Why Monsanto's attempt to "disappear" tumours by using historical control data is invalid

An Earth Open Source briefing
26 September 2012

Summary

Monsanto has invoked "historical norms" to dismiss statistically significant findings of increased tumours and mortality rates in rats fed GM maize NK603, as well as in rats exposed to levels of Roundup claimed by regulators to be safe, in a 2-year study by Professor Gilles-Eric Seralini's research team in France.

Monsanto says that the increased mortality rates and tumour incidence "fall within historical norms for this strain of laboratory rats, which is known for a high incidence of tumours".

By "historical norms" and "within this historical range", Monsanto means historical control data – data from various other studies that they find in the scientific literature or elsewhere.

However, the use of historical control data is an unscientific strategy used by industry and some regulators to dismiss statistically significant findings of toxicity in treated (exposed) groups of laboratory animals in toxicological studies intended to evaluate safety of pesticides, chemicals, and GMOs.

The only scientifically valid control for such experiments is the concurrent control, not historical control data. This is because scientific experiments are designed to reduce variables to a minimum. The concurrent control group achieves this because it consists of animals treated identically to the experimental group, except that they are not exposed to the substance under study. Thus, the only variable is exposure to the substance(s) being tested – in the case of Seralini's experiments, NK603 maize and Roundup.

With this experimental design, any differences seen in the treated animals are very likely to be due to the substance being tested, rather than due to irrelevant factors, as is the case with historical control data.

Historical control data consists of a wide range of data gathered from diverse experiments performed under widely differing conditions. As a result, factors totally irrelevant to the study are responsible for the majority of differences in historical control data. Such factors may include environmental conditions; different diet for the animals; different pesticide residue exposures; different genetic background of the animals; even different years in which the experiments were performed, which is known to affect results for reasons that are poorly understood.

In contrast, using the concurrent controls reduces such variables to a minimum and enables researchers to reach evidence-based conclusions about the effects of the substance being tested.

For these reasons, toxicological studies performed by independent (non-industry) scientists and published in the peer-reviewed literature hardly ever invoke historical control data. They certainly do not use it to dismiss statistically significant findings of harm in treated groups of animals.

Those who do use historical control data in this way include industry-affiliated sources and some regulators. The practice was introduced into risk assessment by the Organisation for Economic Cooperation and Development. Even the OECD, however, advises caution in the use of such data and warns against it in evaluating findings of tumours (such as those seen in Seralini's study). Even by weak OECD standards, Seralini's findings are valid.

Even if we were to follow Monsanto's recommendation and use historical control data in evaluating Seralini's findings, the historical control data cited by Monsanto is invalid because it relates to rats of a different origin (SD rats from Charles River Labs) than Seralini's rats (SD rats from Harlan). Seralini took historical data on the Harlan SD rat fully into account in his study – and the results still show that the tumour increase and other effects were statistically significant. The tumour incidence in the test groups in his study was overall around three times higher than the normal rate observed in the Harlan SD rat strain he used, as reported in the literature.

Finally, the "tumour-prone rat" argument used by Monsanto and others to dismiss Seralini's findings of increased tumours is spurious. The key point about Seralini's tumour findings was that the controls got some tumours, but the treated groups got significantly more tumours, and these appeared sooner and were more aggressive than those of the control groups.

So what matters in this study is not whether the rat strain was or was not prone to develop tumours, but
(a) the earlier and more rapid rate of tumour development in the treated groups of animals, compared with the concurrent control, and
(b) the larger number of tumours caused by the treatment, over and above the "spontaneous" background level.

To illustrate the point by analogy: a small proportion of people who never smoke get lung cancer. If you smoke, your risk of getting lung cancer is about 12 times higher than if you don't smoke. The measurement is called a "relative risk". So even if there were an ethnic group of people with a higher rate of naturally occurring lung cancer, if people in that group smoke, their rate of lung cancer will still increase like that of everyone else.

This is a basic principle of science and it is worrying that attempts are being made by pro-GM lobbyists to override it in the interests of keeping the products of powerful multinational biotechnology companies on the market.

The responsibility now lies with Monsanto to pay for a full 2-year carcinogenicity study on its NK603 maize and the associated pesticide, Roundup. Such a study would, however, have to be carried out, not by industry or its contracted labs, but by independent scientists commissioned by an impartial publicly funded body consisting of a wide range of stakeholders representing the public interest.

