ENSURING OPEN SCIENCE AT EPA

HEARING BEFORE

THE SUBCOMMITTEE ON ENVIRONMENT

COMMITTEE ON SCIENCE, SPACE AND TECHNOLOGY

US HOUSE OF REPRESENTATIVES

TESTIMONY

OF

ELLEN KOVNER SILBERGELD, PhD

PROFESSOR

JOHNS HOPKINS BLOOMERG SCHOOL OF PUBLIC HEALTH

February 11, 2014

I am Ellen K Silbergeld, Professor of Environmental Health Sciences and Epidemiology at the Johns Hopkins Bloomberg School of Public Health. I am appearing at your invitation to testify on the issue of information disclosure and on the discussion draft entitled the "Secret Science Reform Act of 2014." The views and opinions presented here are my own, and do not represent the views and opinions to the Johns Hopkins University. By way of background and experience, I have conducted research related to environmental health for over 40 years. I have also served on numerous expert panels, advisory boards, and as a consultant to the State of Maryland, National Research Council, EPA, DoE, CDC, FDA, NIH, NSF, WHO, ILO, UNIDO, FAO, and UNEP. Thus, I am familiar with the processes by which regulatory and scientific agencies identify and evaluate scientific information as part of the process of regulation. I was a member of several expert groups convened by EPA and the NRC considering the health impacts of lead in the environment, during which a re-evaluation of research data was undertaken. I have submitted my resume to the Committee in advance of this hearing.

First, I want to state that the principles of openness and fairness are fundamental to science, including toxicology, epidemiology, and basic research. I have been a leader in the international movement towards adopting the principles of evidence based decision making in fields beyond clinical medicine and health care. I strongly support access to and sharing of scientific findings within the community of stakeholders in a manner consistent with principles of fairness and adherence to the goal of improving the process of decision making. These principles are described in my paper in ALTEX (attached).

I agree with the statement of Chairman Smith that at present there is an important need to reduce the secrecy that confounds public access to the basis for some EPA decisions. In my experience, the major driver of secrecy in EPA rulemaking is the deference given to industry in terms of shielding its studies from public view. For this reason, I am puzzled as to the uneven nature of the debate on this topic, which we discussed in our commentary published in EHP (attached). The proposed bill would continue to immunize industry from disclosure while increasing the burden on EPA and, by pass through, on non-industry researchers. As noted in an earlier statement by Chairman Lamar Smith (November 2013), the interest of the public in the right to

see data is such high importance that the clouds of secrecy should be dispelled whatever the source.

The problem of nondisclosure by industry was a key issue in the initiation of the High Production Volume Chemicals Challenge Program by the OECD (Organization for Economic Cooperation and Development). I was a member of the US delegation to the OECD Environment Program during the development of this voluntary process, which was initiated following a study by the Environmental Defense Fund (**Toxic Ignorance** http://www.edf.org/sites/default/files/243_toxicignorance_0.pdf of which I was a coauthor). I was proud of the leadership role of American industry in the success of this program through participation in a tripartite partnership among government, industry and NGOs, to overcome the lack of basic toxicity data on most chemicals in commerce and consumer products. The HPV program has revealed that in many cases the critical data had already been generated but not released by industry. As stated by the American Chemistry Council on its website:

> Under the HPV Challenge Program, hundreds of chemical makers volunteered health and environmental information on 2,200 chemical products, representing approximately 95 percent of the commercial market by volume in the United States, to help create a database that is available to the public.
>
> This voluntary initiative demonstrates that collaboration between public and private sectors can be an efficient method of developing safety information to help ensure the safety of the products of chemistry.

With respect, this proposed legislation constitutes a retreat from this highly responsible and effective policy of information disclosure accepted and led by US industry.

We need more information and more information disclosure by industry. Like trees falling unheard in the forest, information withheld is not informative. How much better would West Virginia have been able to respond last month if industry data were available and released on 4-methylcyclohexane methanol (MCHM) instead of the empty Material Safety Data Sheet:

Section III.	Hazards Identification
Acute Health Effects	No specific information is available in our data base regarding the toxic effects of this material for humans. However, exposure to any chemical should be kept to a minimum. Skin and eye contact may result in irritation. May be harmful if inhaled or ingested. Always follow safe industrial hygiene practices and wear proper protective equipment when handling this compound.
Chronic Health Effects	**CARCINOGENIC EFFECTS** : Not available. **MUTAGENIC EFFECTS** : Not available. **TERATOGENIC EFFECTS** : Not available. **DEVELOPMENTAL TOXICITY**Not available. Repeated or prolonged exposure to this compound is not known to aggravate existing medical conditions.

As a scientist. I conclude that the broad sweep of stipulations in the draft bill is without a strong basis in terms of improving science or expanding the evidence base for decision making. I am also the editor in chief of a major peer reviewed journal (**Environmental Research**) and in that role over the past 18 years I have considerable experience in and respect for the process of peer review as a method of quality assessment. The peer review process requires the inclusion of scientific and technical information including, as stated in the bill "materials, data, and associated protocols necessary to understand, assess, and extend conclusions." The rest of the items in Section 2(b) (3) do not contribute to this goal, in my opinion. In science, we recognize that no study is perfect. That is why science has relied on replication as a means of validating the findings and conclusions of a particular study. "Replication" is not to be confused with data re-analysis; replication involves the design and conduct of a wholly independent study (sometimes with different methods) to test the same hypothesis. These are critical criteria for evidence in the standard methods of the Cochrane Collaboration.

Let me also reflect on my experience of data re-analysis as part of the EPA's process of reviewing the science related to associated lead as a risk for children's neurobehavioral development relevant to the Clean Air Act. That re-analysis was demanded by industry and it was accomplished in a non-adversarial way through third party review undertaken by an acknowledged expert in biostatistics not connected with government, industry, or the original investigators. This review elicited some recommendations in terms of restating certain results but the main weight of the study was affirmed. And, of course, since that time, hundreds of independent studies have confirmed and extended the findings of that first publication.

In conclusion, I restate my philosophical support for increasing the transparency of information associated with government regulation. I suggest that we already have the tools to accomplish this goal, in an even handed manner, through the methods of systematic review for evidence based decision making. I hope that your concerns can be reframed to apply to all sources of

information in an effective and efficient manner. Given past history of contended regulations, as a scientist, an editor, and a citizen I am not convinced that the extraordinary and frankly arbitrary measures called for in this legislation will accomplish these goals. Because I know that some of my colleagues in industry have been vocal in calling for these steps, I would challenge them to tear down every wall, in the words of Ronald Reagan, that hides critically important information generated and held by industry.

I am prepared to respond to your questions to the best of my knowledge.

Evidence-Based Toxicology: Strait is the Gate, But the Road is Worth Taking

Ellen Silbergeld and Roberta W. Scherer

Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

Summary

The concept of evidence-based toxicology (EBT) was proposed in 2006, but progress since that time has been impeded by differing definitions and goals. This paper describes the parallels and discontinuities between the approach and methods of evidence-based medicine and health care and those proposed for toxicology. The critical element of an evidence-based approach for either discipline is the adoption of unbiased, transparent methodologies during the collection, appraisal, and pooling of evidence. This approach, implemented during the conduct of a systematic review, allows evaluation of the breadth and quality of available evidence. At present, systematic reviews are rarely done in toxicology by regulatory agencies, international organizations, or academic scientists. Adopting an EBT approach will necessitate significant changes in practice as well as attention to distinctive characteristics of toxicological studies, notably their emphasis on identifying harms and their reliance on experimental animal studies. An evidence-based approach does not obviate the role of judgment and values in decision making; its goal is to ensure provision of all available information in a transparent and unbiased manner.

Keywords: evidence-based toxicology, evidence-based medicine/health care, systematic reviews

1 Introduction

The concept of evidence-based toxicology (EBT) has been under discussion for several years (Hoffmann and Hartung, 2006). EBT is about assembling the evidence related to hazards and risks of exposure, or to the evaluation of methodologies for assessing toxicology for the purpose of using this systematically collected evidence during decision making. In this way it is similar to Evidence-based Medicine and Health Care (EBM/HC), which uses evidence derived from randomized controlled trials on which to base healthcare decisions. EBM/HC is defined as the application of *systematically acquired evidence* within the experience and expertise of the clinician, as well as patient values (Sackett et al., 1996). The essential premise is that decisions should be based on the evidence. It is important that the evidence be obtained in a transparent and systematic manner that is clearly described, enabling other investigators to obtain the same evidence. Like EBM, the impetus for EBT clearly is related to the increasingly important role of the discipline of toxicology in decision making related to public health as well as clinical and preclinical sciences. Progress in EBT has been impeded by differing definitions (Guzelian et al., 2005; Griesinger et al., 2009), both of which advocate the use of methods developed for assessing and using evidence from randomized controlled trials for EBM, an approach that is not feasible for the study of agents suspected of toxicity, as we will discuss below. Efforts also were impeded by a relatively limited focus on the application of evidence-based approaches to the validation and acceptance of alternative methods in applied toxicology (Hartung, 2010).

Evidence-based decision making can be defined as the translation of information into accepted practice using methods that reduce bias and increase confidence (Grimshaw et al., 2006). As in the law, evidence-based methods involve the evaluation of information for its admission into consideration in decision making through the process of applying specified norms and methods. In order to avoid bias, these norms and methods must stand apart from the information under consideration, and their application must be undertaken with complete transparency.

These characteristics differentiate evidence-based approaches from current approaches used in the translation of toxicological studies into decision making by agencies concerned with

Received September 10, 2012; accepted in revised form October 31, 2012.

