Good Laboratory Practices and Safety Assessments
doi:10.1289/ehp.0900884

Having confidence in scientific procedures and data is the sine qua non for determining the safety of chemicals and chemical products. For decisions of safety, there must be rigorous and thorough application of fundamental scientific practices, irrespective of the purpose of the study and where it is conducted—academic, industry, or a contract laboratory.

Investigations must be designed and conducted by experts; whenever possible, standardized and validated test methods and test systems should be used, test devices and instruments must be appropriately calibrated and their accuracy assured, and, most important, all of the data, including raw laboratory records, should be available for independent review. Good Laboratory Practice (GLP) requirements, based on these fundamental scientific principles and practices, are indispensable for providing scientific confidence in studies conducted for chemical safety determinations. These reasons explain why government agencies worldwide require GLP compliance, and why it is entirely appropriate for greater weight to be given to GLP studies than non-GLP studies that are only available as articles in scientific journals. In their commentary Myers et al. (2009) argued that noncompliance with GLP should not be used as the sole criterion for excluding studies from consideration in regulatory decision-making. We agree that GLP should not be the sole criterion, but we strenuously disagree with the authors’ mischaracterization of the purpose and function of GLP and with their conclusion that GLP has no utility for weighting the reliability of studies.

Evaluating the safety of any substance should include review of all relevant studies utilizing a systematic weight-of-evidence framework. Although not all studies that are useful for hazard characterization and risk assessment may be amenable to GLP (e.g., epidemiology and mechanistic studies, studies conducted before the acceptance of current GLP), this does not obviate their consideration. Each study, GLP and non-GLP, should be evaluated and weighed in accordance with fundamental scientific principles. Factors to be evaluated include a) verification of measurement methods and data; b) control of experimental variables that could affect measurements; c) corroborate among studies; d) power (both statistical and biological); e) universality of the effects in validated test systems using relevant animal strains and appropriate routes of exposure; f) biological plausibility of results; and g) uniformity among substances with similar attributes and effects. Regulatory agencies [Food and Drug Administration (FDA) and U. S. Environmental Protection Agency (EPA)] and the National Toxicology Program (NTP) require studies to be conducted in accordance with GLP (FDA 2005; NTP 2006; U.S. EPA 2007a, 2007b), and the Organisation for Economic Co-operation and Development (OECD) GLP principles (OECD 1998) apply to all OECD member countries.

Academic basic research is very different from regulatory research and testing. Academic research focuses on developing and evaluating new hypotheses, on creating novel methods, and on discovering new findings. Academic research is open to wide interpretation and may require significant additional studies to clarify and determine whether and how broadly the results apply. Although novel techniques and discoveries of academic investigations stimulate further research, they must also stand up to the scientific method: hypothesis formulation, hypothesis testing, and validation by independent replication. Independent replication provides critical information on the strength of the hypothesis and reliability of test methods. Inconsistent results can arise from use of novel techniques, different test systems, uncertainty and differences in test chemical composition and purity, and a myriad of other factors. These facts, in conjunction with the more limited availability of actual data in most journal publications, means regulatory agencies can face significant challenges in confirming the quality, performance, or data integrity of results obtained solely from information available from a typical article in peer-reviewed journals. Whereas all study records and data from GLP investigations are available to agencies, rarely, if ever, are such details made available as part of the peer-review process for publishing a manuscript in a scientific journal. This can limit the ability of an agency to independently evaluate conclusions or to conduct alternative analyses of the data. The challenges faced by the peer-review procedures of journals have been recently highlighted (Nature 2006), and it has been pointed out that “…scientists understand that peer review per se provides only a minimal assurance of quality, and that the public conception of peer review as a stamp of authentication is far from the truth” (Jennings 2006). Journal peer review relies on summarization of experimental procedures and results, and does not include examination of laboratory study records or raw data. The purpose for journal peer review is to judge whether the study has been conducted and reported according to internationally recognized, general scientific standards and whether the study meets the interest level for dissemination to scientific community. It is not designed to provide assurance of accuracy or to recalculate raw data, and it does not provide an opportunity for independent audit of the study. Myers et al. (2009) failed to clearly make these distinctions.

