United States Senate
COMMITTEE ON ENVIRONMENT AND PUBLIC WORKS
WASHINGTON, DC 20510–6175

March 17, 2014

Dr. Francesca Grifo
Science Integrity Official
U.S. Environmental Protection Agency
William Jefferson Clinton Building
1200 Pennsylvania Avenue NW (Mail Code: 8105R)
Washington, DC 20460

Dear Dr. Grifo:

Last summer, the U.S. Environmental Protection Agency (EPA) entered into an agreement with Republican Senators from the Environment and Public Works (EPW) Committee of the United States Senate to obtain all research data from certain epidemiological studies that were supported by funding from EPA. The data we requested stem from two key epidemiological studies on the health effects of long-term exposure to air pollution: "An Association between Air Pollution and Mortality in Six U.S. Cities," published in the *New England Journal of Medicine;* and "Particulate Air Pollution as a Predictor of Mortality in a Prospective Study of U.S. Adults," which was published in the *American Journal of Respiratory and Critical Care Medicine*.[1] Both were the subject of follow-up studies,[2] which were used in the benefits analyses of numerous significant air regulations proposed by EPA.

Both the United States Senate and House of Representatives have attempted to work with EPA since 1997 to acquire the underlying data from the researchers. Gina McCarthy, Administrator of EPA, committed to EPW Republicans in 2013 to deliver certain of this data, which essentially provides the backbone for the vast majority of air regulations promulgated by her Agency. A number of excuses, by both EPA and the research institutions, have been proffered for not acquiring or otherwise releasing the underlying data for independent reanalysis. Last fall, as part of a larger attempt to gather the data, EPA transmitted EPW's request for study protocols, questionnaires, software, and other non-confidential information to the four lead researchers and mediated our request for the raw data from the associated institutions (*i.e.*,

[1] Dockery, D.W., C.A. Pope, X. Xu, J.D. Spengler, J.H. Ware, M.E. Fay, B.G. Ferris, Jr., and F.E. Speizer. 1993. "An Association between Air Pollution and Mortality in Six U.S. Cities." *New England Journal of Medicine* 329: 1753-1759, and Pope, C.A., M.J. Thun, M.M. Namboodiri, D.W. Dockery, J.S. Evans, F.E. Speizer, and C.W. Heath, Jr. 1995. "Particulate Air Pollution as a Predictor of Mortality in a Prospective Study of U.S. Adults." *American Journal of Respiratory and Critical Care Medicine* 151:669-674.

[2] Laden, F., J. Schwartz, F.E. Speizer, and D.W. Dockery. 2006. "Reduction in Fine Particulate Air Pollution and Mortality." *American Journal of Respiratory and Critical Care Medicine* 173:667-672, and Pope, C.A. III, R.T. Burnett, M.J. Thun, E.E. Calle, D. Krewski, K. Ito, and G.D. Thurston. 2002. "Lung Cancer, Cardiopulmonary Mortality, and Long-term Exposure to Fine Particulate Air Pollution." *Journal of the American Medical Association* 287:1132-1141.

Harvard University, American Cancer Society, and Health Effects Institute). We received an incomplete data set from Krewski *et al.* (2009) and Lepeule *et al.* (2012).[3]

Congress employed the help of the Texas Commission on Environmental Quality and the National Institute of Statistical Sciences to independently examine the data received, and both separately verified that we were missing critical relative information necessary for reanalysis.[4] More importantly, information provided by the relevant institutions tends to indicate that data may have been corrupted, destroyed, or otherwise intentionally made unavailable for independent reanalysis.

EPA's Scientific Integrity Policy, which you are required to ensure all aspects are upheld by the Agency, describes the Agency's scientific work as that of being of the highest quality.[5] However, relying on data that potentially no longer exists in complete form, or is otherwise of diminished value, actively negates any claim EPA has on its commitment to relying on the highest quality science. To retain a standard of probity in not only the scientific community but in the eye of the American public, the Scientific Integrity Policy describes ways to strengthen "the actual and perceived credibility of Agency science by… ensuring that scientific studies used to support regulatory and other policy decisions undergo appropriate levels of independent peer review."[6] The policy also directs EPA "to expand and promote access to scientific information by making it available online in open formats in a timely manner, including access to data and non-proprietary models underlying Agency policy decisions."[7] Considering the current state of EPA's relationship with Congress regarding the availability of this data, there are clear points of divergence that need to immediately be addressed.