This paper is an elaboration of presentations by Roberta Scherer and Ellen Silbergeld at the workshop on "Evidence-based Toxicology for the 21st Century: Opportunities and Challenges," held on January 24-25, 2012 at the Environmental Protection Agency campus in Research Triangle Park, NC, USA (see http://www.ebtox.com). See Judson et al. (2013), for a written version of another major presentation from the workshop and Stephens et al. (2013), for an overall summary of the workshop.

occupational and environmental health and consumer protection, as we will demonstrate in this paper. In present practice, the identification of relevant primary studies and the norms by which these studies are evaluated in toxicology are largely implicit (the so-called Delphi method). As a result, the process clearly is not transparent and, because of this, it is difficult to avoid or reduce controversies over policy decisions incorporating toxicology. A previous paper commented on the opacity of the Delphi method often used in risk assessment (Silbergeld, 2009), in terms similar to critiques of medical decision making using these methods (Flower et al., 2007).

There is an understandable skepticism on the part of practitioners and experts in a field to the suggestion that the adoption of major changes in practice may be advantageous. This skepticism was expressed in the early days of EBM (Feinstein, 1995; Williams and Garner, 2002; Chalmers, 2005). We acknowledge and respect this natural skepticism in toxicology. This paper makes the case that adoption of evidence-based methods in toxicology may benefit from awareness of the history of evidence-based approaches in medicine and health care (EBM/HC). The goal of this paper is to introduce a consistent vocabulary for EBT and to examine the extent to which our experience in EBM /HC can inform the development of EBT.

At the outset, we recognize that it is reasonable to ask if adopting EBT will increase efficiency and quality of decision making. The history of EBM/HC demonstrates that the evidence-based approach has accomplished these goals in medicine and many health care-related fields (Dickersin and Manheimer, 1998). Moreover, this history shows that a commitment to an evidence-based approach in these fields has stimulated expansion and improvement in the field, specifically through the development of systematic reviews as the instrument for translating information into evidence. Systematic reviews often are considered the highest source of evidence in that primary studies are systematically identified and appraised and the totality of evidence is synthesized. This did not occur without considerable effort. When systematic reviews were initially conducted in medicine in the early ‘80s, many authors noted that methods associated with conducting systematic reviews were wanting in several areas, including reporting the primary studies, methods for identification and appraisal of the data, and methods for statistical pooling of the data (Mulrow, 1987; Oxman and Guyatt, 1988). The need to develop these approaches was not accepted readily by all practitioners (Chalmers, 2005). Nevertheless, over time, standards were developed through consensus for reporting primary studies (e.g., the CONSORT statement and its extensions[1]), for reproducibly searching for these studies (Dickersin et al., 1994), and standardized methods to identify and account for biases in the primary studies (Moher et al., 1996). Also over time, further statistical methods and inferential models were put forward to synthesize similar research efforts. This focus on methods used during the conduct of a systematic review process, in turn, has led both to greater transparency in reporting primary studies and to an increased focus on the quality of the studies comprising the evidence.

Also of interest to the field of toxicology, the focus on study quality in EBM/HC, in turn, has influenced researchers in relevant fields to improve the quality of their research designs and the rigor of their statistical analyses in order to meet the criteria for inclusion in systematic reviews as well as to support evidence-based strategies. From the perspective of the development of toxicological sciences, this may be one of the most important benefits to consider in adopting EBT.

There is concern that an evidence-based approach introduces rigidity into decision making (Gatchel and McGeary, 2002) and through this may exclude valuable information through the use of scoring systems and meta-analysis. In answer to these concerns, it should be noted that in EBM/HC the evidence provided by transparent systematic reviews provides only one stage of the evidence-based process of application of the evidence. This is not dissimilar to the role of toxicology in decision making as part of the overall process of risk management (NRC, 1994.) Any decisions made in EBM/HC or toxicology must include consideration of other factors, such as cost, feasibility, and the bounds of accepted practice. Thus, in medicine, application of systematically acquired evidence is done taking into account the needs and values of the individual seeking health care (Sackett et al., 1996). Moreover, there is no requirement for evidence-based decision making to employ formal meta-analysis or to use forest plots to express integrated findings.[2] The use of systematic tools, when appropriate, is an important means of ensuring reproducibility of analysis, as well as the quality of the review, by ensuring comparability in design and conduct across the individual data sources, and, above all, enhanced transparency of conclusions reached in the systematic review.

We argue that toxicologists should consider key lessons learned over the evolution of EBM/HC. First, such transitions are best managed by the community of researchers and practitioners, rather than by imposition from outsiders (such as regulators and other consumers of toxicological evidence). Second, as demonstrated in current practice in EBM/HC, evidence-based methods do not reduce or replace the importance of expert and experienced judgment. Rather, they simply provide the totality of evidence upon which to base those decisions. Third, the process in itself does not generate decisions. Simply put, an evidence-based approach assists the community by providing systematically collected information using clearly described methods that reliably represent the state of relevant knowledge. Thus, this approach assists decision makers in increasing the acceptability of their decisions by ensuring transparency during evidence collection. Fourth, a systematic and transparent approach to collecting and appraising the available evidence in EBM/HC has had a positive influence on researchers in terms of study design and data analysis.

[1] http://www.consort-statement.org/

[2] http://www.cochrane.org/resources/handbook/index.htm

2 Toxicology is not medicine or health care

Despite the relevance of understanding the history and experience in EBM/HC, there are characteristics of toxicology and its applications in public health that require more than simple adoption of EBM/HC methodologies. Some of these are related to differences in fundamental objectives. EBM/HC focuses primarily on developing evidence of the efficacy of therapy, together with an emerging focus on the accuracy of diagnostic tests, as well as some focus on etiology, prognosis, and screening. In contrast, the main focus of toxicology is on developing evidence for harms (hazard) and the magnitude or likelihood of harms (risk). Although questions of harm have occasionally been the subject of EBM systematic reviews, as discussed below, many study designs utilized in generating evidence in EBM are not specifically intended to detect or characterize harms. Second, EBM/HC draws almost exclusively upon studies conducted in humans and human populations; toxicology draws primarily upon studies conducted in nonhuman animals and nonanimal models in order to achieve its societal goals of preventing disease and disability. Thus it is important to recognize that adoption of evidence-based approaches for toxicology will require considerable work by the community, as discussed below.

3 Assessing current practice in toxicology

To date, there have been relatively few explorations of the application of evidence-based practices to resolving issues of importance in toxicology. Toxicology has matured in the context of increased demands for its information through the growth of public concerns and regulation in environmental and occupational health. The structure of information needs for decision making in these domains of public health is relatively well defined to include understanding the elements of relevant toxicological studies and the major decision rules into which these elements are to be incorporated. For the purposes of this paper, we focus on those toxicological studies related to defining hazard and quantifying risk; exposure assessment, which is the other element of risk-based decision making, involves other disciplines and methodologies. Hazard and risk are common to the practice of risk assessment and to application of the precautionary principle, which has been advanced as a partial alternative to risk assessment based methods related primarily to reducing the burden of information required for undertaking assessments (Silbergeld et al., 2004).

Current evaluations of toxicological information (from human and nonhuman subjects), almost without exception, have failed to utilize systematic or transparent methods. These limitations are exemplified by a review on lead and cancer by one author of this paper (Silbergeld, 2003) and a review of the carcinogenicity of lead compounds by the International Agency for Research on Cancer (IARC, 2006). Both of these examples are distinguished by lack of transparency such that it is not possible to determine or to replicate the process of identifying studies or their selection for review. No information was provided on the search strategy or on screening criteria in terms of study quality. Without this information, it is not possible to ascertain the completeness of the review. There is no disclosure of which studies were discarded or why they were discarded. Further, there is no information on why certain studies were emphasized in the discussion. In the case of experimental studies, a similar lack of transparency informed the identification and selection of studies. A recent comment on the failure of IARC monographs to utilize systematic approaches or to cite systematic reviews echoed these same concerns with additional examples (Straif et al., 2012).

In these two examples, the review of epidemiological studies combined cross-sectional, longitudinal, cohort, or secondary analyses without acknowledgement or discussion of heterogeneity, even though it was unlikely that their results could be combined in any meaningful manner. Similarly, the *in vitro* studies were discussed without consideration of study design, dose or *in vitro* concentration, animal strain or cell line. Other sources of heterogeneity were obvious as well. Sometimes studies actually reported on different endpoints. These problems are increased when multiple experimental tests are used to define an endpoint, such as multiple *in vitro* systems and different animal strains (for example, in current US EPA guidelines for developmental neurotoxicity (Crofton et al., 2004) and endocrine disruption (Daston et al., 2003)). When the methods of such studies are so diverse, it may not be appropriate to combine results except in the most general way. Similarly, in EBM/HC studies are not combined if they show either clinical or statistical heterogeneity.

In place of a formal integration of results using clearly described methods (e.g., formal meta-analyses or focused narrative syntheses of the data), these reviews included only tables that summarize selected findings. The only qualified judgments relate to carcinogenicity using EPA or IARC criteria. Even more disturbing than these examples is the practice in some health assessments to base conclusions on only a few or even one study, judged to be the most appropriate or reliable (on nontransparent criteria). Facing two alternative conclusions, one must "choose" which one, if either, to believe. In contrast, a systematic approach uses all the accepted evidence on which to provide a basis for decision making. The concept of a "key study" is contrary to the notion of a systematic review because of its deliberate exclusion of the body of relevant information. This selective practice was followed in a recent NRC review of mercury, in which a nontransparent decision was made to reject one of two large prospective epidemiological studies on early exposures to methyl mercury and neurodevelopmental outcomes (NRC, 2001). Another approach on this same topic utilized a self-described Bayesian "integrative" approach to examine several studies, but no reason was provided for why only some pertinent studies were included (Axelrad et al., 2007). The recent NTP review of lead (2011) moves closer to the practice of systematic reviews as practiced using an evidence-based approach, but it is still a mixture of transparent and nontransparent methods. There are clear statements related to framing

specific questions and to some extent explicating the initial criteria for searching the literature for relevant primary studies, but it fails to present an explicit means by which these studies were identified or evaluated. In addition, as stated in the report, NTP explicitly relied upon other "authoritative sources" (from US government agencies) to identify citations for review, supplemented by some searches of the literature and consultation of experts rather than systematically reviewing all relevant citations. Thus, it is difficult overall to define the methods by which the primary studies were identified or selected, and it is likely to be difficult to replicate the process in an independent exercise. Most importantly, the document does not describe how these study results were integrated to support qualitative judgments based on IARC criteria. Tables in the document are rated as either "supporting" or "not supporting" these qualitative judgments without defining or describing the criteria used to classify a study as supporting or not. Furthermore, the authors appear to have selected which studies are cited in these tables rather than showing all data. Evaluation also involved nontransparent processes such as expert consultation and review by a selected panel. The conclusions were further influenced by the committee review, as well as by the conclusions of the "authoritative sources," which, as noted above, did not adopt or implement transparency.