Relevant internationally agreed test methods are used by industry to generate toxicity data for safety determinations by regulatory agencies. Incorporation of GLP in these laboratory tests assures that written protocols and standard operating procedures for each study component are developed and carefully and completely followed. GLP also requires meticulous adherence to dosing techniques; the use of adequate group sizes to allow meaningful statistical analysis; characterization (identity, purity, concentration) of test and control substances, including dosing solutions; detailed recording of study measurements and data; and collection of all raw laboratory data in a manner that can be retained and made available for regulatory agencies to audit and reach independent conclusions. Quality control procedures, quality assurance reviews, and facility inspections are also used to monitor and enforce GLP compliance. The relevance, reliability, sensitivity, and specificity of most test methods required by industry by regulatory agencies are well understood because they have been subjected to extensive, round-robin validation programs conducted in numerous laboratories throughout the world. This high level of scientific rigor, in conjunction with the detailed processes of GLP, provides regulatory agencies increased confidence in both the relevance and quality of GLP scientific studies for safety decisions, and it is the reason it is wholly appropriate in regulatory decision making for greater weight and confidence to be afforded to studies conducted in accordance with GLP.

This letter has been reviewed in accordance with the peer- and administrative-review policies of the authors’ organizations. The views expressed here are those of the authors and do not necessarily reflect the opinions and/or policies of their employers.

The authors are employed by trade associations whose members manufacture and use chemicals.

Richard A. Becker
American Chemistry Council
Arlington, Virginia
E-mail: rick_becker@americanchemistry.com
Good Laboratory Practices: Myers et al. Respond
doi:10.1289/ehp.0900884R

We are in complete agreement with the statement by Becker et al. that “having confidence in scientific procedures and data is the sine qua non for determining the safety of chemicals and chemical products.” Our aim in writing the commentary (Myers et al. 2009) was not to challenge the original intent of Good Laboratory Practices (GLP) requirements, which was to establish standards of record keeping in contract laboratory research so as to reduce the likelihood of fraud. Our goal instead was to show—through an analysis of the application of GLP data on bisphenol A (BPA) in regulatory proceedings—that GLP by itself is insufficient to guarantee valid and reliable science. Becker et al. appear to have missed the point of our commentary entirely.

In the case of BPA, three GLP studies have been offered by industry-sponsored laboratories as proof of the chemical’s safety (Cagen et al. 1999; Tyl et al. 2002, 2008). Each has errors in study design and/or data interpretation that are sufficiently serious as to invalidate the conclusions of these studies (Myers et al. 2009). Nevertheless, because the studies were conducted using GLP guidelines, they were judged by regulators as being more reliable than the many National Institutes of Health (NIH)-funded and peer-reviewed studies that have reported adverse effects (Richter et al. 2007; vom Saal et al. 2007).

As our commentary (Myers et al. 2009) clearly establishes, GLP did not guarantee the scientific validity of these three studies. Because previous analyses had identified serious flaws in the first two of those GLP studies, we focused critical attention on the most recent (Tyl et al. 2008), which both the European Food Safety Authority (EFSA 2006) and the U.S. Food and Drug Administration (FDA) had identified as key in their BPA risk assessments (FDA 2008). We found three main flaws: a) the animals were inexplicably insensitive to estrogen; b) the assays were outdated and insensitive compared with methods used in NIH-funded research showing adverse effects; and c) validity of the findings was challenged. For example, the prostate weights of control animals reported by Tyl et al. (2008) were >70% larger (mean, >72 mg) than those reported by numerous laboratories, including a previously published study using CD-1 mice (conducted at RTI, where the study by Tyl et al. 2008) was conducted) that reported mean prostate weights of 46 mg in CD-1 males that were examined at a similar age (Heindel et al. 1995).