In the meantime, NK603 must be immediately withdrawn from the market and all GMOs must be subjected to long-term testing.

Introduction

Monsanto has invoked "historical norms" to dismiss statistically significant findings of increased tumours and mortality rates in rats fed GM maize NK603, as well as in rats exposed to levels of Roundup claimed by regulators to be safe, in a 2-year study by Professor Gilles-Eric Seralini's research team in France.[1]

Monsanto says:

> *"Mortality rates and tumor incidence in all groups fall within historical norms for this strain of laboratory rats, which is known for a high incidence of tumors… The percent survival to study termination in SD rats (Charles River Laboratories) ranges from 17-62.9% in males and from 20-62% in females. Findings by Seralini are within this historical range."*[2]

What does Monsanto mean by "historical norms" and "within this historical range"? The answer: historical control data.

While the term may seem technical, citizens need to understand what historical control data is and how it's often misused by industry and some regulators to "disappear" findings of toxicity such as those found in Seralini's study. This enables risky products to reach, and remain on, the market.

For example, in 2011 Earth Open Source uncovered massive abuse of historical control data by German government agencies to "disappear" increases in malformations/birth defects in laboratory animals in industry's own tests on glyphosate, the main chemical ingredient of Roundup herbicide.[3]

How is historical control data used?

This is how the "historical control data" scam is practised:

1. Statistically significant findings of toxicity are found in the treated group of animals (those exposed to the substance being tested, such as a chemical or a genetically modified food), when compared with the concurrent control group. The concurrent control group consists of animals treated identically to the experimental group except that they were not exposed to the substance under study. For a pesticide company or GMO developer, such a statistically significant finding is inconvenient, as it means that the substance may be banned or restricted.
2. Industry and/or regulators decide to compare the treated group not to the control group within the experiment (the concurrent control group), but to data from various other studies that they find in the scientific literature or elsewhere. Because this "historical control data" has been gathered from a diversity of studies carried out

[1] Séralini, G. E., E. Clair, et al. (2012). Long term toxicity of a Roundup herbicide and a Roundup-tolerant genetically modified maize. Food and Chemical Toxicology.
[2] Monsanto comments: Long term toxicity of a Roundup herbicide and a Roundup-tolerant genetically modified maize. September. http://www.monsanto.com/products/Documents/ProductSafety/seralini-sept-2012-monsanto-comments.pdf
[3] Antoniou, M., M. Habib, et al. (2011). Roundup and birth defects: Is the public being kept in the dark?, Earth Open Source. http://bit.ly/IP2FWH

under a wide range of conditions, the data will have a wide range of values. This variability is due not to variation related to the effects of the substance under study, but to a wide range of irrelevant factors.

3. The statistical measures of data variability determined for the historical control data are, unsurprisingly, very large, while the measures of data variability for the concurrent control are much smaller. When the data from the experimental group are compared with that from the concurrent control group, variability of both groups will be small, and a clear difference between the experimental and control groups will be observed, while when the data from the experimental group is compared with the historical control data, the large variability in the historical control data makes it appear that the experimental group is not significantly different from the control.
4. The statistically significant findings of toxicity in the treated group of animals disappears in the “noise” of irrelevant variables introduced by the historical control data.
5. Based on this spurious comparison, the substance is declared safe enough to market.

What’s wrong with using historical control data?

Using historical control data runs counter to good scientific practice as established over centuries. One scientist told us that use of historical control data was “rewriting the rules of science”.

The only valid control in any experiment is the concurrent control. This is because scientific experiments are designed to minimize variables. Within the experiment, the only difference between the treated group and the control group should be that the treated group is exposed to the substance being tested and the control is not. This is the single variable.

All other conditions must be the same between the treated and control group – for example:

- Strain and origin of animal, so that all animals have similar genetic makeup
- Diet, apart from the addition of the substance being tested in the treatment group’s diet
- Environmental conditions in the lab: exposures to pests, pathogens, stressors, etc. should be equal for the experimental and concurrent control groups.
- Substance tested should be from the same batch, same formulation, storage conditions, etc.

Using the concurrent control ensures that any significant changes seen in the treated group are almost certainly due to the substance being tested and not due to some other factor.

What do other scientists say?

Several published papers confirm that the concurrent control groups are the most valid or the only valid control group and warn against the biasing effect of including historical control data.