4 Why EBT and why now

The need for EBT is arguably driven by several forces: the increased demand for transparency and a stronger scientific basis for decision making in both public and private sectors, as well as longstanding dissatisfaction with the pace and contentious nature of current modes of decision making in public health (EEA, 2001). Examples such as the divergent risk assessments for methyl mercury and bisphenol A in public health policy in the US and the EU (Beronius et al., 2010) do not encourage confidence. Stakeholders with an interest in efficient government and public health should be greatly concerned by the fact that EPA's evaluation of the human health effects of dioxins took 18 years. How the data used to make these decisions was obtained is neither clear nor replicable. EBT mandates the provision of methods used to develop a set of primary studies which are then used as the evidence for decision making. Clearly, the use of EBT can promote reduction of controversies, as all can obtain exactly the same data on which to base decisions; the methods used to obtain, assess, and integrate the data are described clearly enough to allow replication. In addition, through increasing the efficiency of decision making, EBT can respond to societal pressure to decrease the resources of time, money, and vertebrate animals utilized in reaching decisions related to hazard and risk (Rovida and Hartung, 2009). These pressures have increased interest in developing alternative methods that reduce the time required to obtain relevant information (NRC, 2007). For this reason, the need to validate these alternative methods adds further impetus to EBT.

5 Initial steps towards systematic reviews in toxicology

We have carried out some of the more detailed studies using principles of EBM/HC to evaluate the evidence for associations between environmental toxicants and human health risks, and this experience provides some perspective on the challenges in adopting and adapting these methods to EBT (Navas-Acien et al., 2005, 2006, 2007; Maull et al., 2012). These reviews follow the norms of transparency and methods that have been developed for systematic reviews of diagnostics and interventions in medicine and health care. They incorporate the following steps: development and explicit framing of research questions that can be answered by a systematic review plus explicit statement of a publically available protocol for conducting the systematic review. This protocol includes a defined and annotated strategy for locating sources of evidence; *a priori* conditions for exclusion and inclusion; defined analytic procedures to evaluate study designs and statistical methods; criteria for evaluating selected studies; methods for integrating study results. These rules are based on the assumption that all studies are well intentioned but no study is perfect. The goal is to identify all relevant sources of information in an unbiased manner and then to screen this body of information by identifying aspects of each study that can increase bias or uncertainty and to consider the impact of these aspects on analytic confidence.

Our attempts to integrate toxicological studies into our reviews were limited in terms of availability of studies, due in most cases to the variability in study design or in the endpoints selected, as well as to differences or lack of precise information on dosing and dose duration, and uncertainty as to the relevance of measured outcomes to the inference of human health risk. Some of these issues relate to toxicology, in which a range of endpoints often are utilized as relevant indicators of human disease risk; this is related to the lack of accepted phenotypic animal (or *in vitro*) models for many human health endpoints and uncertainty as to mechanisms involved in human disease. Lacking a coherent nosology, toxicological studies are likely to be more varied in design and endpoint than epidemiological or clinical studies. Integration of different endpoints may be possible using a systems biology approach to group endpoints in terms of common pathways, but this has not been tested in practice. These concerns also were cited by Maull et al. (2012).

A similar experience is presented in an excellent recent systematic review of formaldehyde and reproductive and development endpoints (Duong et al., 2011). The review of epidemiological studies is a model in transparency and rigor. In contrast (and similar to our reviews on lead and arsenic mentioned above), the review of experimental animal studies was less transparent. No clear information is presented on search terms and criteria for inclusion or exclusion of studies. Large differences were noted among studies in terms of species, routes of exposure and dose, as well as endpoints, which probably impeded any attempt at integration such that only a summary of "key findings" was presented. A thorough narrative discussion of mechanisms and modes of action also was included.

6 Challenges for EBT

The results of our analyses, along with more recent experience from an expert working group convened by the National Institute of Environmental Health Sciences (NIEHS) to evaluate associations between environmental chemicals and diabetes, indicate that toxicologists have considerable work to do to implement an evidence-based approach (Silbergeld, 2009). Innovations and modifications are especially needed to develop evidence-based methods tailored for toxicology and experimental nonhuman studies. Some of the major limitations noted in our reviews are discussed here for human studies and experimental studies. First, the amount of primary information available from independently conducted epidemiological studies in the published literature is relatively sparse for many exposures of interest. Second, many of the available epidemiological studies have significant problems in terms of study design or data reporting such that it is difficult to identify biases in them. For example, in many studies of arsenic, there are limited or no data on individual exposures and many studies failed to collect or report information on important covariates and confounders or information sufficient to determine heterogeneity. Many studies are relatively small and likely underpowered; many of the studies of larger cohorts (such as NHANES) are not actually independent of each other, and none are longitudinal, and so causality cannot be inferred in terms of exposure preceding outcome. In addition, there are broad differences in definition and measurement of outcomes of interest. This is understandable for toxicological studies, but is also characteristic of many epidemiological studies on, for example, lead and arsenic. For the toxicological studies, there is enormous heterogeneity in all aspects of study design and interpretation, as discussed above and in Duong et al. (2011). These criticisms were similar to the evaluations of the medical literature in the early '80s when systematic reviews in EBM/HC were first widely applied and just beginning to be appreciated (Dickersin and Manheimer, 1998).

Nevertheless, our reviews demonstrated that important elements of the methodology of systematic reviews can be adopted by EBT with little change, notably an allegiance to transparency in methods for searching the available literature for potential evidence, in selecting studies for review, and application of *a priori* criteria for assessing each selected study. Toxicologists can examine existing criteria for systematic reviews of observational epidemiology (Blair et al., 1995; AMS, 2007; Longnecker et al., 1988). When appropriate, some of the methods for integrating results across studies also may be adopted. From our analyses, we also observed that the greatest challenges for developing EBT are related to handling information from experimental nonhuman studies, where there is no consensus on analytic procedures and where even the construction of research questions may be more complex owing to the many test systems and endpoints used in studies on the same topic. In addition, there is no consensus on methods for screening primary studies, for evaluating the selected studies, and on appropriate statistical methods to integrate study results from the range of experimental designs. This challenge will not be met by selecting information only from standard toxicology test guidelines or Good Laboratory Practice requirements as the definition of acceptability for evidence-based decisions. Many of these designs are extremely limited and, while they may produce data of use in standard risk assessment methods, they are underpowered and not robust (Reuter et al., 2003). As has been noted in endocrine disruptor research, these types of studies may be less informative than research studies that are more specifically designed to investigate defined hypotheses rather than to generate minimal information on hazard (Myers et al., 2009). Rather, all relevant studies should be sought and then evaluated using methods for appraising sources of biases identified through a consensus process in order to determine the strength of the evidence provided by each. Achieving this goal will foster a closer relationship between environmental epidemiology and experimental research, going beyond the invocation of experimental research merely to satisfy one of Bradford Hill's recommendations.

Achieving the goals of evidence-based and systematic analysis, as argued by practitioners in EBM/HC, has involved two strategies implemented at the beginning: involvement of a broadly based community for achieving consensus in methods and evaluations and a commitment to complete transparency. These commitments are exemplified within the Cochrane Collaboration. At its inception, the Collaboration included only a few dedicated investigators with a shared vision to help people make good health care decisions. This goal drove the development of systematic reviews and the dissemination of these reviews, which now cover a broad range of topics related to health care interventions. Key principles of transparency and continuous improvement in methods based on empirical evidence underlie the growth of the Cochrane Collaboration and its influence in the field of EBM/HC.[3] This paper argues that these strategies, as well as a commitment to continuous growth and improvement in methods, are equally critical for the successful development and adoption of EBT.

The decision for EBT involves a commitment by the field of toxicology, not only to science but to community. As noted above, practitioners in EBM/HC stress that its success has involved the engagement of a broadly based community for consensus evaluations and a commitment to complete transparency. These steps cannot be rushed by establishing structural frameworks and empty institutions but must be grown from an organic discussion among the community of stakeholders, including scientists, technicians, governments, private sector, and the public (Chalmers, 2005).

Our success may transform the field of toxicology, as well as the practice of decision making in regulation. EBT can contribute to the efficient adoption of alternative methods through consensus agreement on identifying the evidence and on criteria for evaluation, drawing on experience from diagnostic evaluations in EBM/HC. However, there must be a commitment to

[3] http://www.cochrane.org/resources/handbook/index.htm

empirically testing the methodology for systematic reviews of toxicological data; without such methodological studies, the field cannot move forward. This will not be a simple task. Since toxicology is fundamentally a science of prevention (Silbergeld et al., 2004), its aim is to detect likely harms prior to human exposure. For this purpose, experimental studies are the only source of truly preventive information, and thus the focus of EBT should be on experimental toxicology and test methods in the broadest sense.