Trust is not blind in the scientific community; reproducibility and consistency are hallmarks of the scientific method. EPA has circulated memos and established Agency policy guidelines which reaffirm that it "recognize[s] that influential scientific, financial, or statistical information should be subject to a higher degree of quality (for example, transparency about data and methods) than information that may not have a clear and substantial impact on important public policies or private sector decisions."[8] While such affirmations of commitment to transparency and holding influential information to higher degrees of quality are good first steps, circumstances indicate they have not been upheld. The April 8, 1995, EPA memo on "Guidelines for Study Rejection Based on GLP Considerations" discusses good laboratory practice (GLP)

---

[3] Krewski, D., M. Jerrett, R.T. Burnett, R. Ma, E. Hughes, Y. Shi, M.C. Turner, C.A. Pope III, G. Thurston, E.E. Calle, M.J.Thun. 2009. "Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality" *Health Effects Institute (HEI) Research Report*, 140, and Lepeule, J., F. Laden, D. Dockery, and J. Schwartz. 2012. "Chronic Exposure to Fine Particles and Mortality: An Extended Follow-up of the Harvard Six Cities Study from 1974 to 2009." *Environmental Health Perspectives* 120 (7):965-970.

[4] *See* Letter from Dr. Michael Honeycutt to Congressman Andy Harris (November 9, 2013) *available at* http://science.house.gov/sites/republicans.science.house.gov/files/documents/TCEQ%20Letter%20to%20Harris_0.pdf; *See also* Letter from Dr. Stan Young to Senator David Vitter (March 7, 2014).

[5] EPA. "Science Integrity Policy." http://www.epa.gov/osa/pdfs/epa_scientific_integrity_policy_20120115.pdf at p.3

[6] Ibid. at p. 4

[7] Ibid. at p. 4

[8] EPA. 2002. "Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity, of Information Disseminated by the Environmental Protection Agency." EPA/260R-02-008. http://www.epa.gov/quality/informationguidelines/documents/EPA_InfoQualityGuidelines.pdf at p. 20

guideline studies and their review by Agency scientists. Of particular note is Item 8: "Failure to record and archive critical raw data, including study specimens/tissues; inability to reconstruct a study due to the absence of raw data that the Agency has required and/or relied upon to make a regulatory decision [40 CFR 160.130; 40 CFR 160.195]"[9] This point illustrates EPA's acknowledgement that having raw data is of the utmost importance, particularly when the raw data is part of a significant EPA regulation.

While EPA recognizes how important it is for industry to provide the Agency with raw data, EPA clearly feels they are above such a requirement: EPA is promoting a double standard in which EPA requires industry to provide raw data, but EPA can establish regulations on industry without making the raw data available to the public? It is with great irony that EPA would require industry to submit "raw" data whereas EPA has been issuing regulations that are based on data that nobody can reconstruct and evaluate. EPA's willful reticence is akin to somebody being criminally prosecuted based on data that the defense is not allowed to see.

Without raw data, it is nearly impossible to critically evaluate the quality, integrity, and relevance of the study.[10] The soundness, effectiveness, and credibility of EPA's regulations ultimately rest on the scientific and technical bases for these actions. Careful attention to research record keeping can help ensure data quality and integrity.[11] The same guidelines for ensuring information quality goes on to state, "It is important that analytic results for influential information have a higher degree of transparency regarding (1) the source of the data used, (2) the various assumptions employed, (3) the analytic methods applied, and (4) the statistical procedures employed."[12]

Any scientific misconduct in research damages the trust of citizens in science and in government. I believe EPA is guilty of data-related misconduct, described by the Organization for Economic Co-operation and Development (OECD) as:[13]

- Not preserving primary data
- Bad data management, storage
- Withholding data from the scientific community

As stated in OECD's "Best Practices for Ensuring Scientific Integrity and Preventing Misconduct," "Misconduct is a special concern for governmental administrators, who are the

---

[9] EPA. 1995. Memorandum. Subject: Guidelines for Study Rejection Based on GLP Considerations; From: Dan Barolo, Director, Office of Pesticide Programs; To: All Division Directors, Office of Pesticide Programs. http://www.epa.gov/compliance/resources/publications/monitoring/fifra/glp/glpstudyrejection.pdf

[10] See, *e.g.*, Haseman, J.K., Bailer, A.J., Kodell, R.L., Morris, R., and Portier, K. 2001. "Statistical Issues in the Analysis of Low-Dose Endocrine Disruptor Data." *Toxicological Sciences* 61: 201-210, at p. 204 ("In two studies, the raw data had significant errors…These errors were detected (and corrected) only because the Statistics Subpanel had access to summary data in published papers for comparative purposes.")) http://toxsci.oxfordjournals.org/content/61/2/201.full.pdf+html

[11] Birnbaum, L.S. and B.T. Culpepper. 1999. "Research Integrity: A Government Perspective." *Quality Assurance* Oct-Dec; 7(4):217-24. http://www.ncbi.nlm.nih.gov/pubmed/11191122

[12] EPA. 2002. *supra* note 8, at p. 21

[13] Organization for Economic Co-Operation and Development Global Science Forum, "Best Practices for Ensuring Scientific Integrity and Preventing Misconduct." at p.1 *Available at* http://www.oecd.org/sti/sci-tech/40188303.pdf

primary constituency of the OECD Global Science Forum. On behalf of the public, and to achieve societal benefits, they fund, oversee, and evaluate research, much of which is conducted directly in public institutions or is otherwise sponsored by governments."[14] My concern is that EPA, as well as the relative research institutions, may be committing data-related misconduct under the standards set forth by OECD.