Since we published our commentary (Myers et al. 2009), a possible contributor to both the estrogen insensitivity and the enlarged control prostates has been suggested: Approximately 3 years before the experiments that formed the basis of the study by Tyl et al. (2008), there was a polycarbonate fire that released BPA into the RTI laboratory where the research was conducted (Kissinger and Rust 2009). An investigation revealed that animals in the laboratory were exposed to low doses of BPA that government-funded science (Richter et al. 2007) indicates could affect research animals.

Additional uncertainties about Tyl et al.’s study (Tyl et al. 2008) have now been identified by the lead author. Whereas the published paper reports that the animals were examined at approximately 14 weeks of age, Tyl testified at an FDA hearing in September 2008 that they were 6 months of age, and then at a German Environmental Protection Agency hearing in March 2009 that they were 5 months of age (Kissinger and Rust 2009). There she confirmed that the information in the original article was inaccurate. Because an animal’s physiology changes as it ages, these contradictory statements are problematic for all reported outcomes; even at 5–6 months of age, normal, healthy CD-1 male mice would not have the grossly enlarged prostates reported by Tyl et al. (2008).

The use of flawed science, however, is not the only concern. The type of multigeneration testing approach used in these studies is, quite simply, insufficient for the testing of endocrine-disrupting chemicals. This is not a new concept. The need for more specific tests for endocrine-active compounds led in 1998 to the establishment at the U.S. Environmental Protection Agency (U.S. EPA) of the Endocrine Disruptor Screening Program, mandated by Congress (U.S. EPA 1998). After virtually no progress for over a decade, in 2009 the U.S. EPA finally announced a set of testing procedures that will be examined. The proposed “new” methodology, heavily dependent upon traditional toxicologic methods used in multigenerational GLP studies, is still woefully inadequate (Colborn 2009).

The letter by Becker et al. provides a striking example of the reluctance of industry lobbyists to hear this message. In the eyes of the 36 scientific colleagues who coauthored our commentary (Myers et al. 2009), the BPA studies that Becker et al. attempt to defend are so seriously flawed as to be indefensible. Rather than continue to defend a dead issue, we encourage industry representatives to come into the 21st century and help us devise new paradigms for testing endocrine-disrupting chemicals that will safeguard human health.

The authors’ freedom to design, conduct, interpret, and publish this letter was not nor is it compromised by any controlling sponsor as a condition of review and publication. J.P. Myers is CEO/chief scientist for Environmental Health Sciences (EHS), a not-for-profit organization that receives support from several private foundations (listed at http://www.environmentalhealthnews.org/about.html) to support EHS’ mission to advance public understanding of environmental health sciences; no grants to EHS were received to support the writing of this letter. T. Colborn is the president of TEDX (The Endocrine Disruption Exchange), a not-for-profit organization receiving funding from several private foundations (listed at http://www.endocrinedisruption.com/support.php) in support of their mission to educate people about endocrine disruption and other toxic chemicals, and to assist them in their efforts to reduce the production, use, and exposure to chemicals that can interfere with development and function; TEDX also receives individual contributions solicited via the website. S. Jobling is the principal of Beyond the Basics.
Correspondence Limited, a consultancy company that manages scientific projects and advises on scientific research. F. vom Saal is CEO of Xenoanalytical LLC, a small private laboratory that performs assays of xenobiotic compounds.

John Peterson Myers Environmental Health Sciences Charlottesville, Virginia E-mail: jpmyers@ehsic.org

Frederick S. vom Saal Julia A. Taylor Division of Biological Science University of Missouri Columbia, Missouri

Benson T. Akingbemi Department of Anatomy, Physiology & Pharmacology College of Veterinary Medicine Auburn University Auburn, Alabama

Koji Arizono Faculty of Environmental and Symbiotic Science Prefectural University of Kumamoto Tsukide, Kumamoto, Japan

Scott Belcher Department of Pharmacology & Cell Biophysics Center for Environmental Genetics University of Cincinnati Cincinnati, Ohio

Theo Colborn The Endocrine Disruption Exchange Paonia, Colorado

Ibrahim Chahoud Institut für Klinische Pharmakologie und Toxikologie Charité—Universitätsmedizin Berlin Campus Benjamin Franklin Berlin, Germany

All 37 authors of the original commentary signed this letter but only the first 7 are listed here.