Adoption of an evidence-based approach does not mean the adoption of the clinical trial design as the "highest" or only form of reliable information (Silbergeld, 2009). Evidence may come from any type of study, and although many reviews focus on randomized clinical trials, the type of evidence (i.e., study design) required depends on the type of research question (e.g., the use of randomized controlled clinical trials to answer questions of efficacy and cohort or case-control studies to answer questions related to etiology). This has facilitated the development of both "rules of practice" and the *post hoc* evaluation of research results (Dickersin and Manheimer, 1998). EBM/HC also provides a rich source of valuable guidance to EBT in its methods for evaluating observational epidemiology (Blair et al., 1995; Longnecker et al., 1988; AMS, 2007). While we can learn from EBM/HC, as noted at the outset of this paper, the issues of concern to toxicology, for the most part, are not the same as those in medicine and health care. In EBM/HC, the evidence-based approach has been developed most fully for answering questions related to therapy and diagnosis. The evaluation of novel test methods (such as alternative systems) may draw usefully upon methods used in evaluating diagnostics. Systematic reviews using only evidence from randomized controlled trials (RCTs) are not well suited to identifying harms, primarily due to study designs focused on identifying benefit, often with insufficient power to detect adverse effects because of the relatively low number of individuals exposed and the short time frame of many RCTs (Chou and Helfand, 2005).

The investment of our community in developing EBT will be worthwhile. In the absence of an evidence-based process, decision making is dependent upon a pseudo-Delphi process, in which experts are convened to undertake a qualitative process of integrating and weighing information (e.g., the NTP and IARC). This is less and less satisfactory to the public and other stakeholders; it is also highly resource-intensive in terms of repetitious studies and expert consultation (Rovida and Hartung, 2009). EBT will lead us into new domains of science and assessment, but we should remember that, in identifying harms and assessing risks, as in the law, an evidence-based approach does not remove the need for the application of judgment (Sackett et al., 1996; NRC, 1994). The premise and promise of EBT is the reduction of uncertainties by assuring a consistent body of information and enhancing confidence in the selection and evaluation of this information through a fully transparent process dedicated to continuous improvement through experience. These were the goals that inspired Archie Cochrane and the early community of analysts; by adopting them, the community of toxicologists can enhance the development of science and better serve the social goals of health protection and safety assurance.

References

AMS (2007). *Identifying the Environmental Causes of Disease: How Should We Decide What to Believe and When to Take Action*. London, UK: The Academy of Medical Sciences.

Axelrad, D. A., Bellinger, D. C., Ryan, L. M., et al. (2007). Dose-response relationship of prenatal mercury exposure and IQ: an integrative analysis of epidemiologic data. *Environ Health Perspect 115*, 609-615.

Beronius, A., Rudén, C., Håkansson, H., and Hanberg, A. (2010). Risk to all or none? A comparative analysis of controversies in the health risk assessment of Bisphenol A. *Reprod Toxicol 29*, 132-46.

Blair, A., Burg, J., Foran, J., et al. (1995). Guidelines for application of meta-analysis in environmental epidemiology. ISLI Risk Science Institute. *Regul Toxicol Pharmacol 22*, 189-197.

Chalmers, I. (2005). Academia's failure to support systematic reviews. *Lancet 365*, 469.

Chou, R. and Helfand, M. (2005). Challenges in systematic reviews that assess treatment harms. *Ann Intern Med 142*, 1090-1099.

Crofton, K. M., Makris, S. L., Sette, W. F., et al. (2004). A qualitative retrospective analysis of positive control data in developmental neurotoxicity studies. *Neurotoxicol Teratol 26*, 345-352.

Daston, G. P., Cook, J. C., and Kavlock, R. J. (2003). Uncertainties for endocrine disrupters: our view on progress. *Toxicol Sci 74*, 245-252.

Dickersin, K., Scherer, R., and Lefebvre, C. (1994). Identifying relevant studies for systematic reviews. *BMJ 309*, 1286-1291.

Dickersin, K. and Manheimer, E. (1998). The Cochrane Collaboration: evaluation of health care and services using systematic reviews of the results of randomized controlled trials. *Clin Obstet Gynecol 41*, 315-331.

Duong, A., Steinmaus, C., McHale, C. M., et al. (2011). Reproductive and developmental toxicity of formaldehyde: a systematic review. *Mutat Res 728*, 118-138.

EEA (2001). *Late Lessons from Early Warnings: The Precautionary Principle* 1896-2000. Copenhagen, Denmark: European Environment Agency.

Feinstein, A. R. (1995). Meta-analysis: statistical alchemy for the 21st century. *J Clin Epidemiol 48*, 71-79.

Flower, A., Lewith, G. T., and Little, P. (2007). Seeking an oracle: using the Delphi process to develop practice guidelines for the treatment of endometriosis with Chinese herbal medicine. *J Altern Complement Med 13*, 969-976.

Gatchel, R. J. and McGeary, D. (2002). Cochrane collaboration-based reviews of health-care interventions: are they unequivocal and valid scientifically, or simply nihilistic? *Spine J 2*, 315-319.

Griesinger, C., Hoffmann, S., Kinsner, A., et al. (2009). Special issue: Evidence-based toxicology (EBT). Preface. *Hum Exp Toxicol 28*, 83-86.

Grimshaw, J. M., Santesso, N., Cumpston, M., et al. (2006). Knowledge for knowledge translation: the role of the Cochrane Collaboration. *J Contin Educ Health Prof 26*, 55-62.

Guzelian, P. S., Victoroff, M. S., Halmes, N. C., et al. (2005). Evidence-based toxicology: a comprehensive framework for causation. *Hum Exp Toxicol 24*, 161-201.

Hartung, T. (2010). Evidence-based toxicology – the toolbox of validation for the 21st century? *ALTEX 27*, 253-263.

Hoffmann, S. and Hartung, T. (2006). Toward an evidence-based toxicology. *Hum Exp Toxicol 25*, 497-513.

IARC (2006). *Inorganic and Organic Lead Compounds, IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*. Lyon, France: International Agency for Research on Cancer.

Judson, R., Kavlock, R, Martin, M., et al. (2013). Perspectives on validation of high-throughput assays supporting 21st century toxicity testing. *ALTEX 30*, 51-66.

Longnecker, M. P., Berlin, J. A., Orza, M. J., et al. (1988). A meta-analysis of alcohol consumption in relation to risk of breast cancer. *JAMA 260*, 652-656.

Maull, E. A., Ahsan, H., Edwards, J., et al. (2012). Evaluation of the association between arsenic and diabetes: A National Toxicology Program Workshop Review. *Environ Health Perspect*. http://dx.doi.org/10.1289/ehp.1104579

Moher, D., Jadad, A. R., and Tugwell, P. (1996). Assessing the quality of randomized controlled trials. Current issues and future directions. *Int J Technol Assess Health Care 12*, 195-208.

Mulrow, C. D. (1987). The medical review article: state of the science. *Ann Intern Med 106*, 485-488.

Myers, J. P., vom Saal, F. S., Akingbemi, B. T., et al. (2009). Why public health agencies cannot depend on good laboratory practices as a criterion for selecting data: the case of bisphenol A. *Environ Health Perspect 117*, 309-315.

Navas-Acien, A., Sharrett, A. R., Silbergeld, E. K., et al. (2005). Arsenic exposure and cardiovascular disease: a systematic review of the epidemiologic evidence. *Am J Epidemiol 162*, 1037-1049.

Navas-Acien, A., Silbergeld, E. K., Streeter, R. A., et al. (2006). Arsenic exposure and type 2 diabetes: a systematic review of the experimental and epidemiological evidence. *Environ Health Perspect 114*, 641-648.

Navas-Acien, A., Guallar, E., Silbergeld, E. K., et al. (2007). Lead exposure and cardiovascular disease – a systematic review. *Environ Health Perspect 115*, 472-482.

NRC – National Research Council (1994). *Science and Judgment in Risk Assessment*. Washington, DC: National Academies Press.

NRC (2001). *Arsenic in Drinking Water*. Washington: NAS Press.

NRC (2007). *Toxicity Testing in the 21st Century*. Washington, DC: National Academies Press.

NTP – National Toxicology Program (2011). *NTP Monograph on Health Effects of Low Level Lead*, NTP, NIEHS, USDHHS, draft October 14, 2011. http://ntp.niehs.nih.gov/NTP/ohat/Lead/DraftNTPMonographonHealthEffectsofLowLevelLead.pdf (downloaded April 4, 2012).

Oxman, A. D. and Guyatt, G. H. (1988). Guidelines for reading literature reviews. *CMAJ 138*, 697-703.

Reuter, U., Heinrich-Hirsch, B., Hellwig, J., et al. (2003). Evaluation of OECD screening tests 421 (reproduction/developmental toxicity screening test) and 422 (combined repeated dose toxicity study with the reproduction/developmental toxicity screening test). *Regul Toxicol Pharmacol 38*, 17-26.

Rovida, C. and Hartung, T. (2009). Re-evaluation of animal numbers and costs for in vivo tests to accomplish REACH legislation requirements for chemicals – a report by the transatlantic think tank for toxicology (t^4). *ALTEX 26*, 187-208.

Sackett, D. L., Rosenberg, W. M., Gray, J. A., et al. (1996). Evidence based medicine: what it is and what it isn't. *BMJ 312*, 71-72.

Silbergeld, E. K. (2003). Facilitative mechanisms of lead as a carcinogen. *Mutat Res 533*, 121-133.

Silbergeld, E., Lerman, S., and Hushka, L. (2004). Ethics. Human health research ethics. *Science 305*, 949.

Silbergeld, E. (2009). 4.2 Applying an evidence-based approach: arsenic as a health risk. *Hum Exp Toxicol 28*, 127-129.

Stephens, M., Andersen, M., Becker, R. A., et al. (2013). Evidence-based toxicology for the 21st century: Opportunities and challenges. *ALTEX 30*, 74-103.

Straif, K., Stayner, L., Demers, P. A., et al. (2012). Use of meta-analyses by IARC working groups. *Environ Health Perspect 120*. doi:10.1289/ehp.1205397

Williams, D. D. and Garner, J. (2002). The case against 'the evidence': a different perspective on evidence-based medicine. *Br J Psychiatry 180*, 8-12.