Historically, EPA has worked with the lead researchers to stonewall Congressional attempts to gain access to the above referenced data, countering all requests with variations of the same reasons why it refuses to release the data: 1) EPA neither possesses nor own all of the data; 2) EPA and the research institutions are concerned that releasing data would compromise the privacy of study participants. In response to the first point, with enactment of the Shelby amendment[15] and subsequent OMB guidance,[16] the federal government retains the right to "obtain, reproduce, publish, or otherwise use the data" produced from a federal grant and to "authorize others to receive, reproduce, publish, or otherwise use" such data for federal purposes.[17] Congress has a constitutionally-based right of access to information from the executive branch, which EPA continuously disregards as it comes up with excuse after excuse for why it is justified in withholding, or otherwise not attempting to acquire, important research.

EPA previously admitted to Congress that the limited universe of data they shared, which they received from the outside researchers, is insufficient for reanalysis. As the input and output files are fundamental to conducting reanalysis, I repeatedly requested that EPA: (1) obtain all the data files; (2) determine which data files pose a threat to privacy; (3) immediately release all data files that do not pose a threat to privacy; and (4) investigate measures to remove all personal health information from the files that contain confidential data prior to release. The fourth request should have already been completed, as it is hardly a novel undertaking in the scientific field. The U.S. Department of Health and Human Services recently issued guidelines on how to de-identify medical records in order to implement elements of the new healthcare law.[18] Additionally, EPA itself worked with the Centers for Disease Control and Prevention (CDC) to remove personal identifiers from data provided by Harvard University and released information on deaths originally obtained from the National Death Index (NDI), providing evidence that data containing personal information can be de-identified and released.

Finally, there is reason to believe that EPA has manipulated information so that research is not accurately represented in the record. As OECD notes, this is of "special concern for

---

[14] Ibid.

[15] H.Rept. 105-825. The provision was a rider attached to the Treasury and Postal section of the Omnibus Consolidated and Emergency Supplemental Appropriations Act for FY1999. It required that OMB amend section 36 (c) [intangible property] of Circular A-110, "Uniform Administrative Requirements for Grants and Agreements with Institutions of Higher Education, Hospitals, and Other Non-Profit Organizations" (2 C.F.R. 215). Its principal sponsors were Senator Richard C. Shelby and Representative Robert B. Aderholt. The provision is sometimes called the Shelby or Shelby-Aderholt Amendment. It has also been called the Data Access Act.

[16] Office of Management and Budget, "Circular A-110: Uniform Administrative Requirements for Grants and Agreements with Institutions of Higher Education, Hospitals, and Other Non-Profit Organizations," 2 C.F.R. 215 (May 11, 2004).

[17] 2 C.F.R. §§ 215.36 and 215.53.

[18] U.S. Department of Health and Human Services, "Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule" http://www.hhs.gov/ocr/privacy/hipaa/understanding/coveredentities/De-identification/guidance.html

governmental administrators."[19] OECD clearly points out that "misconduct in research damages science"[20] and "society may be harmed if false results become widely known and believed."[21] As EPA, as well as the accommodating institutions, has failed for nearly two decades to be transparent, I am concerned that OECD's apprehensions with such activities are playing out within a critical U.S. Agency. OECD states, "Damage to science through the undermining of the public's trust in science, and of the government's ability to foster and promote research in a competent and responsible manner"[22] can lead to a decline in the credibility of scientific analysis and "advice on issues that have important implications for society."[23]

The matter of data-related misconduct can be of significant consequence to our citizenry. A great many regulatory actions are being taken with impacts in the hundreds-of-billions of dollars. The broader socioeconomic impacts are no less consequential. Accordingly, I ask for your specific guidance on the appropriate steps that need to be taken to ensure guidelines and policy on data misconduct and scientific transparency are not violated by EPA. It is critical that our governmental institutions retain the highest standards of scientific integrity. Instruction from you on how EPA is working to resolve matters of "secret science" regarding the data requested and information that are being withheld from the public, to ensure transparency in the process, is appreciated. I look forward to receiving your response by no later than April 17, 2014.

Thank you,

David Vitter
Ranking Member
Committee on Environment and Public Works

Cc: Douglas W. Dockery, Professor of Environmental Epidemiology, Department Chair, Harvard School of Public Health, 665 Huntington Ave., Boston, M.A. 02115-6021
American Cancer Society
Otis W. Brawley, MD, FACP, Chief Medical Officer and EVP, Research and Cancer Control Service, American Cancer Society, 250 Williams Street, Atlanta, GA 30303-1002

[19] OECD *supra* note 13, at p. 1.
[20] Ibid.
[21] Ibid at p. 4.
[22] Ibid.
[23] Ibid.