REFERENCES

Cagen SZ, Waechter JM Jr, Dimond SS, Bremslin WJ, Butala JH, Jekat PW, et al. 1999. Normal reproductive organ development in CF-1 mice following prenatal exposure to bisphenol A. Toxicol Sci 50:36–44.

Colborn T. 2009. EPA’s New Pesticide Testing Is Outdated, Crude. Available: http://www.environmentalhealthnews.org/ehi/editorial/epa2009a-new-pesticide-testing-is-outdated-crude [accessed 27 April 2009].

EFSA (European Food Safety Authority). 2006. Opinion of the Scientific Panel on Food Additives, Flavourings, Processing Aids and Materials in Contact with Food on a Request from the Commission Related to 2,2-Bis(4-hydroxyphenyl) propane (Bisphenol A). EFSA J 4:20–1–6.

FDA. 2008. Draft Assessment of Bisphenol A for Use in Food Contact Applications. Available: http://www.fda.gov/ohrms/dockets/ac/08/briefing/2008-3038b/1_01_02_FDA%20PA%20Draft%20Assessment.pdf [accessed 14 August 2008].

Heindel JJ, Chapin RE, George J, Gulati DK, Fial PA, Barnes LH, et al. 1995. Assessment of the reproductive toxicity of a complex mixture of 25 groundwater contaminants in mice and rats. Fundam Appl Toxicol 25:19–29.

Kissingler M, Rust S. 2009. Consortium rejects FDA claim of BPA’s safety. Milwaukee Journal Sentinel, 11 April. Available: http://www.journal.com/watchdog/watchdogreports/42858807 [accessed 11 April 2009].

Myers JP, vom Saal FS, Akingbemi BT, Arizono K, Belcher S, Colborn 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.

Richter CA, Birnbaum LS, Farabollini F, Newbold RR, Rubin BS, Talsness CE, et al. 2007. In vivo effects of bisphenol A in laboratory rodents studies. Reprod Toxicol 24(2):199–224.

Tyl RW, Myers CB, Marr MC, Sloan CS, Castillo N, Veselica MM, et al. 2008. Two-generation reproductive toxicity study of dietary bisphenol A in CD-1 (Swiss) mice. Toxicol Sci 104:362–384.

Tyl RW, Myers CB, Marr MC, Thomas BF, Keimowitz AR, Brine DI, et al. 2002. Three-generation reproductive toxicity study of dietary bisphenol A in CD Sprague-Dawley rats. Toxicol Sci 68(1):121–146.

U.S. EPA (U.S. Environmental Protection Agency). 1998. Endocrine Disruption Screening Program (EDSP). Available: http://epa.gov/edsp/index.htm [accessed 3 May 2009].

Von Saal FS, Akingbemi BT, Belcher SM, Birnbaum LS, Crain DA, Eriksen M, et al. 2007. Chapel Hill bisphenol A expert panel consensus statement: integration of mechanisms, effects in animals and potential to impact human health at current levels of exposure. Reprod Toxicol 24(2):131–138.

Electromagnetic Fields and the Precautionary Principle doi:10.1289/ehp.0901111

Since Galileo, debates in science are supported by logical reasoning and reference to statements of fact and not by reference to “authorities.” Consequently, literature serves or should serve two purposes: to give credit to thoughts expressed earlier by others, and to refer to statements of facts. The article by Dolan and Rowley (2009)—employees of the mobile telephone industry—is an example of a compilation of points of views expressed by authorities. No number of references to authoritative statements can replace scientific discourse. The article can be summarized as follows: There is no convincing evidence of harm from exposure to microwaves below levels recommended by the International Commission on Non-Ionizing Radiation Protection (ICNIRP) (1998); therefore, there is no harm, and hence application of the precautionary principle is not indicated.