Correspondence to
Ellen Silbergeld, PhD
Department of Environmental Health Sciences
Johns Hopkins School of Public Health
615 N. Wolfe Street, E6644
Baltimore, MD, 21205
USA
e-mail: esilberg@jhsph.edu

Assuring Access to Data for Chemical Evaluations

Lynn R. Goldman[1] *and Ellen K. Silbergeld*[2]

[1]School of Public Health and Health Services, The George Washington University, Washington, DC, USA; [2]Department of Environmental Health Sciences, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

Background: A database for studies used for U.S. Environmental Protection Agency (EPA) pesticide and chemical reviews would be an excellent resource for increasing transparency and improving systematic assessments of pesticides and chemicals. There is increased demand for disclosure of raw data from studies used by the U.S. EPA in these reviews.

Objectives: Because the Information Quality Act (IQA) of 2001 provides an avenue for request of raw data, we reviewed all IQA requests to the U.S. EPA in 2002–2012 and the U.S. EPA's responses. We identified other mechanisms to access such data: public access databases, the Freedom of Information Act (FOIA), and reanalysis by a third party.

Discussion: Only two IQA requests to the U.S. EPA were for raw data. Both of these were fulfilled under FOIA, not the IQA. Barriers to the U.S. EPA's proactive collection of all such data include costs to the U.S. EPA and researchers, significant time burdens for researchers, and major regulatory delays. The U.S. EPA regulatory authority in this area is weak, especially for research conducted in the past, not funded by the U.S. government, and/or conducted abroad. The U.S. EPA is also constrained by industry confidential business information (CBI) claims for regulatory testing data under U.S. chemical and pesticide laws. The National Institutes of Health Clinical Trials database systematically collects statistical data about clinical trials but not raw data; this database may be a model for data from studies of chemicals and pesticides.

Conclusions: A database that registers studies and obtains systematic sets of parameters and results would be more feasible than a system that attempts to make all raw data available proactively. Such a proposal would not obviate rights under the IQA to obtain raw data at a later point.

Key words: access to information; chemicals, hazardous; pesticides; review, systematic. *Environ Health Perspect* 121:149–152 (2013). http://dx.doi.org/10.1289/ehp.1206101 [Online 11 December 2012]

The U.S. Environmental Protection Agency (EPA) is one among many agencies covered by the Information Quality Act (IQA 2001), an amendment to the Treasury and General Government Appropriations Act for fiscal year 2001 that has been viewed as a mechanism to increase access to such information and to seek corrections if parties think that government agencies have used faulty information and analyses. The Office of Management and Budget (OMB) issued IQA guidelines that apply to all agencies in the Executive Branch: When these agencies provide "influential scientific, financial, or statistical information," they also "shall include a high degree of transparency about data and methods to facilitate the reproducibility of such information by qualified third parties" (OMB 2002). The law was enacted without debate or hearing. In the absence of an extensive legislative history and because both the IQA and OMB guidelines were silent about whether agency responses were judicially reviewable, some had viewed the act as providing a new avenue for legal challenges of agency decisions across the U.S. government. For example, in 2006 the U.S. Fourth Circuit Court of Appeals ruled that plaintiffs did not have standing to sue the Department of Health and Human Services under Title III of the IQA to compel access to a study conducted by the National Heart, Lung, and Blood Institute (NHLBI) that was used to support action by the Food and Drug Administration (FDA) on dietary salt (Salt Institute *v.* Leavitt 2006). A number of industry groups had petitioned the NHLBI to make the raw data from the study available so that they could do subgroup reanalyses. The court found that the plaintiffs had received no injury from being denied access to the NHLBI data and thus did not have standing. However, the court also noted that the petitioners had a longstanding right to request the raw data from the study using the Freedom of Information Act (FOIA 1966). In its response, the NHLBI noted that it was preparing a public access data set for release, which it later made available (NHLBI 2005). Although this case was resolved under existing FOIA mechanisms, in the wake of this litigation there has been concern that the IQA does not provide outside parties sufficient access to the data for studies that underlie regulatory decisions made by U.S. government agencies. There is increasing interest in improving the methods by which chemical and pesticide hazards and risks are evaluated not only by government but also by independent scientists (Bucher et al. 2011; Woodruff et al. 2011). This interest has spurred increased demand for transparency and disclosure of the data used by the U.S. EPA to make evaluations that support regulatory decisions for chemicals and pesticides. In this context, we examine the role of the IQA in making such data more accessible and suggest alternative approaches.

Review of Requests for Data

To find out how responsive the U.S. EPA has been to requests for raw data under the IQA, we reviewed 79 requests filed with the U.S. EPA between 2002 and 2012 either to correct or to reconsider the data that the U.S. EPA used in evaluations supporting its regulatory decisions during that period. Under OMB guidance for the IQA (OMB 2002), parties can request that agencies reconsider or correct any information used to support regulatory decisions; usually these requests are made in the form of letters. The U.S. EPA posted these 79 requests on its web site, according to OMB guidelines (U.S. EPA 2012a). Interestingly, only two of these requested raw data.

The first request for raw data was filed in December 2003 by the Perchlorate Study Group, an industry consortium of manufacturers and users of perchlorate (Aerojet, American Pacific Corporation, Kerr-McGee Chemical, and Lockheed Martin). They requested that the U.S. EPA provide raw data from experimental studies (Girard 2003). The U.S. EPA granted this request in September 2004 and provided access to brain images and contractor's reports (Gilman 2003).

The second case was filed by the Association of Battery Recyclers (ABR) in October 2008 (Steinwurtzel 2008). Now called America's Battery Recyclers, and formerly called the Secondary Lead Smelters Association, the ABR is a group of auto and industrial battery recyclers, primary lead producers, and users of recycled lead (America's Battery Recyclers 2012). The ABR requested raw data from a study of lead toxicity (Lanphear et al. 2005) that was among several published studies relied upon by the U.S. EPA in its development of the National Ambient Air Quality Standard (NAAQS) for lead under the Clean Air Act Amendments (1990). Because the ABR and others had taken the U.S. EPA to court to overturn the lead NAAQ rule at the same

Address correspondence to L.R. Goldman, School of Public Health and Health Services, The George Washington University, 2175 K St., NW, Suite 500, Washington, DC 20037 USA. Telephone: (202) 994-5179. E-mail: goldmanL@gwu.edu

The authors declare they have no actual or potential competing financial interests.

Received 4 October 2012; accepted 5 December 2012.

time, the U.S. EPA opted to postpone consideration of the request under the IQA pending the decision of the court. In its response to the request, the U.S. EPA noted that concerns about the data analysis had been noted in comments during the rule-making process and that the U.S. EPA had commissioned new external peer reviews of the study (U.S. EPA 2012a) in addition to a reanalysis of the data of Lanphear et al. (Rothenberg and Rothenberg 2005). After the lead NAAQS was upheld in July 2010, the ABR again requested that the U.S. EPA provide access to the Lanphear data (Steinwurtzel 2010). Meanwhile, litigation was filed over the delay in providing the data. This litigation was dropped when the U.S. EPA FOIA office worked out an agreement with the Cincinnati Children's Medical Center to obtain the Lanphear study data (Lanphear BP, personal communication; Pohl *v.* U.S. EPA et al. 2012). U.S. EPA attorneys determined that access to the data was required under the 1998 Shelby Amendment, which makes federally funded research data accessible to the public under FOIA (Treasury and General Government Appropriations Appropriations Act 1998). Thus, as for the request to the NHLBI to provide data concerning the salt study (Salt Institute *v.* Leavitt 2006), the resolution of the request was managed under FOIA.

Because requests for raw data are few and far between, it has not been onerous for the U.S. EPA to provide such data. Existing mechanisms have provided the ability to reanalyze data by *a*) development and availability of a public-access database (with suitable protections for the human subjects involved in such studies); *b*) provision of raw data via FOIA, for cases in which data are in possession of or can be obtained by the agency (e.g., the perchlorate case cited above); and *c*) reanalysis of data by a third party. As an example of the third mechanism, the widely publicized results from the Harvard Six Cities Study (Dockery et al. 1993) were used by the U.S. EPA in 1997 as a basis for developing new standards for fine particulate matter (≤ 2.5 μm in aerodynamic diameter) air pollution (U.S. EPA 1997). Interested parties, mostly from industry, raised questions about study analysis and interpretation. The raw data were not in the possession of the U.S. EPA, and the U.S. EPA could not compel the submission of these data from Harvard University or the funding source, the American Cancer Society. Under pressure from government agencies and industry, Harvard and the American Cancer Society voluntarily requested that the Health Effects Institute (HEI) step in as a third party to supervise a reanalysis of their data. The HEI [a consortium of industry, academic, and government scientists established by the Clean Air Act Amendments (1990)] provided the data for reanalysis by a third party selected by a science advisory committee that included representation from interested parties who had argued for an independent reanalysis, thus providing a process to address the uncertainties about the analysis and interpretation (HEI 2000).

Discussion

Over time, the U.S. EPA has come to rely increasingly on a large number of scientific studies to complete reviews for a single chemical. This is illustrated by the case of 2,3,7,8-TCDD (2,3,7,8-tetrachlorodibenzo-*p*-dioxin). In its recent assessment of TCDD, the U.S. EPA identified some 2,000 studies directly relevant to its review of dioxin toxicity. From these, the U.S. EPA selected 10 "key" epidemiologic studies and 74 "key" experimental animal studies. Even for this smaller subset of "key" studies, the raw data for each human study and animal experiment are substantial, and most of the data on TCDD were not in the possession of the U.S. EPA (2012b).

There are several mechanisms through which the U.S. EPA might obtain these data. The U.S. EPA could require that investigators submit their raw data to the agency upon completion of their research as a condition of U.S. EPA funding, but this would not completely solve the problem. Most research evaluated by the U.S. EPA for regulatory decision making is not funded by the U.S. EPA. In these cases, the U.S. EPA would have to undertake an extensive collection of raw data from study investigators, which would be costly to the U.S. EPA and burdensome to the research community. Not insignificantly, this would create major delays in rule making. In terms of resource allocation, it is reasonable to ask how much of the U.S. EPA's budget could be allocated to accomplish this, and where this would rank relative to other priorities, such as increasing the numbers of priority assessments to meet the U.S. EPA's statutory goals.

