



REPRODUCIBILITY AND RIGOR IN REE'S PORTFOLIO OF RESEARCH

September 20, 2016

Background

The Science Advisory Council was established in FY2016 as a subcommittee of the National Agricultural Research, Extension, Education, and Economics (NAREEE) Advisory Board to provide advice and guidance, on a scientific basis, on the overall strength, practicality, and direction of agricultural research, including emerging technology and scientific issues and report any findings publicly to the NAREEE Advisory Board.

In spring of 2016, the USDA Science Advisory Council was first charged by the Chief Scientist to examine a number of controversial and challenging issues. The first was to address the general subject of reproducibility in the agricultural and nutrition related sciences. General background reading was assigned to members and first review of the issue was conducted in a face to face meeting on May 24th, 2016. This report reflects the understanding and consensus of the advisory council in this matter.

The issue of rigor should actually be appreciated across three specific understandings. This approach was recently discussed in a National Science Foundation report of a year ago (Ref 1) where these understandings were very well outlined. Our goal in guiding USDA on this issue is to help ensure that all studies, extramural and intramural, achieve the goal of being reproducible, replicable and generalizable. Throughout this report we will take the term rigor to include the three concepts of reproducibility, replicability and generalizability.

- **Reproducibility** – The concept of reproducibility is to be able to take the original data and/or materials used in a prior set of experiments and to reproduce the results by conducting a separate but identical approach of analysis. Failure to reproduce the same results could be due to differences in how the data were collected and handled, differences in how the statistical analysis was applied, differences in the actual operations of the statistical tools utilized, and difference in execution of the reproduction that introduced other errors. Each of these variances is less likely to happen if methodology is shared in a clear and transparent manner by the research team.
- **Replicability** – The concept of replicability is to be able to replicate the study by generating a new set of data using the exact same set of methods and the same set of analyses to arrive at the very same set of results and conclusions. To fail at replicability may indicate a flaw in the original experiment. It could also indicate a flaw in the second experiment that is attempting to

reproduce the original results. Alternatively, such a failure could reflect similar results simply because it lacks the sample that may be needed to assure statistical significance between the two studies. Finally, a study may not be replicable because the underlying contention that the conditions/samples were the same for both experiments was false.

- **Generalizability** – The concept of generalizability refers to the ability of a study to be used in a more general context with other populations or contexts that differ significantly from the original study. It reflects the degree to which relationships found in the study apply in different situations and the degree of generalizability can reflect great value of the study. Failure at generalizability usually points to highly limiting conditions that fail to allow the findings to be applicable much beyond the specific conditions of the study.

The pursuit of rigor in our research programs directly reflect on the reputation of all parties -- researcher, sponsor and publisher. Additionally the pursuit of rigor directly affects the citizen's ability to trust the quality and findings of the research.

Report Focus

The Science Advisory Council (the council) has been asked to reflect on the issue of research rigor including reproducibility, replicability and generalizability, in the agriculture and nutrition areas. The council has met by phone and met face to face in discussion regarding the issues of reproducibility. The council drafted a version of this report and then met by conference call to discuss the details of the report. The council primarily addressed the following question: ***What actions can and should USDA take to foster reproducibility and rigor in USDA-supported research?*** To help the council address this we broke the issue into six component parts that are presented as questions to be addressed:

- 1) **General Contributing Factors: Are the typically identified contributory factors to irreproducibility in biomedical and social sciences research also likely contributory factors to irreproducibility in agricultural and nutrition research?**
 - The Science Advisory Council agrees that the issues identified by the biomedical and social science are largely applicable and are also fairly complete for agriculture and nutrition research. An in-depth discussion of these factors can be found in an Academy of Medical Sciences report (Ref 2) outlining issues and possible directions. The council did discuss each in relationship to the world of agricultural and nutrition research. Nutrition researchers present expressed details of where their scientific society and journals were moving positively to address the issue of rigor. Agricultural researchers present were less specific in details regarding actions by their societies and related publishers on this matter.
 - The council discussed the following challenges:
 - **Data Dredging** — This is the perhaps one of the most common challenges in the desire to publish. Data dredging is the practice of repeatedly searching the data for "significant" differences. Alternatively, some use different statistical procedures in the