These include Haseman (1984),[4] Hardisty (1985),[5] and Cuffe (2011). Cuffe cites an example in which using such data can lead to Type II errors.[6] In the case of a toxicology study on a chemical or GMO, Type II errors would mean false negatives, in which a real finding of harm goes unnoticed and the substance is falsely claimed to produce no effect.

Some of the variables that can creep into the dataset when using historical control data include:

- Different strain, origin, or batch of animal with different genetic makeup.
- Different formulations of the substance being tested. For example, with pesticides, formulation ingredients vary widely, as explained elsewhere in this report.
- Different impurities in the substance being tested.
- Different diet. The standard feed (chow) given to experimental animals can vary from year to year, between manufacturers, countries, labs, etc. The ingredients and pesticide residues will be different from batch to batch – sometimes dramatically so.
- Different storage conditions for the tested substance. Chemical reactions may take place within formulated products between times of manufacture and use, depending on the ingredients and conditions they are kept in.
- Different pathogens in the environment in which the animals are kept.
- Different years in which the experiments were performed. Interestingly, this has been identified as the major problem with using historical control data. For reasons that are not understood, the incidence of certain endpoints, e.g. birth defects, tumours etc., is cyclical – that is, in some years there is a higher incidence of that particular effect than in other years.[7]
- Different laboratory, again, for reasons that are not understood,[8] but that probably relate to changes in one or more of the above types of factors.

The resulting wide range of effects resulting from all these different variables mean that any statistically significant effects from the treatment will be lost in the “noise” of irrelevant data.

With concurrent controls, these variables are kept to a minimum. So statistically significant changes in the treated group as compared with the control group are almost certainly due to the substance being tested.

[4] Haseman JK. Statistical issues in the design, analysis and interpretation of animal carcinogenicity studies. Environmental Health Perspectives. 1984; 58: 385–392.
[5] Hardisty JF. Factors influencing laboratory animal spontaneous tumor profiles. Toxicol Pathol. 1985; 13(95–104).
[6] Cuffe RL. The inclusion of historical control data may reduce the power of a confirmatory study. Stat Med. Mar 22 2011.
[7] Haseman JK. Statistical issues in the design, analysis and interpretation of animal carcinogenicity studies. Environmental Health Perspectives. 1984; 58: 385–392.
[8] Haseman JK. Statistical issues in the design, analysis and interpretation of animal carcinogenicity studies. Environmental Health Perspectives. 1984; 58: 385–392

How do we know when historical control data is being used?

People who use historical control data to "disappear" inconvenient findings of toxicity do not always call it that. Alternative terms include "within the normal range of biological variation", "within the historic range", or "within the range expected for this species/strain".

Is it ever OK to use historical control data?

Haseman says there are a few rare instances where historical control data can be useful, such as in cases where effects seen are borderline and show only a marginal increase over the concurrent controls, or in the case of rare tumours, where there is a genuine shortage of concurrent control data because the effect is seldom seen.

Even then, he stipulates that extreme care must be taken to ensure that all sources of variability in the historical control data are identified[9] – an exercise that we believe is not followed by industry and regulatory bodies that use such data in risk assessments.

In contrast, most chronic toxicity studies conducted by independent (non-industry) scientists use only concurrent controls.

Those who use historical control data in issues that affect the public interest must be required to provide the complete data set, along with an analysis of all the potential variables that they have controlled for in selecting/eliminating data from their historical dataset. Such an in-depth analysis is essential, before historical control data can be accepted as valid.

How did historical control data come to be used in risk assessments?

The responsibility for allowing historical control data to be used in risk assessments lies with the Organisation for Economic Cooperation and Development (OECD), which introduced the concept into its guidelines for industry tests performed for regulatory purposes.[10]

However, even the OECD guidelines encourage the use of concurrent controls in chronic (long-term) toxicity tests and give strict conditions for the use of historical control data. It says such data should only be used:

> *"provided that the data chosen are from studies that are comparable with the study being investigated. It is widely recognized that large differences can result from disparities in factors such as pathology nomenclature, strain, husbandry, pathologists."*

In reality, we have never seen evidence in a risk assessment that the conditions of the various experiments from which historical control data were drawn were "comparable".

---

[9] Haseman JK (1984) Statistical issues in the design, analysis and interpretation of animal carcinogenicity studies. Environmental Health Perspectives 58: 385–392.