In addition to the burden on the U.S. EPA, there would be a significant burden on the scientific community that produces most of the relevant research, and it is very likely that there would be significant pushback from the academic community under the Paperwork Reduction Act (1995). In fact, the Paperwork Reduction Act, which was enacted to reduce the total amount of paperwork handled by the U.S. government, would not allow the U.S. EPA to undertake such a massive data collection without establishing that the burden imposed upon the research community would be justified by the benefits of providing the data.

At the least, scientists would need funding to respond to requests that are generated as a consequence of the use of their studies by the U.S. EPA rather than any action taken by the investigators themselves. Burdened by other responsibilities and unable to fund such activities from grants provided by sources other than the U.S. EPA, scientists are not likely to voluntarily provide the U.S. EPA with raw data from studies conducted months to decades in the past simply because the U.S. EPA has decided to include those studies in their latest assessment.

Moreover, the U.S. EPA would not have clear legal authority to compel the submission of data from industry, federally funded studies conducted prior to the 1998 Shelby Amendment, studies funded by other federal agencies, or studies that are not funded by the U.S. government, including studies from non-U.S. investigators. We therefore conclude that a regulatory approach, in which the U.S. EPA compels the submission of raw data for all studies reviewed for rule making on pesticides and chemicals, would not be tenable. It could in fact have a chilling effect on the engagement of the global scientific community in research relevant to the protection of human health and the environment. Certainly, this is not in the best interests of science-based policy.

In addition, there are other feasibility issues. In the case of older studies, raw data may not exist or may be difficult to access because of storage on outdated media such as tapes. For epidemiologic studies, consideration would need to be given to ethical issues governing studies of human subjects. These include protection of confidentiality and privacy, and prevention of abuse of the data, for example, by marketing companies who may wish to identify patients with particular medical conditions. Clinical-trials investigators have been working for years to develop ways to disclose data from human studies, including mechanisms for placing data behind a barrier to universal access, so that it is accessible only to those who meet conditions of use. In the case of clinical trials, there are studies in which removal of all identifying data negates its scientific value; therefore access to the data would need to be limited to protect privacy (Hrynaszkiewicz et al. 2010). With adequate resources and planning, these obstacles could be anticipated and/or overcome.

In the case of research data concerning chemicals and pesticides, the U.S. EPA also is constrained by legal constructs that have defined regulatory testing of pesticides as "confidential business information" (CBI) and that require the U.S. EPA to redact certain data and obtain affirmations from recipients that they will not give the remaining data to multinational companies that might seek to register the pesticide to market it in other countries (U.S. EPA Office of Pesticide Programs 2010). The U.S. EPA could improve the web access to summaries and analyses of

these data, which are publically available but often difficult to find in web searches. This would not be the same as providing access to raw data. We therefore suggest that, in the short run, industry should work with the U.S. EPA to identify approaches to provide more robust data sets for studies that they submit to the U.S. EPA. The U.S. EPA also could invite companies to voluntarily waive CBI claims on tests of pesticides and chemicals. In the long run, we think that Congress should amend the Toxic Substances Control Act (1976) and the Federal Insecticide, Fungicide, and Rodenticide Act (1972) as amended by the Food Quality Protection Act (1996) to ease CBI protections from pesticide and chemical test data.

In an ideal world we would always favor more disclosure over less, but it is not clear how this should be done, or who should pay for it. The HEI, which has an independent governing board and is supported by a consortium of funders including the U.S. EPA and the automobile and petrochemical industries, may be a useful precedent. The HEI requires that data from all HEI-funded studies be made available as expeditiously as possible,

> [taking] into consideration the legitimate intellectual interests of the investigator to have the opportunity to benefit from his or her intellectual endeavors and to publish subsequent analyses from the data set (including additional analyses funded by HEI). (HEI 2010)

The HEI attempts to balance the interests of investigators with those of interested parties in cases of "studies of particularly high regulatory importance being used to inform decisions over a short time frame," and encourages its principal investigators to share the data except in situations where "providing the data would place an undue burden on the investigator" (HEI 2010). For example, in cases when there have been so many requests that it was difficult for the investigators to continue their research, the HEI has assisted investigators with data sharing. In addition, the HEI requires that data requesters provide "reasonable reimbursement for both the direct costs of providing the data, and for the time of the investigator and/or HEI staff to gather, transmit, and explicate the data" (HEI 2010). HEI also "will consider requests from the investigator for a reasonable budget of data archiving funds, to be provided as part of the project budget" (HEI 2010). From this precedent, it seems that proponents of increased access to raw data need to consider not only financial and time burdens on investigators, but also a way to reasonably balance the need for data access with the ability of investigators to realize the fruits of their own intellectual endeavors.

Another useful precedent that could serve as a model for data sharing is the National Institutes of Health (NIH) clinical trials database (ClinicalTrials.gov; NIH 2012). It does not contain "raw data" but rather contains detailed and useful information about clinical-trial study designs and statistics that not only convey results in a standardized fashion but also identify important quality parameters (e.g., drop-out rates). Required by law (Section 113 of the Food and Drug Administration Modernization Act 1997), the clinical trials database was developed by the NIH with input from the FDA and the National Library of Medicine (NLM). Currently, many medical journals require that trials be registered in ClinicalTrials.gov prior to their publication; as of 3 December 2012, 136,605 studies in 182 countries were registered. Although many researchers are now are calling for access to raw data for all clinical trials (Gotzsche 2011), the ClinicalTrials.gov database has greatly increased access to information about drug efficacy trials and drug safety, and the development of such a database for studies of chemicals and pesticides would be a major step toward increasing the transparency of the U.S. EPA's evaluations and making data more accessible to third parties.

Conclusions

At present, there does not seem to be a large demand for raw data related to U.S. EPA decision making; however, this may change as formal evidentiary reviews of environmental health research become increasingly common (Maull et al. 2012). Compared with clinical trials, the acquisition of raw data for chemicals and pesticides would be much more complex, in part because it would require a framework that can accommodate data from numerous types of studies: observational and experimental, animal, human, *in vitro*, and high throughput screening studies.

For human epidemiologic studies, clear and complete documentation would need to be provided for interpretation of the variables collected in such studies. This is no simple task given, for example, *a*) the wide range of possible study designs and the intricacies of design of questionnaires and subsequent coding and transformation of variables; *b*) environmental and biomarker sample-collection procedures, chain-of-custody and sample processing and storage, laboratory analyses, data analysis, and coding; and *c*) imputation of missing variables or laboratory nondetects. Although it is a standard practice to carefully document all of these details, there is currently no generally agreed-on manner in which to upload such data into an electronic database. There is a risk that people who were not involved in data collection can misunderstand these details and thus obtain erroneous results. Some effort would be required to develop a standardized system for reporting this kind of information. For experimental animal studies, there should be parameters related to quality assessment (e.g., blinding of investigators, randomization, housing and care of animals).

If the U.S. EPA chose this path, the first step might be to develop a framework similar to ClinicalTrials.gov that would capture statistics and other parameters but would not necessarily require uploading raw data. With adequate funding, involvement of the NLM might provide more sophisticated informatics expertise to make the data more usable, and the NLM or the National Toxicology Program (NTP) could perhaps provide a "home" for the data. The U.S. EPA and other environmental agencies could also require (or request) that investigators register their studies with the database, and journals could require registration as a condition of publication (as some journals currently do for results of clinical trials) or suggest that it be done. Given resource limitations, especially for investigators in developing countries, this step might be difficult for many investigators compared with researchers who perform clinical trials.

A system that provides raw data might be possible if the U.S. EPA could pilot the development of a system that could handle raw data using data already in its possession [e.g., results of its intramural research, results of U.S. EPA-funded extramural research (where available), and any raw data that it has requested from investigators in support of risk-assessment activities]. Other federal agencies, such as the NTP and the National Institute for Occupational Safety and Health could contribute as well. The NTP already publishes all of its data and methods in its reports; however, it does not publish raw data or studies with nonstandard protocols online. In any case, busy investigators may oppose this not only because of the effort and resources required but also because they would be relinquishing exclusive access to their own raw data (and therefore the risk of being "scooped") for the possibility of future requests for reanalysis. Even in cases where investigators contemplate no further data analyses, they may have concerns about the effort to respond to questions about repeat analyses. In any case, additional resources would be required, and this is not a time of plenty for research in the United States or anywhere else. In short, as in all of life, there is no free lunch. We already have mechanisms for disclosure of data used by the U.S. EPA in decision making and even for obtaining raw data. It is doubtful that we can afford the luxury of having this information available for release prior to any request, and it is uncertain who should be responsible for the cost and effort required to provide it.

We conclude that, as is the case for clinical trials, a registry for studies that could handle a wide variety of methodologies and methods of analysis and provide a more complete and

standardized presentation of statistical results and other parameters than is possible in the peer-reviewed literature would be a tremendous resource to society for increasing transparency and improving assessments of pesticides and chemicals. However, at present, there is no evidence that there is a net social benefit to requiring collection of and access to raw data for all studies utilized by the U.S. EPA prior to requests for such data from interested parties. As a first step the U.S. EPA, NTP, and NLM should begin to generate discussions among agencies and with interested outside parties, including academic researchers and the regulated industry, on the possible creation of a reporting system for environmental health studies of chemicals and pesticides that would systematically collect results and data about studies—but not raw data.

References

America's Battery Recyclers. 2012. America's Battery Recyclers Homepage. Available: http://www.americasbatteryrecyclers.com/association.html# [accessed 11 November 2012].