hunt for "significance". A-priori designed studies with theoretically driven hypotheses and preplanned testing as appropriate should be expected for USDA sponsored research. NIH funded clinical trials are required to be registered and to declare their main hypotheses before starting the trial, and most reputable journals subscribe to this rule, and require a CONSORT format for articles to be published.

- **Omitting Null Results** — This is a common practice among editors and publishers. Sponsors, reviewers and editors need to be more proactive in declaring null results as a valid and important finding so this censoring is less likely to occur.
 - **Under Powered Study** — Statistical power is directly linked to ensuring the amount of data collected is sufficient to identify an actual statistical difference beyond pure luck. Scientists that ensure their studies offer enough power generally seek a power of 0.8, which corresponds to an 80% chance of identifying a real effect.
 - **Errors** — It may seem obvious but occasionally technical error occur in the execution of the study. These errors can fundamentally flaw the results but often go undetected.
 - **Under Specified Methods** — Occasionally, a principle investigator will choose to be obtuse or less detailed in their description of their approach purposefully to "throw off" would be attempts to reproduce the work. This effort must be countered by grant reviewers and publishers, perhaps by counting reproduced results with equal worth to publication citations. Seeking patterns during the earlier development of the research may help with transparency and reduce secrecy.
 - **Weak Experimental Design** — This is where the design chosen by a principle investigator is fundamentally flawed. The flaws could prevent any degree of reliability or validity, and perhaps, could be addressed with either more aggressive pre-award review or post-award oversight of funded research.
 - **Weak Design Execution** — When resources are not used appropriately or according to protocols, and/or inadequate supervision of execution is provided.
- 2) **Unique Contributing Factors: Are there any unique factors that may contribute to a lack of rigor in USDA's broad portfolio of research? If so, how can these unique factors be addressed?**
- For agricultural research, there is a particular high likelihood that under powered studies may be carried out, as the length of production seasons for crops may limit the duration of experiments. This limitation can be addressed by pooling experiments across time, but given the unique production factors (climate, soil) of different parcels, pooling has limits as well. For some social science research, the infrequency of some data sources (Census, yields, and seasonal prices) may also limit the power of studies.

- Human research is expensive, and due to many ethical issues, may lack adequate controls. The interaction of diet, genetics, environment and metabolic factors is complex, and many times difficult to reproduce as time, geography and culture may drastically change these interactions. USDA has funded human nutrition laboratories, and research centers to increase control conditions, however, a strictly controlled situation may make reproduction of results in real-life situations difficult and may limit generalizability of results to society in general. Additionally, by the nature of some nutrition work, identifying the power of a study a priori can be challenging. Observational studies in humans can be hugely expensive and therefore may warrant smaller "directional" studies, however these type so design should be clearly identified as such by principal investigators in their proposal to sponsors.
- Ecosystem research -- some areas of research in the fields of ecology have specific design constrains that should be identified and at least made aware of to all prior to research design and funding. Examples of these limitations include sample size and spatial scale that limit replicability.
- In general, the Chief Scientist should encourage the discussion and use of power analysis across all USDA related fields of research should be encouraged where appropriate and recognize as a cultural change for the fields of study.

