criteria.”

After a comprehensive evaluation of all input received from the field and the AAMC, we have concluded that there is agreement on many issues. The overwhelming majority stated that: productivity should be assessed during the VA Merit Review process; assessment should apply uniformly across all four VA Research Services; and that a *peer review committee* (representing the peers of the applicant) should jointly complete the productivity assessment, not the ORD administrative staff.

Issue 2

2) How should this productivity assessment process be guided?

We agree, as noted by one responder:

“It is close to impossible to satisfy everyone for any one decision or policy. And we all know that change of any kind is painful.”

Though there were some who endorsed the process of trying to develop rigid quantitative guidelines (in fact, to a greater extension than what ORD initially proposed), this did not represent the vast majority of the responses. We strived to identify an approach that would satisfy the great majority. A few sample comments supporting rigid and quantitative guidelines included:

“[The] VHA does need a better means to track productivity and performance of its VA funded researchers and clinical research scientists. The creation of a grading system to quantitatively assess performance and productivity, especially if the goals and expectations were realistic, would be of value.”

“The (ORD) proposed guidelines make sense and would lead to fair, unbiased assessment of productivity.”

The quotes highlighted above, while we believe have merit, were not the predominant thrust of input received. Instead, many VA researchers in the field encouraged a more balanced approach but in a less prescriptive and rigid format. These individuals suggested that productivity evaluations should be based on broad qualitative guidelines. A sample of such comments appears below:

“ORD’s proposed productivity assessment system is an attempt to deal with a real problem. However, there seem to be at least two serious flaws in ORD’s proposed

productivity assessment. First, the system by including numerical targets (N of publications, N of grants) overemphasizes counting over the more difficult judgments of quality and importance, or considerations of the amount of work required for publication in various areas. Second, only the criteria for an extreme scenario are included (full professor with extremely high fractions on VA research). Reviewers are instructed to adjust this target for lower fractions and less seniority. But research has clearly established that humans are very bad at making adjustments to numerical targets.”

“...the performance criteria as written are too rigid, and would only promote ‘salami science’ (slicing up your work, like a salami, into the least publishable unit).”

“I commend your work on this difficult, but important task...(However) a formulaic assessment would preclude any meaningful evaluation from a committee of scientific peers. It would become a VACO ORD administrative staff operation.”

“...a document that would describe the various categories important to the VA should be developed and the peer review committees could be asked to provide a separate score to both the proposal and the PI based on these guidelines.”

“It is quality, not quantity, of publications that should be the basis of evaluation, and no numerical schema can substitute here for experience, expertise, and seasoned judgment.”

“I urge VA R&D and the CRADO to focus on providing guidelines for reviewers’ analysis of the variables, NOT a quantitative measure that cannot possibly measure the many variables that a competent reviewer will consider in her/his analysis of a proposal.”

“I [would] favor a quantitative assessment of productivity by the peer review committee. ...these assessments have not always been carefully documented. ... Because of the broad disciplines covered, these standards cannot be applicable to all the investigators reviewed by a single subcommittee, let alone all covered by R&D as a whole. ... The ultimate measure of productivity is to what degree the results of the investigation advance the field, and specifically improve the care of veterans, and this may be only distantly related to any absolute number of publications.”

We strongly agree with all of the responses highlighted above. We believed from the start that it would be nearly impossible to develop a universal formula when assessing productivity. Further, we agree if this was done, it would translate to “VACO ORD administrative staff operations” as one noteworthy responder stated. I was particularly struck by the “salami” comment and agree with the point that quantitative metrics for publication could lead to diluting the quality of any individual manuscript, “gaming” the

system, and rewarding less scholarly efforts/publications simply through counting numbers.

Issue 3

3) What should be considered and should junior investigators also be assessed?

The following are representative quotes received from the field on the topic of who should be assessed and the criteria that should be used when evaluating productivity:

“The concept of using external funds to leverage and maximize VA research is important...The NIH model of peer review includes an ‘Investigators’ category when determining a grant’s score. While there is a broad definition of what to evaluate, the reviewers evaluate the impact the investigators have had in their specific field. Experts in the field that are familiar with the relevant factors for the investigator being reviewed apply criteria.”

“Any productivity score should be an overall score, taking into account the whole picture of funding, publishing (including actually looking at perhaps 3 papers selected by the applicant as most representative, to judge how substantial they are), national recognition (service on study sections, national committees, talks at annual meetings), etc. The VA should not impose upon reviewers strict numerical criteria such as exact number of papers...”

“An important part of the mission of the VA, and of investigators within the VA, is the training of the next generation of researchers.”