Indeed, the precautionary principle is not intended as a response to unfounded fears of the public or to aim at zero risk, but as a risk management strategy in case of scientific uncertainty about the existence or magnitude of a risk. Apparently Dolan and Rowley (2009) are not aware that their subjective reasoning does not differ from the unfounded reasoning does not differ from the unfounded risk of a guideline level for power frequency fields and increase of body temperature for high-frequency fields, not because there are no other effects, even at levels far below the guideline levels derived from these acute effects, but because the panels came to the consensus that these other effects cannot (yet) form the basis for the derivation of guidelines. For example, the International Agency for Research on Cancer (IARC 2002) classified power frequency magnetic fields as a possible human carcinogen. In that case, the subjectivity of the assessment is fully transparent: The basic rules of IARC were violated, as the panel questioned whether epidemiologic evidence can be causally interpreted in spite of evidence that neither bias nor confounding accounts for the increased childhood leukemia risk. The exposure level for which there is evidence of an increased childhood leukemia risk is far below the international standards, but the panels setting the standard did not use this evidence as a basis for the derivation of a guideline level for power frequency fields. There are surely many arguments for this decision. However, none are scientific. This is not meant as a reproach, because we recognize the fact that guidelines cannot be derived from scientific statements alone.

Refraining to the World Health Organization (WHO 2000), Dolan and Rowley (2009) stated: “The corresponding advice to governments is to adopt science based guidelines and not to undermine confidence by incorporating additional arbitrary safety factors.” The expression “science-based guidelines,” if taken literally, is a contradiction in terms. Although public health guidelines should be based on a thorough risk assessment, neither the assessment itself nor the reasoning that is applied to derive a guideline can be scientific. No scientific evidence can define a margin of safety; no scientific evidence can replace the value judgment of which evidence to rely on, which evidence to dismiss, and so forth. Safety factors are always—at least to certain degree—arbitrary. For example, we very rarely have scientific evidence about the distribution of sensitivity to a toxic agent in the population; therefore, we apply arbitrary factors for taking interindividual differences into account. What is important, and nearly always neglected in the area of electromagnetic fields (EMF), is to clearly state where value judgments and arbitrary decisions entered the argument and the derivation of guidelines.

The international standards for EMF (ICNIRP 1998; IEEE 2006) are based on immediate effects of exposure, such as excitation of nerve or muscle cells for low-frequency fields and increase of body temperature for high-frequency fields, not because there are no other effects, even at levels far below the guideline levels derived from these acute effects, but because the panels came to the consensus that these other effects cannot (yet) form the basis for the derivation of guidelines. For example, the International Agency for Research on Cancer (IARC 2002) classified power frequency magnetic fields as a possible human carcinogen. In that case, the subjectivity of the assessment is fully transparent: The basic rules of IARC were violated, as the panel questioned whether epidemiologic evidence can be causally interpreted in spite of evidence that neither bias nor confounding accounts for the increased childhood leukemia risk. The exposure level for which there is evidence of an increased childhood leukemia risk is far below the international standards, but the panels setting the standard did not use this evidence as a basis for the derivation of a guideline level for power frequency fields. There are surely many arguments for this decision. However, none are scientific. This is not meant as a reproach, because we recognize the fact that guidelines cannot be derived from scientific statements alone.

It would be much more appropriate if Dolan and Rowley expressly stated that they are completely satisfied with the international standards and that the industry does not want to be bothered by allusions to precaution.
EMF and the Precautionary Principle: Dolan and Rowley Respond
doi:10.1289/ehp.0901111R

In their letter commenting on our article (Dolan and Rowley 2009), Kundi et al. attempt to discredit us as “employees of the mobile telephone industry” and imply that we are merely advocating an industry position of support for the international radiofrequency exposure guidelines rather than addressing issues of substance regarding the application of the precautionary principle to mobile telephony. A careful reading of our article clearly demonstrates this is not the case.