Bucher JR, Thayer K, Birnbaum LS. 2011. The Office of Health Assessment and Translation: a problem-solving resource for the National Toxicology Program [Editorial]. Environ Health Perspect 119:A196–A197.

Clean Air Act Amendments. 1990. Public Law 101–549.

Dockery DW, Pope CA III, Xu X, Spengler JD, Ware JH, Fay ME, et al. 1993. An association between air pollution and mortality in six U.S. cities. N Engl J Med 329(24):1753–1759.

Federal Insecticide, Fungicide, and Rodenticide Act. 1972. 7 USC 136.

Food and Drug Administration Modernization Act. 1997. Public Law 105–115.

Food Quality Protection Act. 1996. Public Law 104–170.

Freedom of Information Act. 1966. Public Law 85–619.

Gilman P. 2003. Letter from P. Gilman, U.S. EPA, to Girard M, Perchlorate Study Group. Re: Request for Correction #13679, 15 September 2003. Available: http://www.epa.gov/quality/informationguidelines/documents/13679-response.pdf [accessed 11 November 2012].

Girard M. 2003. Letter from M. Girard, Perchlorate Study Group, to the U.S. Environmental Protection Agency. Correction Pursuant to Agency-specific Information Quality Guidelines, 3 December 2003. Available: http://www.epa.gov/quality/informationguidelines/documents/13679.pdf [accessed 11 November 2012].

Gotzsche PC. 2011. Why we need easy access to all data from all clinical trials and how to accomplish it. Trials 12:249; doi:10.1186/1745-6215-12-249 [Online 23 November 2011].

HEI (Health Effects Institute). 2000. Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality. Available: http://pubs.healtheffects.org/getfile.php?u=273 [accessed 11 November 2012].

HEI (Health Effects Institute). 2010. Appendix D: Policy on the Provision of Access to Data Underlying HEI-Funded Studies. Available: http://www.healtheffects.org/RFA/RFA-AppendixD.htm [accessed 30 September 2012].

Hrynaszkiewicz I, Norton ML, Vickers AJ, Altman DG. 2010. Preparing raw clinical data for publication: guidance for journal editors, authors, and peer reviewers. Trials 11:9; doi:10.1186/1745-6215-11-9 [Online 29 January 2010].

IQA (Information Quality Act). 2001. Public Law 106–554, Sec. 515. Available: http://www.fws.gov/informationquality/section515.html [accessed 17 December 2012].

Lanphear BP, Hornung R, Khoury J, Yolton K, Baghurst P, Bellinger DC, et al. 2005. Low-level environmental lead exposure and children's intellectual function: an international pooled analysis. Environ Health Perspect 113:894–899.

Maull EA, Ahsan H, Edwards J, Longnecker MP, Navas-Acien A, Pi J, et al. 2012. Evaluation of the association between arsenic and diabetes: a National Toxicology Program workshop review. Environ Health Perspect 120:1658–1670.

NHLBI (National Heart, Lung, and Blood Institute). 2005. Dietary Approaches to Stop Hypertension - Sodium Study (DASH Sodium). Available: https://biolincc.nhlbi.nih.gov/studies/dashsodium/ [accessed 26 September 2012].

NIH (National Institutes of Health). 2012. ClinicalTrials.gov. Available: http://clinicaltrials.gov/ct2/home [accessed 17 December 2012].

OMB (Office of Management and Budget). 2002. Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies. Final Guidelines. Available: http://www.whitehouse.gov/omb/fedreg_reproducible [accessed 17 December 2012].

Paperwork Reduction Act of 1995. 1995. Public Law 104–13.

Pohl *v.* the United States Environmental Protection Agency, et al. 2012. Civil Action No. 09-1480. U.S. District Court for the Western District of Pennsylvania, 6 March 2012.

Rothenberg SJ, Rothenberg JC. 2005. Testing the dose–response specification in epidemiology: public health and policy consequences for lead. Environ Health Perspect 113:1190–1195.

Salt Institute *v.* Leavitt. 2006. Case No. 05-1097. 2006 U.S. App. Lexis 5580. U.S. Court of Appeals for the Fourth Circuit, Richmond, VA, 6 March 2006.

Steinwurtzel RN. 2008. Request for Correction of Information Disseminated by EPA in Support of Its Proposed National Ambient Air Quality Standard for Lead. Available: http://www.epa.gov/quality/informationguidelines/documents/09001.pdf [accessed 11 November 2012].

Steinwurtzel RN. 2010. Request for Correction Number 09001. Available: http://www.epa.gov/quality/informationguidelines/documents/09001-rel4.pdf [accessed 11 November 2012].

Toxic Substances Control Act. 1976. Public Law 94–469.

Treasury and General Government Appropriations Act. 1998. Public Law 105–277.

U.S. EPA (U.S. Environmental Protection Agency). 1997. National Ambient Air Quality Standards for particulate matter. Fed Reg 62(138):38652–38760. Available: http://www.gpo.gov/fdsys/pkg/FR-1997-07-18/pdf/97-18577.pdf [accessed 17 December 2012].

U.S. EPA (U.S. Environmental Protection Agency). 2012a. EPA Information Quality Guidelines–Requests for Correction (RFC) and Requests for Reconsideration (RFR) Submitted to EPA. Available: http://www.epa.gov/quality/informationguidelines/iqg-list.html [accessed 11 November 2012].

U.S. EPA (U.S. Environmental Protection Agency). 2012b. EPA's Reanalysis of Key Issues Related to Dioxin Toxicity and Response to NAS Comments, Volume 1. (CAS No. 1746-01-6). In Support of Summary Information on the Integrated Risk Information System (IRIS). EPA/600/R-10/038F. Available: http://www.epa.gov/iris/supdocs/dioxinv1sup.pdf [accessed 17 December 2012].

U.S. EPA (U.S. Environmental Protection Agency), Office of Pesticide Programs. 2010. FIFRA Information Security Manual. Available: http://www.epa.gov/oamhpod1/oppts_grp/0913541/ism.pdf [accessed 20 December 2012].

Woodruff TJ, Sutton P, The Navigation Guide Work Group. 2011. An evidence-based medicine methodology to bridge the gap between clinical and environmental health sciences. Health Aff (Millwood) 30(5):931–937.

Heather M. Stapleton
Nicholas School of the Environment
Duke University
Durham, North Carolina
E-mail: heather.stapleton@duke.edu

References

Allen JG, McClean MD, Stapleton HM, Webster TF. 2008a. Critical factors in assessing exposure to PBDEs via house dust. Environ Int 34:1085–1091.

Allen JG, McClean MD, Stapleton HM, Webster TF. 2008b. Linking PBDEs in house dust to consumer products using X-ray fluorescence (XRF). Environ Sci Technol 42(11):4222–4228.

Harrad S, Abdallah M, de Wit C, Östman C, Bergh C, Covaci A, et al. 2010. Indoor contamination with hexachlorocyclododecanes, polybrominated diphenyl ethers, and perfluoroalkyl compounds: an important exposure pathway? Environ Sci Technol 4:3221–3231.

Stapleton HM, Kelly SM, Allen JG, McClean MD, Webster TF. 2008. Measurement of polybrominated diphenyl ethers on hand wipes: estimating exposure from hand to mouth contact. Environ Sci Technol 42(9):3329–3334.

U.S. Environmental Protection Agency. 2011. Exposure Factors Handbook: 2011 Edition. EPA/600/R-09/052F. Washington, DC:U.S. Environmental Protection Agency. Available: http://www.epa.gov/ncea/efh/pdfs/efh-complete.pdf [accessed 8 March 2013].

Watkins DJ, McClean MD, Fraser AJ, Weinberg J, Stapleton HM, Sjödin A, et al. 2011. Exposure to PBDEs in the office environment: evaluating the relationships between dust, handwipes, and serum. Environ Health Perspect 119:1247–1252.

Watkins D, McClean M, Fraser A, Weinberg J, Stapleton HM, Sjödin A, et al. 2012. Impact of dust from multiple microenvironments and diet on pentaBDE body burden. Environ Sci Technol 46:1192–1200

Webster TF, Harrad S, Millette JR, Holbrook RD, Davis JM, Stapleton HM, et al. 2009. Identifying transfer mechanisms and sources of decabromodiphenyl ether (BDE 209) in indoor environments using environmental forensic microscopy. Environ Sci Technol 43(9):3067–3072.

Weschler CJ, Nazaroff WW. 2010. SVOC partitioning between the gas phase and settled dust indoors. Atmos Environ 44:3609–3620.

Wu N, Herrmann T, Paepke O, Tickner J, Hale R, Harvey E, et al. 2007. Human exposure to PBDEs: associations of PBDE body burdens with food consumption and house dust concentrations. Environ Sci Technol 41(5):1584–1589.

Access to Chemical Data Used in Regulatory Decision Making

http://dx.doi.org/10.1289/ehp.1206438

It is clear from our commentary (Goldman and Silbergeld 2013), that we disagree with Lutter et al. (2013) about whether the public disclosure of all raw data used by the U.S. Environmental Protection Agency (EPA) for making regulatory decisions for chemicals is necessary to ensure the scientific basis for such decisions, and about the extent to which preemptive disclosure (prior to any request) is practical. However, the most important disagreement between us is the basis asserted by Lutter et al. in their commentary for this change in policy. Lutter et al. argued that it is necessary for the U.S. EPA—and anyone else who desires to do so—to reanalyze all data used in their assessments in order to "replicate" the findings and conclusions of the original investigators.

Lutter et al. (2013) repeatedly used the terms "replicability" and "replication" as synonymous with an "independent analysis" of raw data from an existing study. Replication in science is quite different; it involves performance of an independent study with the same hypothesis and then testing the extent to which this independent study reaches the same conclusions. Recalculation of study statistics or other reanalysis of an existing study data set is not a replication. Designing and conducting a replication study does not require access to raw data from the original study; this would abrogate the concept of independence. Moreover, an independent study will by definition utilize different sets of animal models or human populations, and as a consequence may employ different statistical techniques.