[10] OECD (2011) Guidance document 116: On the conduct and design of chronic toxicity and carcinogenicity studies, supporting test guidelines 451, 452 & 453 – 2nd ed. http://www.oecd.org/env/chemicalsafetyandbiosafety/testingofchemicals/seriesontestingandassessmenttestingforhumanhealth.htm

Typically the risk assessment contains vague phrases such as, “the [toxicity] overlapped the normal range of historical controls”, without establishing the relevance of the data used. In many cases, the normal range is not specified and no peer-reviewed published evidence is cited to show that this is truly the normal range.

Typically also, the historical control data is unpublished. So we cannot check that the data is valid. Whether and how often the risk assessors see such data is questionable.

The “tumour-prone rats” argument is spurious

Monsanto’s attempt to dismiss the tumour findings on the claimed basis that the type of rat used is prone to tumours is spurious. The key point about Seralini's tumour findings was that the controls got some tumours, but the treated groups got significantly more tumours, which began sooner and were more aggressive, than those of the control groups.

In the context of the controlled experiment, what matters is not the number of “spontaneous” tumours in the control group, but the additional number of tumours caused by the treatment, over and above the "spontaneous" background level.

To illustrate the point by an analogy: a small proportion of people who never smoke get lung cancer. If you smoke, the risk of getting lung cancer is about 12 times higher than if you don't smoke. The measurement is called a "relative risk". So even if there were an ethnic group of people with a higher rate of naturally occurring lung cancer, we know that if people in that group smoke, their rate of lung cancer will still increase like everybody else.

This is a basic principle of science and it is worrying that attempts are being made by pro-GM lobbyists to override it in the interests of keeping the products of biotechnology companies on the market.

Conclusions on historical control data and Seralini’s findings

The OECD’s allowance of historical control data in risk assessments has done a disservice to science and to public health.

Nevertheless, with regard to Seralini’s findings that NK603 maize and Roundup increased tumour incidence, even the OECD guidelines on the use of historical control data favour Seralini’s interpretation of the data (that a treatment-related increase exists), not Monsanto’s (that no treatment-related increase exists).

The OECD says:

> *“It should be stressed that the concurrent control group is always the most important consideration in the testing for increased tumour rates.”*

In Seralini’s study, comparison of the experimental group to the concurrent controls clearly shows that GM maize NK603 and Roundup increase tumour rates, especially in the female animals. Even the weak OECD guidelines prioritize concurrent controls in assessing increased tumour rates. So applying the OECD guidelines supports the conclusion that GM maize NK603 and Roundup increase tumour rates.

We conclude that the OECD guidelines do not support Monsanto's attempt to "disappear" the increased tumour rates through use of historical control data.

Let's for the sake of argument assume for a moment that Monsanto is correct and we should use historical control data in interpreting Seralini's findings. Monsanto said:

> *"The percent survival to study termination in SD [Sprague-Dawley] rats (Charles River Laboratories) ranges from 17-62.9% in males and from 20-62% in females. Findings by Seralini are within this historical range."*

But Seralini got his SD rats from Harlan, not Charles River Labs. This is just one of many variables that indicates that the Monsanto historical control data is invalid as a comparator with the Seralini data.

Seralini made specific reference in his paper to the Harlan SD strain and took into account peer-reviewed publications that have evaluated tumour rates in these animals. These data supported the conclusion that tumours significantly increased in rats fed NK603 maize and rats exposed to Roundup. The tumour incidence in the test groups in his study was overall around three times higher than the normal rate observed in the Harlan SD rat strain he used, as reported in the literature (Brix et al, 2005), including in the largest study with 1329 SD female rats (Chandra et al, 1992).

Thus the most valid historical control data supports Seralini's findings.

From the point of view of rigorous scientific principles, however, the only valid controls that must be used to assess Seralini's results are the concurrent controls. These clearly show statistically significant increased tumour rates in treated animals.

The responsibility now lies with Monsanto to pay for a full 2-year carcinogenicity study on its NK603 maize and the associated pesticide, Roundup. Such a study would, however, have to be carried out not by industry or its contracted labs but by independent scientists commissioned by a publicly funded body consisting of a wide range of stakeholders representing the public interest.

In the meantime, NK603 must be immediately withdrawn from the market and all GMOs must be subjected to long-term testing.