With few exceptions, countries around the world have implemented the international guidelines based on many scientific reviews undertaken by experts appointed by national governments over the past decade. Kundi et al. simply ignore these reviews and their conclusions because they do not agree with them. Although Kundi et al. are entitled to hold and promulgate their own views, they should acknowledge that they are acting as advocates for lower guidelines (based on their own subjective analysis of existing scientific evidence) and they should not simply dismiss anyone who does not agree with their point of view.

Kundi et al. seem to think that the precautionary principle should be applied whenever there is some scientific doubt or uncertainty, without recognizing that its use is limited by national regulatory and legal constraints, which we addressed in our commentary (Dolan and Rowley 2009). In particular, the concerns should be plausible (World Commission on the Ethics of Scientific Knowledge and Technology 2005). As an example, the European Court of First Instance (Pfizer Animal Health SA v. Council of the European Union 2002) has ruled that [A] preventive measure cannot properly be based on a purely hypothetical approach to the risk, founded on mere conjecture which has not been scientifically verified . . . . [A] preventive measure may be taken only if the risk, although the reality and extent thereof have not been fully demonstrated by conclusive scientific evidence, appears nevertheless to be adequately backed up by the scientific data available at the time when the measure was taken.

The many scientific review bodies we referred to in our article (Dolan and Rowley 2009) have not considered the existing health data on mobile telephony adequate to trigger the application of the precautionary principle. We accept that decisions regarding application of the precautionary principle are not to be made by scientists alone because they “have neither democratic legitimacy nor political responsibilities” (Pfizer Animal Health SA v. Council of the European Union 2002). However, governments should not simply disregard scientific advice or adopt populist policies and “fall prey to public fear when it is baseless” (Telstra Corporation Limited v. Hornsby Shire Council 2006).

In the conclusion of our commentary (Dolan and Rowley 2009), we made it clear that there are many things that governments and industry can do to better address public concern, including supporting ongoing research and conducting education and information programs for the public who, when fully informed, are better able to take their own personal precautionary measures if they wish to do so. What should be avoided is the rush to adopt measures—justified by reference to the precautionary principle—to reassure the public, because this has been shown to actually increase public concern (Weidemann and Schütz 2005).

The views expressed in this paper are those of the authors and do not necessarily represent the views of any organizations or companies with which they are professionally associated. Their freedom to design, conduct, interpret, and publish research is not compromised by any controlling sponsor as a condition of review and publication.

Both authors are employed by trade associations representing the mobile communications industry.

Mike Dolan
Mobile Operators Association
London, United Kingdom
E-mail: mikedolan@ukmoa.org

Jack Rowley
GSM Association
London, United Kingdom
E-mail: jrowley@gsm.org

References

Dolan M, Rowley J. 2009. The precautionary principle in the context of mobile phone and base station radiofrequency exposures. Environ Health Perspect 117:1329–1332; doi:10.1289/ehp.0900727 [Online 18 May 2009].

International Commission on Non-Ionizing Radiation Protection. 1998. Guidelines for limiting exposure to time-varying electric, magnetic, and electromagnetic fields (up to 300 GHz). Health Phys 74(4):494–522.

IEEE. 2006. IEEE Standard for Safety Levels with Respect to Human Exposure to Radio Frequency Electromagnetic Fields, 3 kHz to 300 GHz. IEEE Std C95.1-2005. New York:IEEE.

WHO (World Health Organization). 2000. Electromagnetic Fields and Public Health: Mobile Telephones and Their Base Stations. Fact Sheet No. 193. Available: http://www.who.int/mediacentre/factsheets/fs193/en [accessed 2 June 2009].