Their second argument is that disclosure of raw data will assist in identifying sources of scientific bias. We consider this unlikely because the most important sources of bias are usually related to problems in study design or limitations of the data collected. This is not identifiable through data recalculation; however, this type of bias can usually be identified in the text of the original study publication.

Lutter et al. (2013) noted (correctly) that applicants to the U.S. EPA for pesticide registrations must provide raw data from regulatory testing as part of the package submitted to the U.S. EPA. This is a very special case, in that these studies are neither peer reviewed nor accessible to the public because of the protection sought by industry and extended by law for confidential business information (CBI). The assumption of bias related to these studies is not unreasonable, given that they are conducted by or on behalf of commercial entities seeking to obtain pesticide registration. These studies are rarely published in the scientific literature or in any way subject to independent peer review other than review by the U.S. EPA. Many scientists and public policy practitioners consider the CBI cloak as a major impediment to transparency and confidence. Industry could demonstrate their commitment to transparency by declining this protection, thereby increasing the confidence of all.

Finally, Lutter et al. (2013) attempted to support their proposal by claiming that journals [*Nature* and the *Proceedings of the National Academy of Sciences of the United States* (*PNAS*)] and an expert body (the Bipartisan Policy Center) agree with them. However, these bodies have neither supported the concept of requiring that all raw data be reported to the the U.S. EPA nor that the U.S. EPA carry out its own independent recalculation. Rather, *Nature* and *PNAS* require authors to agree to make data sets (as well as materials and protocols) available to editors, and to others, upon request (Nature Publishing Group 2012; PNAS 2012). One of us (L.R.G.) was a member of the Science for Policy Project; its final report (Bipartisan Policy Center 2009) also recommended this practice. Many journals require data, such as DNA and protein sequences, macromolecular structures, microarray data, and crystallographic data, to be made available on publicly accessible databases, but most of these are not "raw data" in the sense that Lutter et al. proposed. *Nature* also recommends that authors submit clinical trials data to external clinical trials databases (Nature Publishing Group 2012).

In summary, we disagree with the argument that raw data from every study used by the U.S. EPA to support a regulatory assessment should be made available to the agency and to the public. This proposal does not serve the purpose of "replication" or identification of bias, as asserted by Lutter et al. (2013). In practice, it may generate obstacles to good science and discourage researchers from studying issues of importance in environmental health. This proposal would also limit the U.S. EPA from using the results of research published in the peer-reviewed scientific literature by placing studies off-limits if the authors did not submit raw data sets to the the U.S. EPA.

Finally, there is no obvious need for these changes. When the U.S. EPA has determined a need to reanalyze data, the current regulatory practice has not impeded such activities. Past history indicates that difficult cases are rare and do not warrant an intrusive and burdensome new requirement for the automatic submission of data from all studies.

L.R.G. lists her affiliation for the purpose of identification only.

The authors declare they have no actual or potential competing financial interests.

Lynn R. Goldman
George Washington School of Public Health and Health Services
Washington, DC
E-mail: goldmanl@gwu.edu

Ellen Silbergeld
Department of Environmental Health Sciences
Bloomberg School of Public Health
Johns Hopkins University
Baltimore, Maryland

References

Bipartisan Policy Center. 2009. Science for Policy Project: Improving the Use of Science in Regulatory Policy: Final Report. Washington, DC:Bipartisan Policy Center. Available: http://bipartisanpolicy.org/sites/default/files/BPC%20Science%20Report%20fnl.pdf [accessed 8 March 2013].

Goldman LR, Silbergeld EK. 2013. Assuring access to data for chemical evaluations. Environ Health Perspect 121:149–152.

Lutter R, Barrow C, Borgert CJ, Conrad JW Jr, Edwards D,

Felsot A. 2013. Data disclosure for chemical evaluations. Environ Health Perspect 121:145–148.
Nature Publishing Group. 2012. Availability of Data and Materials. Available: http://www.nature.com/authors/policies/availability.html [accessed 11 December 2012].
PNAS (Proceedings of the National Academy of Sciences of the United States). 2012. Editorial Policies: Materials and Data Availability. Available: http://www.pnas.org/site/authors/journal.xhtml [accessed 8 March 2013].

Access to Chemical Data: Lutter et al. Respond

http://dx.doi.org/10.1289/ehp.1206438R

We appreciate the attention paid by Goldman and Silbergeld (2013) to the issue of data disclosure and agree that there has been "increased demand for transparency and disclosure of the data used by the U.S. EPA [Environmental Protection Agency] to make evaluations that support regulatory decisions."

In their letter, Goldman and Silbergeld contend primarily that "replication" in science means to independently repeat a prior study to see if the same results can be obtained. They suggest that public availability of the prior study's data is unnecessary because a subsequent study will generate its own data. In 2011, a special section of *Science* (Vol. 334, No. 6060) addressed replicability and reproducibility and made two general points. First, "replication," as defined by Goldman and Silbergeld, while perhaps the cornerstone of the scientific method, can be difficult in many settings because of the uniqueness of the precise conditions surrounding field observations, the expense and time required to collect data (e.g., for longitudinal studies), and ethical constraints (e.g., Jasny et al. 2011). Second, in those cases where conduct of a second experiment may be impossible or infeasible, review and reanalysis of the first study's data is still a meaningful step along the "reproducibility spectrum," assists in understanding the differences between competing analyses, and "may be sufficient to verify the quality of the scientific claims" (Peng 2011; see also Ioannides and Khoury 2011; Santer et al. 2011).

Other empirical work also supports the view that data availability promotes reproducibility. In empirical economics, a discipline that uses large-scale statistical models broadly similar to those of epidemiologists, a famous study of replication of peer-reviewed research suggested that inadvertent errors may be "commonplace rather than rare occurrences" (Dewald et al. 1986). The *American Economic Review* (AER 2013) subsequently adopted a policy "to publish papers only if the data used in the analysis are clearly and precisely documented and are readily available to any researcher for purposes of replication." Further, the *AER* conducted a recent evaluation of its policy and reported that about 80% of 39 sampled papers met the spirit of the data availability policy (Glandon 2010). Importantly, independent efforts at replication of 9 selected papers found no serious errors (almost exact replication for 5 studies and "several small discrepancies … immaterial to the conclusions" for another 4.) This result represents a marked improvement relative to the results of the original 1986 study of replication. The difference is presumably attributable, at least in part, to the difference in care and quality of work associated with the *AER*'s current policy of data availability. Although analytic methods underlying papers published in the *AER* are different from those used in chemical evaluation, the experience of the *AER* suggests that there is merit in promoting data availability for the purpose of improving the reliability of the results of published, peer-reviewed scientific papers, at least in disciplines that use complex statistical models.

Finally, we, like Goldman and Silbergeld, "disagree with the argument that raw data from every study used by the U.S. EPA to support a regulatory assessment should be made available to the agency and to the public." Unlike Goldman and Silbergeld, we recommend that the U.S. EPA, when it uses results of a published study in a regulatory assessment, ask the authors for underlying data (Lutter et al. 2013). If the U.S. EPA does not receive such data, it should explain how it used the study results in light of the fact that data sufficient to assess reproducibility was not forthcoming. We believe our approach would facilitate and not obstruct good science and that it would not discourage researchers from studying issues of importance in environmental health. Moreover, it would not, as Goldman and Silbergeld state,

> limit the U.S. EPA from using the results of research published in the peer-reviewed scientific literature by placing studies off-limits if the authors did not submit raw data sets to the U.S. EPA.

R.L., an independent consultant, consults for CropLife America (CLA) and received financial support from the CLA to moderate a forum and serve as principal author of this letter. C.B. consults for Dow AgroSciences LLC, an R&D-based agrochemical producer, registrant, and marketer. C.J.B. received CLA funding to review and analyze scientific literature on data quality. J.W.C. has previously received funding from the American Chemistry Council to author work on the quality of scientific research evaluating chemicals. D.E. consults for a variety of pesticide manufacturers and for the CLA. A.F. has consulted with nonprofit organizations funded by the CLA about pesticide issues.

Randall Lutter
Independent Consultant
Bethesda, Maryland
E-mail: rwlutter@gmail.com

Craig Barrow
Craig Barrow Consulting
Gibsonia, Pennsylvania

Christopher J. Borgert
Applied Pharmacology and Toxicology Inc.
Gainesville, Florida

James W. Conrad Jr.
Conrad Law & Policy Counsel
Washington, DC

Debra Edwards
Independent Consultant
Alexandria, Virginia

Allan Felsot
Food and Environmental Quality Lab
Washington State University
Richland, Washington

References

AER (American Economic Review). 2013. The *American Economic Review*: Data Availability Policy. Available: http://www.aeaweb.org/aer/data.php [accessed 26 February 2013].
Dewald WG, Thursby JG, Anderson RG. 1986. Replication in empirical economics: the *Journal of Money, Credit and Banking* project. Am Econ Rev 76(4):587–603.
Glandon P. 2010. Report on the American Economic Review Data Availability Compliance Project. Available: http://www.aeaweb.org/aer/2011_Data_Compliance_Report.pdf [accessed 11 March 2013].
Goldman LR, Silbergeld EK. 2013. Assuring access to data for chemical evaluations. Environ Health Perspect 121:149–152.
Ioannidis JP, Khoury MJ. 2011. Improving validation practices in "omics" research. Science 334:1230–1232.
Jasny B, Chin G, Chong L, Vignieri S. 2011. Introduction: again, and again, and again …. Science 334:1225.
Lutter R, Barrow C, Borgert CJ, Conrad JW Jr, Edwards D, Felsot A. 2013. Data disclosure for chemical evaluations. Environ Health Perspect 121:145–148.
Peng R. 2011. Reproducible research in computational science. Science 334:1226–1227.
Santer BD, Wigley TML, Taylor KE. 2011. The reproducibility of observational estimates of surface and atmospheric temperature change. Science 334:1232–1233.