“Although poor productivity may in general be a reasonable reason to refuse further funding, there should be some kind of an ‘escape hatch’. We have all seen productive investigators hit a rough patch and then recover because of other academic or clinical responsibilities, moving between institutions, losing a grant, personal problems, etc. Although non-productive investigators often remain non-productive, there ought to be a way that a previously productive investigator who has had a problem, that has now clearly been corrected, may be given another chance if the reviewers feel appropriate.”

“You should be applauded for the apparent effort to promote the young investigator by eliminating productivity analysis for applications from investigators who have not previously had VA funding ... a different approach would indeed be utilized by any of our current teams of reviews. For example, if an investigator who applies for a ‘first time VA funding award’ has a PhD or MD with postdoctoral experience but has (to be extreme) never published a paper from that experience, why shouldn’t the applicant be deemed less worthy than one who has a proposal of equal scientific merit but from a similarly young investigator who published 10 papers in peer reviewed journals while in training?”

Not considering the productivity of either means that the proposed system runs the risk of promoting the mediocre.”

We agree wholeheartedly with the comments from the field regarding the factors that should be discussed and considered at the peer review committee when addressing productivity. These include:

- the individual’s stage in their career
- the time they devote to research
- the field they pursue
- the quality of the body of work they have produced including its originality and impact on the field (the consideration here should not be solely limited to peer reviewed manuscripts, but consider things such as patents, national and international awards and recognitions, etc.)
- the individual’s grant portfolio (federal grants (VA, non-VA), national foundation grants, regional foundation grants, others)
- time devoted to mentoring junior faculty

Also, we agree with the individual who suggested an “escape hatch” to consider significant life events. We will recommend to reviewers that they consider issues such as time out for child bearing, moving from one academic institution to another, personal illness, or other reasons that might explain a temporary reduction in performance when evaluating the individual’s productivity.

Initially, it was ORD’s plan to exclude junior investigators from a productivity evaluation. Though responses from the field were uncommon, those that did address this issue expressed comments similar to the ones quoted. Having reconsidered this issue, we agree with the field that expectations relative to being a junior investigator should be considered along with the other factors.

Peer Review Plan

The quotes above represent only a sample of the responses we received. However, we believe that these represent an accurate and cogent reflection of the broad opinions of the field. We hope these quotes illustrate a concise snapshot of the sentiments of all VA researchers who responded. Clearly, we could not include all responses.

After taking into account the comments received, I have re-evaluated the process by which the peer review committee should assess productivity (presented below). We agree with the position that these standards should apply to all Services. We also agree that the evaluation of productivity will be done at the time of the peer review committee.

Notably, I have also concluded that ORD should replace the term “Productivity” with “Scientific Contribution” in future documents as a more precise description and title for this index. Thus, it is ORD’s intention to implement the following policy for the peer review committees:

Productivity Scoring

Each reviewer will be asked to evaluate the individual's prior productivity. In evaluating productivity, the peer review committee will be asked to consider the following criteria: the scientific contributions in peer-reviewed literature (as assessed by impact on the field and originality of the work), extent of grant portfolio (e.g., federal grants—VA and non-VA, national and regional foundation grants), mentorship of junior investigators, productivity of the individual and productivity of the research team, evidence of national and international recognition, time devoted to research, stage of career, and patenting of new technologies.

In addition, we agree with the field that it is important to consider one's career holistically. When evaluating productivity, the peer review committee will have the flexibility to consider life events that can cause researchers to have a temporary reduction in productivity including: family leave for child raising, move between academic institutions, illness or other family-related problem, or other unexpected catastrophic events (such as the flood I experienced in Houston VA in 2002). When submitting a proposal, researchers will be encouraged to discuss situations such as these that may have caused a significant hindrance in productivity. The peer review committee will be instructed to consider this information during the productivity assessment. Committee members will then be asked to give a quantitative score from 10-50 that will also correspond to one of the following categories: Excellent, Very Good, Good, Marginal, or Poor.

Scientific Scoring

Next, each member of the peer review committee will be asked to judge the scientific merit of the proposal considering: significance (relevance to VHA mission, scientific contribution); approach (conceptual framework, design issues); innovation (new concept and/or potential to change existing paradigms); and, environment (resources to ensure success). Based on these criteria, a quantitative score from 10-50 will be given that also corresponds to one of the following categories: Excellent, Very Good, Good, Marginal, or Poor.

Overall Score

Lastly, each member will be asked, using the same scale, to give an overall score for the proposal that takes into account both the scientific merit of the proposal and the individual's productivity. At no time will explicit guidelines be given as to what level of productivity should equate to a numerical score, nor will we give guidelines as to how the two scores should be weighted in developing the overall score. A quantitative score will be assigned from 10-50 that also corresponds to one of the following categories: Excellent, Very Good, Good, Marginal, or Poor.

The overall score will be the score that will be utilized in the determination of funding.