3) **Proposed Solution: What solutions for addressing the common contributory factors to lack of rigor and irreproducibility are likely to enhance USDA-support research?**

- OPEN DATA — The federal initiatives launched across all agencies requiring open access to data fits well with this solution. Access to data and metadata of a specific project can allow the rigor and reproducibility of a project to be examined in a post hoc mode. The Chief Scientist should see that USDA fully embraces the open data initiative and clearly instruct principal investigators (intra- & extramural proposals) on how, when and what data must be shared. This should specifically include the expectation that null results must be shared as a valid end point of a sponsored project.
- PRE-REGISTRATION – Pre-registration of results for clinical trials, which include design, hypotheses, collected variables, statistical analysis which defines expected outcomes, will validate post-trial publications. An example of an approach to this is the CONSORT Statement which is minimum set of recommendations for reporting randomized trials. Yet another related example is the Strobe Statement aimed at strengthening the reporting of observational studies in epidemiology. In fields of study where systems of preregistration do not exist, sponsors should consider encouraging publishers to establish such a system.
- COLLABORATION – The Chief Scientist should see that USDA continues its efforts to create, source, and develop cooperative data sharing agreements with researchers for data sets

that would be cost prohibitive to institutions and researchers otherwise (ERS scanner, ARMS and FoodAPS), to facilitate more statistical power in exploring ag and food related social science research. This approach could clearly be applied more broadly to all other areas of study. The Chief Scientist should consider encouraging USDA-related discipline bodies such as scientific societies to open up a dialog supporting expanded collaboration and establish mechanisms to enable cross-disciplinary collaborations.

- **AUTOMATION** – Laboratory Management Information Systems that use computer programs for automation of data collection, management and control of data organization and efficient record keeping help manage data and reduce inadvertent loss or induction of errors.
 - **OPEN METHODS** – Sharing of methods within disciplines helps to establish best practices and ensure rigor in experimental design. The Chief Scientist should see that USDA encourages and facilitates this approach by working with publishers and scientific societies. Standard reference materials and methods can be shared while at the same time allowing room for new methods to be explored.
 - **POST PUBLICATION** – In the past, hard-copy publications limited the information available due to space. Currently, the majority of the journals offer electronic publications of additional information and these should be encouraged and fully taken advantage of by principal investigators. The Chief Scientist should encourage publishers and scientific societies to address this need to consider more formally moving into this post publication space.
 - **REPORTING GUIDELINES** – The Chief Scientist should see that USDA includes guidelines for internal reports for monitoring and supervision of the research, and ensure that specific guidelines for external periodic (quarterly or annual reports to the funding agency) are developed. In addition, other regulatory reports such as those for the Institutional Review Boards should be able to be considered as related materials in a post publication presentation of the work.
- 4) **Should research funding requests submitted to NIFA and ARS be required to address rigor and reproducibility similar to requests submitted to the NIH?**
- Yes, and post-award management of researchers should be just as fully invested in addressing rigor and reproducibility as the pre-award process that facilitates the selection of the strongest research projects.

- Request for funding should also allow for funding to facilitate early sharing and publication of results, such as funding to attend meetings for sharing of databases, and expenses for making the results available in the public domain.
- We recommend that the Chief Scientist calls upon NIFA and ARS to require funding requests to address rigor and reproducibility in their calls for proposals.

5) **What current actions, practices and/or resources, if any, are being taken by USDA intramural research agencies (e.g., ARS) to foster rigor and reproducibility in research?**

- A number of policies, handbooks and documents are used by USDA to help guide researchers to support rigor and reproducibility in research.
 - USDA INTEGRITY POLICY --This policy helps guide decision makers as public policies are created by ensuring that they are informed by sound science relevant to food, agriculture, natural resources, rural development, and related issues.
 - Policy: https://www.ocio.usda.gov/sites/default/files/docs/2012/DR%201074-001_0.pdf
 - USDA SCIENTIFIC INTEGRITY POLICY HANDBOOK – This Scientific Integrity Policy (SIP) Handbook identifies the procedures USDA used to implement the Department’s Scientific Integrity Policy.
 - Handbook: <http://www.usda.gov/documents/usda-scientific-integrity-policy-handbook.pdf>
 - ARS ETHICS DOCUMENT -- This document present a series of policies and procedures that focus on research misconduct and provide instruction and guidance to employees on the methods and principles for reviewing allegations of research misconduct.
 - ARS Ethics document <https://www.afm.ars.usda.gov/ppweb/pdf/129-0-ars.pdf>
 - HUMAN SUBJECTS POLICY -- The Human Subjects Policy guides researchers when they are involving humans as subjects in their research.
 - P&P 605.1, Protection of Human Subjects <https://www.afm.ars.usda.gov/media/10444/pp605-1.pdf>
- We recommend that the Chief Scientist take the complied recommendations and discussion on this matter as a basis for new policies that support a commitment to rigor and reproducibility while advancing the quality and quantity of science produced by USDA.

6) **What actions, if any, need to be taken by the Chief Scientist to address issues raised in this report and yet to be addressed?**

- The Chief Scientist should convene meetings of publishers associated with the agriculture, food and fiber related research outlets. This meeting should discuss [1] the need for publication of results, including null results, from NIFA funded projects, [2] the need for standardized procedures and reporting guidelines where possible for the science presented and [3] the need for a post publication review/discussion platform for public comment on publications. USDA may negotiate agreements with the journals on the use of information from registered trials and obligations for the public release of information.
- The Chief Scientist should convene meetings of scientific societies for the purpose of discussing the need for platforms to allow for [1] open data access, [2] open methods that detail study protocols and [3] reporting guidelines for certain fields of study.
- The Chief Scientist should convene a set of open forums to encourage the interaction with potential principal investigators for the purpose of discussing rigor and reproducibility. These forums could lead to new guidelines regarding expectation of grant proposals in dealing with rigor and reproducibility issues.
- The Chief Scientist should see that USDA guides principle investigators in proposal writing:
 1. to ensure data dredging is recognized as an unacceptable practice.
 2. to establish appropriate statistical power of their proposed work to ensure sample size is adequate for the proposed work.
 3. to be fully transparent in the specified methods being proposed and that adequate experimental design be fully disclosed in the proposal.
 4. to explicitly include statisticians where appropriate as co-PIs with a full say in the statistical components of the study.

References:

1. Bollen, K., et. al., Social, Behavioral, and Economic Sciences Perspectives on Robust and Reliable Science -- Report of the Subcommittee on Replicability in Science Advisory Committee to the National Science Foundation Directorate for Social, Behavioral, and Economic Sciences, May 15, 2015
2. Reproducibility and reliability of biomedical research: improving research practice – Symposium Report, Academy of Medical Sciences, October, 2015
<http://www.acmedsci.ac.uk/researchreproducibility>
3. Consort Statement: <http://www.consort-statement.org>
4. Strobe Statement: <http://www.strobe-statement.org/index.php?id=strobe-home>

Report Developed by the Science Advisory Council, a Subcommittee of the NAREEE Advisory Board

Dr. Mark McLellan (Chair), Vice President of Research and Dean of the School of Graduate Studies, Utah State University (representing National Food Science Organization)

Dr. Patsy Brannon, Professor and Nutritionist, Cornell University (representing National Human Health Association)

Dr. Adriana Campa, Associate Professor of Nutrition, Florida International University (representing National Nutritional Science Society)

Dr. Steven Daley-Laursen, Senior Executive, Office of the Vice President of Research and Economic Development and Professor of Forest Ecology and Bio-geosciences, College of Natural Resources, University of Idaho (representing National Forestry Group)

Dr. Govind Kannan, Dean and Director, College of Agriculture, Family Sciences, and Technology, Fort Valley State University (representing National Food Animal Science Society)

Dr. Neil Olson, Dean, College of Veterinary Medicine, University of Missouri (representing American Colleges of Veterinary Medicine)

Dr. Robert Taylor, Dean and Director of Land Grant Programs, Florida A&M University (representing National Crop, Soil, Agronomy, Horticulture, or Weed Science Society)

Dr. Dawn Thilmany, Professor, Department of Agriculture and Resource Economics, Colorado State University (representing National Social Science Associations)